

XLIX Simpozijum o operacionim istraživanjima  
XLIX International Symposium on Operational Research



# SYM-OP-IS 2022

Vrnjačka Banja, 19-22. septembar 2022.

## ZBORNIK RADOVA PROCEEDINGS

Urednici / Editors:

**prof. dr Zorica Mladenović**

**dr Mladen Stamenković**



UNIVERZITET U BEOGRADU  
Ekonomski fakultet  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business

*Izdavač / Publisher*

**Univerzitet u Beogradu – Ekonomski fakultet**

**Centar za izdavačku delatnost**

Kamenička 6, Beograd

<http://cid.ekof.bg.ac.rs/>

[cid@ekof.bg.ac.rs](mailto:cid@ekof.bg.ac.rs)

The Publishing Centre of the Faculty of Economics in Belgrade

University of Belgrade - Faculty of Economics and Business

Kamenička 6, Belgrade, Serbia

*Za izdavača / In behalf of publisher*

**prof. dr Žaklina Stojanović**

*Urednici / Editors*

**prof. dr Zorica Mladenović**

**dr Mladen Stamenković**

*Štampa / Printed by*

**JP „Službeni glasnik”, Beograd**

[www.slglasnik.com](http://www.slglasnik.com)

*Tiraž / No. of copies*

**50**

*Godina izdavanja / Publishing year*

**2022**

**ISBN: 978-86-403-1750-4**

## **IZVRŠNI ORGANIZATOR / EXECUTIVE ORGANIZERS**

Univerzitet u Beogradu – Ekonomski fakultet

University of Belgrade – Faculty of Economics and Business

## **SUORGANIZATORI / CO-ORGANIZERS**

Rudarsko-geološki fakultet, Beograd

Ministarstvo odbrane Republike Srbije

Vojska Republike Srbije

Matematički institut, Beograd

Univerzitet u Beogradu – Matematički fakultet

Visoka građevinsko-geodetska škola, Beograd

Ekonomski institut, Beograd

Društvo operacionih istraživača Srbije

Univerzitet u Beogradu – Fakultet organizacionih nauka

Univerzitet u Beogradu – Saobraćajni fakultet

Institut Mihajlo Pupin, Beograd

Fakultet tehničkih nauka, Novi Sad

Univerzitet u Banjoj Luci

## PROGRAMSKI ODBOR / PROGRAMME COMMITTEE

Zorica Mladenović, EF, Beograd, predsednik

Milan Dražić, MF, Beograd

Dragan Urošević, MI SANU, Beograd

Vule Aleksić, VGGŠ, Beograd

Mirjana Čangalović, FON, Beograd

Goran Ćirović, FTN, Novi Sad

Ivan Nikolić, EI, Beograd

Tatjana Davidović, MI SANU, Beograd

Branka Dimitrijević, SF, Beograd

Rade Doroslovački, FTN, Novi Sad

Boban Đorović, UO, Beograd

Đorđe Dugošija, UNP, Novi Pazar

Radoje Banković, VGI, Beograd

Irena Janković, EF, Beograd

Jelena Kočović, EF, Beograd

Nenad Mladenović, MI SANU, Beograd

Vera Kovačević-Vujčić, FON, Beograd

Jozef Kratica, MI SANU, Beograd

Đuro Kutlača, IMP, Beograd

Marija Kuzmanović, FON, Beograd

Duško Letić, TFZR, Zrenjanin

Dragana Makajić-Nikolić, FON, Beograd

Milan Martić, FON, Beograd

Igor Miljanović, RGF, Beograd

Snežana Mitrović, VGGŠ, Beograd

Feđa Netjasov, SF, Beograd

Dragan Nikolić, VGGŠ, Beograd

Dalibor Petrović, MO RS

Nataša Petrović, FON, Beograd

Slavica Petrović, EF, Kragujevac

Željko Praštalo, RGF, Beograd

Ljubiša Preradović, AGGF, Banja Luka

Dragan Radojević, IMP, Beograd

Aleksandar Savić, MF, Beograd

Gordana Savić, FON, Beograd

Milica Šelmić, SF, Beograd

Stanko Stanić, EF, Banja Luka

Zorica Stanimirović, MF, Beograd

Milan Stanojević, FON, Beograd

Milorad Stanojević, SF, Beograd

Dušan Starčević, FON, Beograd

Dejan Stojković, MO RS

Milija Suknović, FON, Beograd

Milorad Vidović, SF, Beograd

Mirko Vujošević, FON, Beograd

Katarina Vukadinović, SF, Beograd

Daniel Aloise, Polytechnique Montreal, Canada

Abdelhakim Artiba, Université Polytechnique  
Hauts-de-France, France

Emilio Carrizosa, University of Sevilla, Spain

Abraham Duarte, Universidad Rey Juan Carlos,  
Spain

Anton Eremeev, Omsk Branch of Sobolev  
Institute of Mathematics, Russia

Laureano Escudero, Universidad Rey Juan  
Carlos, Spain

Saïd Hanafi, INSA Hauts-de-France, France

Miroslav Hudec, University of Economics,  
Bratislava, Slovakia

Yuri Kochetov, Sobolev Institute of Mathematics,  
Russia

Martine Labbé, Université Libre de Bruxelles,  
Belgium

Belén Melián Batista, University La Laguna,  
Spain

Athanasios Migdalas, Luleå University of  
Technology, Sweden

José A. Moreno-Perez, University of La Laguna,  
Spain

Markos Papageorgiou, Technical University of  
Crete, Greece

Günther Raidl, Vienna University of Technology,  
Austria

Said Salhi, University of Kent, United Kingdom

Marc Sevaux, Université de Bretagne-Sud, France

Angelo Sifaleras, University of Macedonia, Greece

Kenneth Sörensen, University of Antwerp,  
Belgium

## POČASNI PROGRAMSKI ODBOR / HONORARY PROGRAMME COMMITTEE

Marko Andrejić, VA, Beograd

Marko Backović, EF, Beograd

Vladan Batanović, IMP, Beograd

Siniša Borović, FMMS, Beograd

Obrad Čabarkapa, MO RS

Dragoš Cvetković, SANU, Beograd

Branislav Đorđević, GF, Beograd

Slobodan Guberinić, IMP, Beograd

Aleksandar Ilić, MO RS

Mitar Kovač, MO RS

Slobodan Krčevinac, FON, Beograd

Katalin Mesaroš, EF, Subotica

Miodrag Mihaljević, MI SANU, Beograd

Gradimir Milovanović, SANU, Niš

Spasoje Mučibabić, MO RS

Ilija Nikolić, FGM, Beograd

Endre Pap, Univerzitet Singidunum, Beograd

Zoran Ognjanović, MI SANU, Beograd

Serafim Opricović, GF, Beograd

Radivoj Petrović, IMP, AINS, Beograd

Milan Rakić, IMP, Beograd

Slobodan Vujić, RI, Beograd

Svetozar Vukadinović, SF, Beograd

Jovo Vuleta, EF, Beograd

Tomislav Zečević, EF, Beograd

## ORGANIZACIONI ODBOR / ORGANIZING COMMITTEE

Dragan Azdejković, predsednik, Univerzitet u Beogradu – Ekonomski fakultet, predsednik

Irena Janković, Univerzitet u Beogradu – Ekonomski fakultet

Đorđe Stakić, Univerzitet u Beogradu – Ekonomski fakultet

Emilija Maksimović, Univerzitet u Beogradu – Ekonomski fakultet

## NOSIOCI POVELJE ZA ZASLUGE U RAZVOJU OPERACIONIH ISTRAŽIVANJA / ACKNOWLEDGEMENT FOR MERITS IN DEVELOPMENT OF OPERATIONAL RESEARCH

Božidar Kraut (1983)

Alojzij Vadnal (1983)

Dragoslav Marković (1983)

Branislav Ivanović (1984)

Ljubomir Martić (1984)

Radivoj Petrović (1984)

Jovan Petrić (1988)

Sanjo Zlobec (1990)

Radoslav Stanojević (1991)

Svetozar Vukadinović (1993)

Slobodan Krčevinac (1993)

Slobodan Guberinić (1993)

Jovo Vuleta (1993)

Vera Kovačević-Vučić (1998)

Dušan Teodorović (1998)

Vlastimir Matejić (2000)

Mirko Vujošević (2000)

Siniša Borović (2001)

Tomislav Zečević (2001)

Slobodan Vujić (2003)

Dragan Radojević (2006)

Mirjana Čangalović (2010)

Nenad Mladenović (2010)

Spasoje Mučibabić (2010)

Milan Martić (2012)

Dragoš Cvetković (2013)

Đorđe Dugošija (2017)

Goran Ćirović (2017)

Milorad Vidović (2019)

Dragan Urošević (2021)



## PREDGOVOR

Simpozijum o operacionim istraživanjima (Sym-Op-Is) održava se 2022. godine po 49. put i to u Vrnjačkoj Banji od 19. do 22. septembra. Glavni organizator 49. Sym-Op-Is-a je Ekonomski fakultet Univerziteta u Beogradu, koji 2022. godine obeležava značajan jubilej – 85 godina od osnivanja.

Zbornik radova 49. Sym-Op-Is-a sadrži 91 rad i 23 apstrakta sa ukupno 265 koautora. Većina autora radova je iz Srbije. Inostrani istraživači su iz sledećih zemalja: Bangladeš, Bosna i Hercegovina, Italija, Izrael, Kina, Rusija, SAD, Severna Makedonija, Slovenija i Španija. Radovi su grupisani u 23 tematske sekcije.

U izboru i pripremi radova Zbornika učestovalo je 39 recenzenata (od toga 27 članova Programskog odbora Sym-Op-Is-a). Dugujemo svima iskrenu zahvalnost za posvećenost recenzentskom procesu.

Na 49. Sym-Op-Is-u biće održana dva plenarna predavanja po pozivu. Predavači su cenjeni profesori visoke naučne reputacije oblasti višekriterijumskog odlučivanja: dr Salvatore Corrente, sa Univerziteta u Kataniji i dr Miłosz Kadziński, sa Tehnološkog Univerziteta u Poznanju. Takođe, predavanje po pozivu u okviru sekcije Matematičko programiranje održaće ugledni profesor oblasti matematičke logike, dr Mirjana Ilić sa Ekonomskog fakulteta Univerziteta u Beogradu.

Nadamo se da će ovogodišnji Sym-Op-Is, kao i ranijih godina, biti centar intenzivne i korisne razmene naučnih ideja operacionih istraživača.

Urednici  
Zorica Mladenović  
Mladen Stamenković

## PREFACE

The Symposium on Operational Research (Sym-Op-Is) will be held for the 49th time in Vrnjačka Banja from September 19 to 22, 2022. The main organiser of the 49th Sym-Op-Is is the Faculty of Economics and Business, University of Belgrade, which in 2022 is celebrating a significant anniversary – 85 years since its foundation.

The Proceedings of the 49th Sym-Op-Is contain 91 papers and 23 abstracts, with 265 coauthors overall. Most of the authors are from Serbia, while foreign researchers are from the following countries: Bangladesh, Bosnia and Herzegovina, Italy, Israel, China, Russia, USA, North Macedonia, Slovenia, and Spain. The papers are grouped into 23 thematic sections.

Overall, 39 reviewers participated in the selection and preparation of the Proceedings (including 27 members of the Program Committee of Sym-Op-Is). We owe everyone sincere gratitude for their dedication to the review process.

The 49th Sym-Op-Is will include two plenary lectures by invited lecturers. The lecturers are respected professors of high scientific reputation in the field of multiple criteria decision aiding: Dr. Salvatore Corrente from the University of Catania and Dr. Miłosz Kadziński from the Poznań University of Technology. Also, a lecture by invitation within the Mathematical Programming section will be held by a respected professor of mathematical logic, Dr. Mirjana Ilić, from the Faculty of Economics and Business, University of Belgrade.

As in previous years, we hope that this year's Sym-Op-Is will be a centre for an intensive and useful exchange of scientific ideas among operational researchers.

Editors  
Zorica Mladenović  
Mladen Stamenković





## SADRŽAJ / TABLE OF CONTENTS

<b>PLENARNA PREDAVANJA / PLENARY SESSIONS</b>	<b>1</b>
INTERACTIVE EVOLUTIONARY MULTIOBJECTIVE OPTIMIZATION GUIDED BY PREFERENCE INFORMATION .....	3
<i>Salvatore Corrente</i>	
RECOMMENDING MULTIPLE CRITERIA DECISION ANALYSIS METHODS WITH A NEW TAXONOMY-BASED DECISION SUPPORT SYSTEM .....	5
<i>Miłosz Kadziński</i>	
<b>ANALIZA PERFORMANSI / PERFORMANCE ANALYSIS</b>	<b>7</b>
ANALIZA EFIKASNOSTI KOMPANIJA U STICANJU KONKURENTSKE PREDNOSTI PRIMENOM DEA METODE .....	9
<i>Ivona Jovanović, Milan Radojičić, Gordana Savić</i>	
ANALIZA INDEKSA PERFORMANSI PROCESA U KOMPANIJAMA U SRBIJI .....	15
<i>Dragana Stojanović, Ivan Tomašević, Barbara Simeunović, Dragoslav Slović</i>	
ANALIZA POKAZATELJA PERFORMANSI UPRAVLJANJA PRIHODIMA U RESTORANIMA .....	21
<i>Branislava Hristov Stančić, Bojan Zečević, Igor Kovačević</i>	
ODREĐIVANJE NAJVAŽNIJIH KRITERIJUMA PRILIKOM SELEKCIJE KANDIDATA U IT INDUSTRIJI PRIMENOM MACBETH METODE .....	27
<i>Tijana Nanuševski, Anđela Mrdak, Bisera Andrić Gušavac, Milena Popović</i>	
PRODUCTIVITY OF LEADING R&D ENTERPRISES IN THE WORLD-LEADING ECONOMIES: A META-FRONTIER APPROACH .....	33
<i>Aleksander Aristovnik, Guoliang Yang, Yaoyao Song, Dejan Ravšelj</i>	
<b>EKOLOŠKI MENADŽMENT I UPRAVLJANJE PRIRODNIM RESURSIMA / ENVIRONMENTAL MANAGEMENT AND NATURAL RESOURCES MANAGEMENT</b>	<b>35</b>
GOING GREEN IN DIGITAL BANKING: A CASE STUDY .....	37
<i>Jovan Trtica, Jelena Andreja Radaković</i>	
PRISTUPI, TEHNIKE I METODE OPERACIONIH ISTRAŽIVANJA U SMANJIVANJU RIZIKA OD KATASTROFA .....	43
<i>Jelena Andreja Radaković, Dragana Makajić-Nikolić</i>	
<b>EKONOMSKI MODELI I EKONOMETRIJA / ECONOMIC MODELS AND ECONOMETRICS</b>	<b>49</b>
BAYESIAN APPROACH IN ANALYSIS OF THE EFFECTS OF REGULATORY FOCUS ON CONSUMER DECISION-MAKING .....	51
<i>Lazar Čolić, Ivana Prica</i>	

DETERMINANTE STOPE VIŠKA SMRTNOSTI PO ZEMLJAMA U PERIODU PANDEMIJE COVID-19. ....	57
<i>Aleksandra Nojković, Miloško Arsić, Emilija Maksimović</i>	
FISKALNI I TEKUĆI BILANS U ZEMLJAMA CIE: PRIMENA DINAMIČKOG PANEL MODELA. ....	63
<i>Jelena Rašković</i>	
GDP GROWTH DETERMINANTS IN SELECTED POST-TRANSITION EUROPEAN COUNTRIES. ....	69
<i>Milutin Ješić</i>	
OCENA PERZISTENTNOSTI INDEKSA NEIZVESNOSTI EKONOMSKE POLITIKE: NEKI EMPIRIJSKI REZULTATI. ....	75
<i>Zorica Mladenović</i>	
RANJIVOST MLADIH U SRBIJI. ....	81
<i>Gorana Krstić, Aleksandra Anić</i>	
STATISTICAL PROPERTIES OF THE LIMIT ORDER BOOK OF THE NASDAQ STOCK MARKET. ....	87
<i>Dragana Radojičić, Nina Radojičić, Željko Jović</i>	
THE LINK BETWEEN TECHNICAL AND TECHNOLOGICAL PROGRESS, THE DEVELOPMENT OF NEW DIGITAL SERVICES FOR PUBLIC SECTOR GOVERNANCE AND ECONOMIC DEVELOPMENT. ....	93
<i>Petra Vujković</i>	
ULOGA OPERACIONIH ISTRAŽIVANJA U EKONOMIJI: BIBLIOMETRIJSKA ANALIZA SRPSKIH ISTRAŽIVAČKIH REZULTATA. ....	95
<i>Mladen Stamenković</i>	

## **ELEKTRONSKO POSLOVANJE / ELECTRONIC BUSINESS 101**

ELEKTRONSKO POSLOVANJE U ŽELEZNIČKOM SAOBRAĆAJU. ....	103
<i>Zoran Pavlović</i>	

## **FINANSIJE I BANKARSTVO / FINANCE AND BANKING 109**

ANALIZA FINANSIRANJA IZDATAKA UNIVERZITETA ODBRANE ZA NABAVKU ARTIKALA ISHRANE. ....	111
<i>Milena Knezević, Aleksandra Mitić</i>	
CAN A DYNAMIC CORRELATION FACTOR IMPROVE THE PRICING OF INDUSTRY PORTFOLIOS?. ....	117
<i>Miloš Božović</i>	
EFEKAT UKLJUČIVANJA INFLACIJE PRI OBRAČUNU REZERVACIJA ZA ŠTETE. ....	119
<i>Jelena Kočović, Marija Koprivica, Željko Jović</i>	
EFFECTS OF APPLICATION OF DIGITAL TECHNOLOGIES IN INSURANCE. ....	125
<i>Nevena Petrović</i>	
HETEROGENEOUS EXPECTATIONS OF INVESTORS AND STOCK PRICES. ....	131
<i>Dragana Draganac</i>	

IDENTIFIKACIJA I METODOLOGIJA PROCENE NEMATERIJALNIH ULAGANJA U BANKAMA . . . . .	137
<i>Dejan Kokanović, Ognjen Vasiljević, Nina Milenković</i>	
KRATKOROČNE FLUKTUACIJE REALNOG EFEKTIVNOG DEVIZNOG KURSA ZEMALJA ZAPADNOG BALKANA KAO INDIKATOR DEVIZNOG RIZIKA U FINANSIRANJU MEĐUNARODNOG POSLOVANJA . . . . .	139
<i>Aleksandra Đorđević, Ivana Popović Petrović</i>	
OPTIMIZATION OF THE PORTFOLIO OF PRIME LISTING SHARES OF THE BELGRADE STOCK EXCHANGE . . . . .	145
<i>Zoran Popović, Aleksandar Božović</i>	
TIPOVI PROMENA KONCENTRACIJE U BANKOVNOM SEKTORU SRBIJE: DEKOMPOZICIJA INDEKSA HIRŠMANA-HERFINDALA . . . . .	147
<i>Rajko Bukvić</i>	
TRETMAN INFLACIJE U MODELU DISKONTOVANOG NOVČANOG TOKA . . . . .	153
<i>Nina Milenković, Dejan Kokanović</i>	
ZELENE OBVEZNICE U EU – ANALIZA DETERMINANTI PRINOSA . . . . .	155
<i>Irena Jankovic, Vladimir Vasić, Jelena Basarić</i>	

## **GEOINFORMACIONI SISTEMI / GEOINFORMATION SYSTEMS 161**

CARTOGRAPHIC GENERALIZATION OF SETTLEMENTS ON TOPOGRAPHIC MAPS . . . . .	163
<i>Marko Simić, Marko Stojanović, Radoje Banković, Jasmina Jovanović, Tanja Janković, Marija Stojanović</i>	
KORIŠĆENJE MAŠINSKOG UČENJA ZA PROCENU VREDOSTI STANOVA U GIS OKRUŽENJU . . . . .	169
<i>Stanislava Bosiočić, Zoran Srdić, Dragan Savić</i>	
MODEL UPOTREBE PRENOSNIH TABLET RAČUNARA U POSTUPKU PROVERE I DOPUNE SADRŽAJA TOPOGRAFSKIH KARATA . . . . .	175
<i>Siniša Drobnjak, Marko Stojanović, Nenad Galjak, Dejan Đorđević</i>	
MULTICRITERIA DECISION MODEL OF FOREST RISK ASSESSMENT BY GIS APPLICATION . . . . .	181
<i>Ljubomir Gigović, Darko Lukić, Nenad Galjak, Dragoljub Sekulović</i>	
OPTIMIZACIJA PROFESIONALNIH MALIH BESPILOTNIH LETELICA ZA POTREBE GEODEZIJE I GIS-A . . . . .	187
<i>Radoje Banković, Boban Milojković, Dragoljub Sekulović, Miloš Basarić, Ivan Garić</i>	
RAD SA PROSTORNIM PODACIMA . . . . .	195
<i>Saša Bakrač, Boris Vakanjac, Dejan Đorđević, Ivan Potić, Siniša Drobnjak</i>	
REMOTE MONITORING OF ENERGY PLANT STATE . . . . .	201
<i>Nenad Galjak, Miroslav Vujasinović, Miodrag Regodić, Miloš Blagojević</i>	
SATELLITE OBSERVATION WITH COPERNICUS PROGRAM . . . . .	207
<i>Nenad Galjak, Miroslav Vujasinović, Miodrag Regodić, Slađana Stanišić</i>	
UTICAJ VOJNIH INTERNET STVARI NA GEOPROSTORNU ANALIZU BOJIŠTA . . . . .	213
<i>Ivan Vulić, Radomir Prodanović</i>	

<b>GRAFOVI I MREŽE / GRAPHS AND NETWORKS</b>	<b>215</b>
ANALIZA SIMETRIJE U EKONOMSKOM SISTEMU PRE, TOKOM I POSLE EKONOMSKE KRIZE PRIMENOM TEORIJE GRAFOVA .....	217
<i>Vojin Stević, Marija Rašajski, Marija Mitrović Dankulov</i>	
METRIC DIMENSION OF COMPLETE SPLIT GRAPHS .....	223
<i>Milica Milivojević Danas, Jozef Kratica</i>	
OCCURRENCE OF FIBONACCI AND LUCAS SEQUENCES IN COUNTING RESTRICTED PERMUTATIONS .....	227
<i>Vladimir Baltić</i>	
SOME PAIRS AND TRIPLETS OF EQUIENERGETIC GRAPHS .....	233
<i>Irena Jovanović</i>	
<b>HEURISTIKE / HEURISTICS</b>	<b>239</b>
ALTERNATIVE LOCAL SEARCH ALGORITHMS IN GREEDY HEURISTIC ALGORITHMS FOR THE P-MEDIAN PROBLEM .....	241
<i>Lev Kazakovtsev, Ivan Rozhnov, Guzel Shkaberina, Alena Stupina, Igor Masich</i>	
BEE COLONY OPTIMIZATION FOR MULTI-LABEL FEATURE SELECTION .....	243
<i>Luka Matijević</i>	
COMPARATIVE ANALYSIS OF GREEDY CONSTRUCTIVE HEURISTICS FOR SOLVING CAPACITATED VEHICLE ROUTING PROBLEM .....	249
<i>Anđela Đorđević, Milan Stanojević</i>	
IMPROVED VARIABLE NEIGHBORHOOD SEARCH FOR FUZZY CLUSTERING ON COMPLEX NETWORKS .....	255
<i>Filip Vidojević, Lazar Mrkela, Dušan Džamić, Miroslav Marić</i>	
METODA PROMENLJIVIH OKOLINA ZA PROBLEM P-REZERVNOG CENTRA .....	261
<i>Jelena Tasić, Zorica Stanimirović</i>	
PRIMENA KARUSEL HEURISTIKE NA PROBLEM RUTIRANJA VOZILA .....	267
<i>Jana Vugdelija, Dragana Makajić-Nikolić</i>	
THE MODIFIED U-GRASP METHOD FOR THE TWO-DIMENSIONAL VECTOR BIN PACKING PROBLEM .....	273
<i>Đorđe Stakić, Ana Anokić</i>	
<b>INFORMACIONI SISTEMI I TEHNOLOGIJE / INFORMATION SYSTEMS AND TECHNOLOGIES</b>	<b>279</b>
GENERATOR PROGRAMA – LUČA .....	281
<i>Siniša Vlajić, Vojislav Stanojević, Miloš Milić, Ilija Antović, Dušan Savić</i>	
POSLEDICE PRIMENE DIGITALNE KOLABORACIJE .....	287
<i>Rajko Ivanišević</i>	
PRIMENA BENCHMARKDOTNET BIBLIOTEKE U MERENJU PERFORMANSI ALATA ZA OBJEKTNO-RELACIONO PRESLIKAVANJE .NET OKVIRA .....	293
<i>Tijana Milošević, Saša D. Lazarević, Tatjana Stojanović</i>	

THE INFLUENCE OF INTERNET PLATFORMS FOR COMMUNICATION ON LEARNING AND EDUCATION OUTCOMES .....	299
<i>Aleksandra Zečević, Đorđe Stakić</i>	
TRANSFORMACIJA LOGIČKE SCHEME BAZE PODATAKA: OD RELACIONOG (T-SQL) MODELA DO DOKUMENTACIONO-ORIJENTISANOG (MONGODB).....	305
<i>Nemanja Radulović, Tatjana Stojanović, Saša D. Lazarević</i>	

## **ISTRAŽIVANJE I RAZVOJ / RESEARCH AND DEVELOPMENT 311**

GAPS AND NEEDS ANALYSIS OF RESEARCH INFRASTRUCTURES IN THE WESTERN BALKAN REGION .....	313
<i>Lazar Živković, Đuro Kutlača</i>	
MERENJE EFIKASNOSTI NACIONALNIH SISTEMA ZASNOVANO NA TRIPLE HELIX MODELU .....	319
<i>Milica Jovanović, Gordana Savić, Maja Levi-Jakšić, Milan Martić</i>	
PRILOG METODOLOGIJI PRAĆENJA REALIZACIJE 4S – POUKE PRAKSE EU.....	321
<i>Dušica Semenčenko, Đuro Kutlača</i>	
THE ROLE OF SMART SPECIALISATION STRATEGY IN ECONOMIC RECOVERY AND SUSTAINABLE DEVELOPMENT IN THE POST COVID-19 ERA .....	327
<i>Dijana Štrbac, Nikola Vasilić</i>	

## **KOMBINATORNA OPTIMIZACIJA / COMBINATORIAL OPTIMIZATION 333**

A REVERSE RANDOMIZED GREEDY ALGORITHM FOR THE MINIMUM POSITIVE INFLUENCE DOMINATING SET PROBLEM IN SOCIAL NETWORKS .....	335
<i>Kristina Kostić, Zorica Stanimirović</i>	
HYBRID OPTIMIZATION ALGORITHMS BASED ON GREEDY HEURISTICS AND THE BRANCH-AND-BOUND METHOD FOR LOGICAL ANALYSIS OF DATA .....	341
<i>Igor Masich, Lev Kazakovtsev, Alena Stupina, Ekaterina Kraeva</i>	
NEW IDEAS TO SPEED-UP FLOYD-WARSHALL SHORTEST PATHS ALGORITHM .....	343
<i>Giuseppe Lancia, Franca Rinaldi</i>	
THE CHOICE OF HEURISTICS FOR FORMATION OF PARTIAL LOGICAL RULES IN THE METHOD OF LOGICAL ANALYSIS OF DATA .....	349
<i>Roman Kuzmich, Alena Stupina, Katerina Ponomareva, Vladislav Stasiuk</i>	

## **LOGISTIKA I LANCI SNABDEVANJA / LOGISTICS AND SUPPLY CHAIN MANAGEMENT 355**

AN APPROACH TO OPTIMAL SCHEDULING OF COLLECTION VEHICLES FOR CONSTRUCTION AND DEMOLITION WASTE .....	357
<i>Milorad Vidović, Nenad Bjelić, Branislava Ratković</i>	
APPLICABILITY OF INDUSTRY 4.0 TECHNOLOGIES IN LOGISTICS CENTERS.....	363
<i>Snežana Tadić, Smiljka Miškić, Željko Stević, Mladen Krstić</i>	
DEFINISANJE SKUPA FAZI PRAVILA U FAZI SISTEMU ZA IZBOR SNABDEVAČA.....	369
<i>Gordana Radivojević, Milica Mitrović, Dražen Popović</i>	

DYNAMIC DISCRETE INVENTORY CONTROL MODEL IN MULTISTOP DELIVERY MODE FOR TWO WAREHOUSES . . . . .	375
<i>Slobodan Antić, Lena Đorđević Milutinović</i>	
EVOLVEMENT OF BLOCKCHAIN IN SUPPLY CHAINS . . . . .	381
<i>Anđela Đorđević</i>	
IDENTIFICATION OF RISK IN SUPPLY CHAINS . . . . .	387
<i>Tamara Bojanić, Branislav Stevanov, Nenad Novaković</i>	
IMPACT OF LOCATION ON DELIVERY DURATION IN URBAN AREAS . . . . .	389
<i>Mladen Krstić, Miloš Veljović, Snežana Tadić, Slobodan Zečević</i>	
INTEGRATING BLOCKCHAIN INTO SUPPLY CHAIN MANAGEMENT . . . . .	395
<i>Milan Todorović, Luka Matijević</i>	
SUSTAINABLE SUPPLY CHAIN MANAGEMENT . . . . .	401
<i>Milica Grujić, Aleksandar Đukić</i>	

## **MATEMATIČKO PROGRAMIRANJE / MATHEMATICAL PROGRAMMING** **407**

Invited speaker of the session . . . . .	409
RELEVANT REASONING . . . . .	409
<i>Mirjana Ilić</i>	
Regular Communication . . . . .	415
DIRP MODEL PREDIKCIJE SKUPA BROJEVA . . . . .	415
<i>Siniša Vlajić, Dušan Savić, Ilija Antović, Miloš Milić, Vojislav Stanojević</i>	
A PARAMETRIC SOLUTION TO WEIGHTED SUM BI-CRITERIA MODELS. . . . .	421
<i>Moshe Eben-Chaime</i>	

## **MEKO RAČUNARSTVO / SOFT COMPUTING** **423**

ANALYTIC DESCRIPTION OF FUZZY SET SOLUTIONS TO FUZZY "LP" PROBLEMS. . . . .	425
<i>Bogdana Stanojević, Milan Stanojević</i>	
ANALYSIS AND PREDICTION OF NEW INFECTIONS DURING COVID-19 PANDEMIC USING ENSEMBLE ALGORITHMS . . . . .	431
<i>Nikola Đorđević, Aleksandar Rakićević and Nemanja Milenković</i>	
FAZI SISTEM ZA ALGORITAMSKO TRGOVANJE ZASNOVAN NA TEHNIČKOJ ANALIZI . . . . .	433
<i>Andrija Matić, Milica Zukanović, Aleksandar Rakićević</i>	
PRIMENA FAZI KLASTEROVANJA ZA KATEGORISANJE TENISKIH IGRAČA. . . . .	435
<i>Luka Tasić, Ana Poledica, Milica Zukanović</i>	
SMS SPAM CLASSIFICATION USING COMPUTATIONAL INTELLIGENCE. . . . .	441
<i>Danilo Janjušević, Ivana Dragović, Pavle Milošević</i>	
TALENT RECOGNITION IN ELEMENTARY SCHOOL STUDENTS USING NEURAL NETWORKS AND FUZZY LOGIC . . . . .	447
<i>Natalija Damjanović, Ivana Dragović, Pavle Milošević</i>	

**MENADŽMENT / MANAGEMENT 453**

- INFLUENCE OF SOCIO-CULTURAL CHARACTERISTICS OF CONSUMERS ON THE PURCHASE OF GLOBAL BRANDS . . . . . 455  
*Stefan Zdravković, Aleksandar Jovanović*

**NAUKA O PODACIMA / DATA SCIENCE 461**

- ALGORITHM FOR VISUAL ASSESSMENT OF THE CLUSTERING TENDENCY WITH SPECIAL DISTANCE MEASURES . . . . . 463  
*Guzel Shkaberina, Lev Kazakovtsev, Natalya Rezova, Ivan Rozhnov*
- IMPROVING QUALITY OF CONVOLUTIONAL NEURAL NETWORK MODELS FOR OFFLINE SIGNATURE VERIFICATION . . . . . 465  
*Siniša Stanivuk, Dejan Simić*
- NEKI ASPEKTI DEMOKRATIZACIJE NAUKE O PODACIMA . . . . . 471  
*Jasna Soldić Aleksić, Rade Stankić, Biljana Chroneos Krasavac*
- PRIMENA STABALA ODLUČIVANJA U ANALIZI UZROKA I VRSTA SAOBRAĆAJNIH NEZGODA NA TERITORIJI REPUBLIKE SRBIJE . . . . . 477  
*Aleksa Maksimović, Vladica Stojanović*
- THE PARETO PRINCIPLE AS A UTILITY MEASURE OF MACHINE LEARNING MODELS . 483  
*Boris Delibašić, Sandro Radovanović, Andrija Petrović, Milija Suknović*
- WIND POWER PREDICTION USING MACHINE LEARNING AND AI: A CASE STUDY . . . . 485  
*Israt Haque Zarin, Hasnain Hasnain, Amzad Hossain, Md Abu Helal, Viswanath Ramakrishna*

**POSLOVNA ANALITIKA / BUSINESS ANALYTICS 487**

- DETERMINING ENTREPRENEURS' PREFERENCES TOWARDS ONLINE MARKET RESEARCH USING CHOICE-BASED-CONJOINT ANALYSIS . . . . . 489  
*Ognjen Nikolić, Marija Kuzmanović*
- EVALUACIJA UTICAJA PANDEMIJE COVID-19 NA POTRAŽNJU ZA PIVOM U REPUBLICI SRBIJI . . . . . 495  
*Ognjen Anđelić, Zoran Rakićević, Aleksandar Rakićević*
- TESTING ANALYTICS . . . . . 501  
*Dragan Azdejković, Slavica Manić*
- THE COVID-19 INFLUENCE ON SMES IN A DEVELOPING COUNTRY: A DELPHI METHOD . . . . . 505  
*Mimoza Arifi, Violeta Cvetkoska*

**PRIMENE OI U ODBRANI / OR APPLICATIONS IN MILITARY DEFENCE 511**

- DEFINING CRITERIA FOR IDENTIFYING SECURITY CHALLENGES, RISKS AND THREATS USING THE MATHEMATICAL MODEL OF MULTICRITERIA DECISION MAKING . . . . . 513  
*Dragan Bojanić, Marina Bojanić, Vladimir Ristić*
- MODEL RECENZIJE NAUČNOISTRAŽIVAČKIH PROJEKATA U MINISTARSTVU ODBRANE I VOJSCI SRBIJE . . . . . 519  
*Srđan Dimić, Srđan Ljubojević, Dragan Kostadinović*

MULTI-CRITERIA ANALYSIS IN DEVELOPMENT OF ACADEMIC STUDY PROGRAMS FOR EXPERTISE ON HAZARDOUS MATERIALS SAFETY AND PROTECTION . . . . .	525
<i>Jovica Bogdanov, Zoran Bajić, Zlate Veličković, Mihael Bučko</i>	
ODREĐIVANJE VELIČINE POPREČNOG PRESEKA ELASTIČNE GREDE BEZLEŽAJNOG ROTORA BESPILOTNOG HELIKOPTERA . . . . .	529
<i>Dalibor Petrović, Vladimir Stanković, Ivan Mudri</i>	
VARIATION OF THE RESULTS OF MULTI-CRITERIA SELECTION DEPENDING ON APPLIED MULTI-CRITERIA DECISION MAKING METHOD . . . . .	535
<i>Srđan Ljubojević, Srđan Dimić, Dragan Kostadinović, Aleksandar Gošić</i>	

## **RUDARSTVO, GEOLOGIJA I ENERGETIKA / MINING, GEOLOGY AND ENERGETICS** **541**

RANKING OPEN-PIT MINES USING EDAS++ METHOD . . . . .	543
<i>Željko Praštalo, Anđelka Štilić, Stevan Ćorluka, Pavle Stjepanović</i>	
THE SERBIAN CLIMATE CHANGE ACT: THE PERSPECTIVE OF IMPLEMENTATION OF CARBON CAPTURE AND STORAGE PROJECTS IN THE OIL AND GAS INDUSTRY . . . . .	549
<i>Serguei Fominykh</i>	

## **SAOBRAĆAJ, TRANSPORT I KOMUNIKACIJE / TRAFFIC, TRANSPORTATION AND COMMUNICATION** **555**

A DECISION TREE MODEL FOR AIRPORT ACCESS MODE CHOICE: SERBIAN AIR TRANSPORT MARKET . . . . .	557
<i>Katarina Kukić, Slavica Dožić, Danica Babić, Milica Kalić</i>	
AIRCRAFT ENCOUNTER CHARACTERIZATION: A SCENARIO-BASED SIMULATION APPROACH . . . . .	563
<i>Lidia Serrano-Mira, Feđa Netjasov, Luis Pérez Sanz, Javier Alberto Pérez-Castán</i>	
ANALYSIS OF METEOROLOGICAL CONDITIONS FOR POTENTIAL AIRPORT DEVELOPMENT . . . . .	569
<i>Nikola Lukačević, Bojana Mirković, Ana Uzelac, Katarina Kukić</i>	
INFLUENCE OF AUTOMATION CHANGES ON GO-AROUND OPERATIONS: A DATA-DRIVEN APPROACH . . . . .	575
<i>Doroteja Timotić, Bojana Mirković, Irene Buselli, Feđa Netjasov, Chen Xia, Christian Eduardo Verdonk Gallego</i>	
ODREĐIVANJE RASPOREDA RADA SKLADIŠNIH DIZALICA U LUČKIM KONTEJNERSKIM TERMINALIMA . . . . .	581
<i>Miloš Nikolić, Ivana Vukićević Biševac, Jovko Jaćimović, Ivana Jovanović, Katarina Vukadinović, Branka Dimitrijević</i>	
PREDIKCIJA POSLEDICA SAOBRAĆAJNIH NEZGODA POMOĆU MODELA BAZIRANOG NA ALGORITMU LOGISTIČKE REGRESIJE . . . . .	587
<i>Predrag Grozdanović, Katarina Kukić, Ana Uzelac, Slađana Janković</i>	
PREDVIĐANJE KAŠNJENJA LETOVA U DOLASKU KORIŠĆENJEM MAŠINSKOG UČENJA . . . . .	593
<i>Nenad Golubović, Slađana Janković, Ana Uzelac, Snežana Mladenović</i>	



## **SIMULACIJA I STOHAŠTIČKI MODELI / SIMULATION AND STOCHASTIC MODELS** **599**

BOUNDS ON THE COUPON COLLECTOR PROBLEM . . . . .	601
<i>Jelena Jocković, Bojana Todić</i>	
METODOLOGIJA KREIRANJA VEŽBI NA BRODSKOM SIMULATORU . . . . .	603
<i>Slobodan Radojević, Milan Kresojević</i>	
PREPARATION OF THE SIMULATION ON THE SHIP SIMULATOR NAVI-TRAINER PROFESSIONAL 5000 . . . . .	605
<i>Milan Kresojević, Slobodan Radojević</i>	
THE SIMULATION MODEL OF BLOCKCHAIN TRANSACTION. . . . .	607
<i>Ana Savić, Hana Stefanović, Goran Bjelobaba, Nikola Popović</i>	

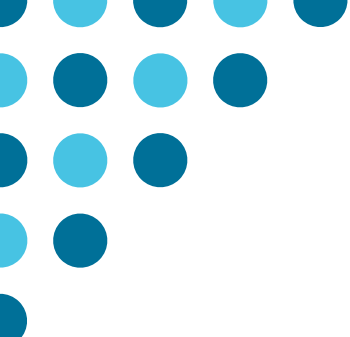
## **STATISTIČKI MODELI / STATISTICAL MODELS** **613**

ANALIZA BROJA OBOLELIH OD VIRUSA COVID-19 U SRBIJI PRIMENOM ARIMA MODELA. . . . .	615
<i>Emilija Maksimović</i>	
ANALIZA VARIJABILITETA PRIHODA U ZDRAVSTVENIM USTANOVAMA - NA PRIMERU PRIVATNIH BOLNICA U BEOGRADU . . . . .	621
<i>Vesna Rajić, Tatjana Rakonjac-Antić, Ivana Ivković, Jasna Babić</i>	
BENFORD'S LAW APPLICATION ON ATTRACTING INVESTMENT AND NEW EMPLOYMENT IN THE REPUBLIC OF SERBIA DURING THE PERIOD OF 2016-2022 . . .	627
<i>Jelena Stanojević, Milena Lutovac Đaković, Željko Jović</i>	
QUALITY EVALUATION OF CLUSTERING RESULTS: SILHOUETTE COEFFICIENT . . . . .	633
<i>Milan Stamenković, Marina Milanović</i>	
STATISTICAL CLASSIFICATION OF NON-ECONOMIC FACTORS IN THE FUNCTION OF IMPROVING BUSINESS EFFICIENCY IN BANKING OPERATIONS . . . . .	635
<i>Nina Uremović, Mirjana Landika, Jana Aleksić, Maja Đurić</i>	

## **VIŠEKRITERIJUMSKA ANALIZA I OPTIMIZACIJA / MULTIPLE CRITERIA ANALYSIS AND OPTIMIZATION** **637**

OMNI-CHANNEL RETAIL: MULTIPLE CRITERIA DECISION AIDING APPROACH . . . . .	639
<i>Mladen Stamenković, Dragan Stojković, Aleksa Dokić</i>	





# PLENARNA PREDAVANJA

---

## PLENARY SESSIONS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business



## INTERACTIVE EVOLUTIONARY MULTIOBJECTIVE OPTIMIZATION GUIDED BY PREFERENCE INFORMATION

SALVATORE CORRENTE<sup>1</sup>

<sup>1</sup> Department of Economics and Business, University of Catania, salvatore.corrente@unict.it

---

**Abstract:** *In multiobjective optimization one aims to simultaneously optimize a certain number of objective functions possibly under one or more constraints. However, since objectives are conflicting so that maximizing one of them implies that some others are reduced, there is not any solution that optimizes simultaneously all of them. Instead, a set of Pareto optimal solutions exists, containing solutions for which it is not possible to improve one objective without deteriorating someone other. In recent years evolutionary algorithms have been applied to multiobjective optimization problems to find a good approximation of the Pareto front. However, finding the whole set of Pareto optimal solutions, as well as some approximation, does not solve the optimization problem since one or some of these solutions, being the most preferred for the user, have to be discovered. Interactive evolutionary multiobjective optimization methods aim, therefore, to address the search to the most interesting part of the Pareto front by considering some preferences provided by the user during the procedure. In this talk we shall present two interactive evolutionary multiobjective optimization methods being NEMOICh and XIMEA-DRSA. Both take into account the preferences expressed by the user during the optimization procedure. On the one hand, NEMOICh tries to discover solutions that are the most preferred for at least one value function, formulated in terms of Choquet integral, compatible with the preference information provided by the user, while, on the other hand, XIMEA-DRSA guides the evolutionary optimization towards the most preferred part of the Pareto front using “if ..., then...” decision rules induced from the preferences expressed by the user.*

**Keywords:** *Multiobjective Optimization, Interactive Evolutionary Multiobjective Optimization, Multiple Criteria Decision Aiding, Choquet integral, Dominance-Based Rough Set Approach*



## RECOMMENDING MULTIPLE CRITERIA DECISION ANALYSIS METHODS WITH A NEW TAXONOMY-BASED DECISION SUPPORT SYSTEM

MIŁOSZ KADZIŃSKI<sup>1</sup>

<sup>1</sup> Poznań University of Technology, Institute of Computing Science, Poland, milosz.kadzinski@cs.put.poznan.pl

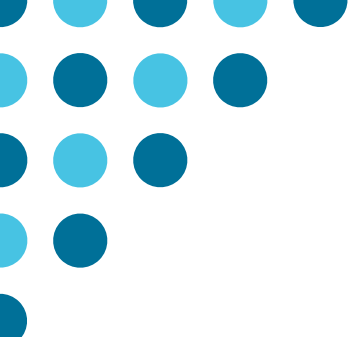
---

**Abstract:** *The aim of the talk is to present a new methodology to lead the selection of Multiple Criteria Decision Analysis (MCDA) methods. It is implemented in the Multiple Criteria Decision Analysis Methods Selection Software (MCDA-MSS), a decision support system that helps analysts answer a recurring question in decision science: “Which is the most suitable Multiple Criteria Decision Analysis method (or a subset of MCDA methods) that should be used for a given Decision-Making Problem (DMP)?”. The MCDA-MSS provides guidance to lead decision-making processes and choose among an extensive collection (>200) of MCDA methods. These are assessed according to an original comprehensive set of problem characteristics. The accounted features concern problem formulation, preference elicitation and types of preference information, desired features of a preference model, and construction of the decision recommendation. The applicability of the MCDA-MSS has been tested on several case studies. The MCDA-MSS includes the capabilities of (i) covering from very simple to very complex DMPs, (ii) offering recommendations for DMPs that do not match any method from the collection, (iii) helping analysts prioritize efforts for reducing gaps in the description of the DMPs, and (iv) unveiling methodological mistakes that occur in the selection of the methods. A community-wide initiative involving experts in MCDA methodology, analysts using these methods, and decision-makers receiving decision recommendations will contribute to the expansion of the MCDA-MSS.*

**Keywords:** *Decision analysis, Multiple criteria, Taxonomy, Decision support system, Method recommendation*







# ANALIZA PERFORMANSI

---

# PERFORMANCE ANALYSIS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business



## ANALIZA EFIKASNOSTI KOMPANIJA U STICANJU KONKURENTSKE PREDNOSTI PRIMENOM DEA METODE

### ANALYSIS OF COMPANY EFFICIENCY IN GAINING COMPETITIVE ADVANTAGE USING DEA METHOD

IVONA JOVANOVIĆ<sup>1</sup>, MILAN RADOJIČIĆ<sup>2</sup>, GORDANA SAVIĆ<sup>3</sup>

<sup>123</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, ivona.jovanovic@fon.bg.ac.rs, milan.radojicic@fon.bg.ac.rs, gordana.savic@fon.bg.ac.rs

**Rezime:** Cilj ovog rada je merenje efikasnosti kompanija sistemom samoprocene, kroz faktore unapređivanja procesa i njihovog uticaja na konkurentsku prednost kompanije. Primenom različitih inicijativa za unapređivanje procesa, može se značajno uticati na povećanje konkurentnosti kompanije na tržištu i efikasnosti kompanije. Za merenje efikasnosti kompanija su primenjena dva DEA modela za optimizaciju mere efikasnosti baziranoj na izravnavajućim promenljivama (Slack Based Measures – SBM), sa konstantnim prinosom na obim i sa varijabilnim prinosom na obim. Podaci o ulaznim i izlaznim parametrima su prikupljeni online anketom, u kojoj je učestvovalo 82 ispitanika. Rezultati pokazuju razliku u prosečnoj efikasnosti u delatnosti u kojoj kompanija posluje, a najmanja prosečna efikasnost je utvrđena za kompanije koje se bave proizvodno-uslužnom delatnošću. Na osnovu dobijenih rezultata istraživanja, date su preporuke za analizu trenda efikasnosti kompanija iz različitih delatnosti.

**Ključne reči:** unapređivanje procesa, konkurentna prednost, SBM DEA model

**Abstract:** The aim of this research is to measure the efficiency of companies using a self-assessment system, through the factors of process improvement and their influential competitive advantage of the company. By implementing different initiatives for process improvement, it can significantly increase market competitiveness and company efficiency. To measure the efficiency of companies, two DEA models were applied to optimize the efficiency measure, based on Slack Based Measures (SBM), with a constant return on volume and with a variable return on volume. Data on input and output parameters were collected by online survey, in which 82 respondents participated. The results show a difference in average efficiency in the industry in which the company operates. The lowest average efficiency was determined for companies engaged in both production and service industry. Based on the obtained research results, recommendations are given for the analysis of the efficiency trend of companies from different industries.

**Keywords:** process improvement, competitive advantage, SBM DEA model

## 1. UVOD

Sa visokim nivoom konkurentnosti na globalnom tržištu, za kompanije postaje sve važnije da unaprede performanse procesa kako bi bile uspešnije u sticanju konkurentne prednosti. Iako primena različitih metoda unapređivanja procesa u cilju povećanja konkurentnosti u praksi raste tokom godina, studije pokazuju da je oko 80% inicijativa za unapređenje procesa bilo neuspešno [2, 4, 26]. Kao osnovni razlog neuspeha se javlja složenost implementacije [5], jer kompanije žele da ostvare rezultate i postignu uspeh za što kraće vreme [24, 28]. Pošto je konkurentna prednost jedan od glavnih izvora za dostizanje kontinuiranog unapređenja poslovnih procesa [1], važno je da se kompanije fokusiraju na različite strategije povećanja konkurentnosti i utvrde šta najbolje dovodi do poboljšanja performansi procesa. Kompanija treba da unapredi svoje procese, kao i ljudske resurse, kako bi bila u ravnoteži sa turbulentnim poslovnim okruženjem i održavala konkurentnu prednost na tržištu [14, 17]. Ako kompanija strateški ne poboljša performanse, njihovi konkurenti hoće [3].

Iako postoje istraživanja o tome kako unapređenje procesa utiče na performanse kompanije, postoji nedostatak praktičnih studija o uspešnim implementacijama u kompanijama. Stojanović i dr. [24] ukazuju na nedostatak strateškog fokusa menadžmenta za unapređivanje procesa i njihovu implementaciju na nivou liderstva, dok su inicijative za unapređenje procesa fokusirane uglavnom na operativne performanse. Poboljšanje procesa mora biti usvojeno kroz usklađivanje sa organizacionom strukturom i osnaživanjem ljudi

[18]. Važno je uspostaviti ravnotežu između zaposlenih i menadžmenta kompanije u inicijativama za unapređenje procesa u kompaniji. Ako zaposleni nemaju osnovne veštine i znanja za podršku inicijativama za unapređenje procesa, ili ako menadžment ne učestvuje u poboljšanjima, efekti mogu biti nezadovoljavajući [3]. Zbog toga, ovim istraživanjem se ispituje koje inicijative za unapređenje procesa mogu biti održiva praksa koja pomaže kompanijama da ostvare konkurentsku prednost i željene rezultate.

Ovo istraživanje analizira efikasnost kompanija primenom metode analize obavljanja podataka (*Data Envelopment Analysis* – DEA), u vidu unapređivanja poslovnih procesa (*Business Process Improvement* - BPI) i uticaja na konkurentsku prednost kompanije. Rezultati istraživanja se baziraju na sistemu lične percepcije ispitanika, pa je bilo potrebno da se kroz *online* anketu izrazi lični stav o kompaniji i njenom poslovanju. Istraživanje se bazira na sličnim istraživanjima iz ove oblasti [15, 23]. Takođe, skala ocenjivanja je formirana tako da omogućava ispitaniku da proceni važnost svakog faktora. Takođe, bilo je potrebno da ispitanici navedu osnovne karakteristike svoje kompanije, čime je bilo interesantno ispitati uspešnost kompanija iz različitih delatnosti – proizvodne, uslužne i proizvodno-uslužne delatnosti. Rezultati istraživanja su omogućili da se izmeri efikasnost kompanija sistemom samoprocene uz pomoć DEA metode, korišćenjem dva modela, a podaci o ulazima i izlazima su dobijeni iz ankete formiranjem kompozitnih indikatora.

Rad je strukturiran na sledeći način: sledeće poglavlje prikazuje izbor metodologije istraživanja, treće poglavlje predstavlja analizu i rezultate istraživanja, koje je praćeno zaključkom i budućim smernicama u okviru četvrtog poglavlja.

## 2. METODOLOGIJA

Unapređivanje procesa je posmatrano iz perspektive menadžmenta i zaposlenih u kompaniji, kako bi se utvrdio međusobni uticaj i na koji način može dovesti do povećanja konkurentске prednosti kompanije u zavisnosti od delatnosti u kojoj kompanija posluje. Izjave (Tabela 1) su modifikovane prema istraživanju Wolf & Harmon [27].

**Tabela 1:** Izjave o unapređivanju procesa, modifikovano prema [27]

Unapređivanje procesa	Izjave
Menadžment	Rukovodilac pruža vidljivu podršku (budžetom, aktivnostima, itd.)
	Rukovodioci poštuju organizacione ciljeve, prioritete i strategije
	Imamo mogućnost da identifikujemo prepreke unapređivanja procesa
	Imamo mogućnost da prevaziđemo prepreke unapređivanja procesa
	Postavili smo kratkoročne i dugoročne ciljeve unapređivanja procesa
	Plan unapređivanja procesa je dokumentovan
	Planove unapređivanja procesa treba detaljnije razraditi
	Zaposleni
Zaposleni	Uspešni smo u dobijanju podrške zaposlenih za unapređivanje procesa
	Zaposleni razumeju potrebu za unapređivanjem procesa
	Zaposleni vide svoju ulogu u unapređivanju procesa
	Uloga svakog zaposlenog je detaljno saopštena u unapređivanju procesa
	Sopstveni ciljevi se odražavaju na dugoročne ciljeve unapređivanja procesa

U okviru istraživanja je bilo potrebno da ispitanik navede u kojoj meri se slaže sa različitim tvrdnjama o konkurentskoj prednosti svoje kompanije. Konkurentska prednost je posmatrana kroz četiri dimenzije: cena, kvalitet, zadovoljstvo klijenata i plasman proizvoda. Ove merljive karakteristike kvaliteta proizvoda važne su za postavljanje strategije unapređenja procesa. Prema Kumaru i dr. [13] i Stojanović i dr. [24], zadovoljstvo klijenata predstavlja važan pokazatelj uspešnosti unapređenja procesa, kao i kvaliteta proizvoda [16]. Trihatmoko i dr. [26] ističu plasman proizvoda i inovativnost proizvoda kao važan fenomen u poslovnoj tržišnoj konkurenciji, pri čemu kompanija treba da razmotri i cenu proizvoda za sticanje konkurentске prednosti [7].

DEA predstavlja neparаметarsku metodu koja ocenjuje da li je jedinica odlučivanja (*Decision Making Units* – DMU) koja se posmatra efikasna ili ne, u odnosu na ostale jedinice koji se analiziraju, i koje koriste iste ulaze za proizvodnju istih izlaza [21]. Prvi DEA model, CCR model, kreiran je 1978. godine [6]. Rasprostranjenost primene DEA metode se ogleda u tome što se može koristiti za merenje efikasnosti u onim sistemima gde se posmatra više ulaza i više izlaza, koji mogu biti raznorodni i mogu biti predstavljeni različitim mernim jedinicama [21, 22].

U okviru ovog istraživanja, izabran je model za optimizaciju mere efikasnosti baziranoj na izravnavajućim promenljivama (*Slack Based Measures* – SBM) [25]. Efikasnost se izražava u razlomljenoj formi, a vrednost

funkcije cilja ne zavisi od mernih jedinica ulaza i izlaza. Model za procenu indeksa efikasnosti kao osnovu sadrži model razlomljenog programiranja, a funkcija cilja u modelu se zasniva na modifikovanoj Raselovoj meri efikasnosti [9, 8, 20].

Pretpostavimo da  $DMU_j, j = 1, \dots, n$  koristi vrednosti ulaza  $x_{ij}, i = 1, \dots, m$  za kreiranje izlaza  $y_{rj}, r = 1, \dots, s$ . Osnovni SBM model za merenje efikasnosti  $DMU_k$  dat je jednačinama (1)-(6).

$$(\min)\rho_k = \frac{1 - \frac{1}{m} \sum_{i=1}^m s_i^- / x_{ik}}{1 + \frac{1}{s} \sum_{r=1}^s s_r^+ / y_{rk}} \quad (1)$$

p.o.

$$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{ik}, i = 1, \dots, m \quad (2)$$

$$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = y_{rk}, r = 1, \dots, s \quad (3)$$

$$\lambda_j \geq 0, j = 1, \dots, n, \varepsilon > 0 \quad (4)$$

$$s_i^- \geq 0, i = 1, \dots, m, s_r^+ \geq 0, r = 1, \dots, s \quad (5)$$

$$\sum_{j=1}^n \lambda_j = 1 \quad (6)$$

gde  $\rho$  predstavlja neradikalnu relativnu efikasnost  $DMU_k$ .

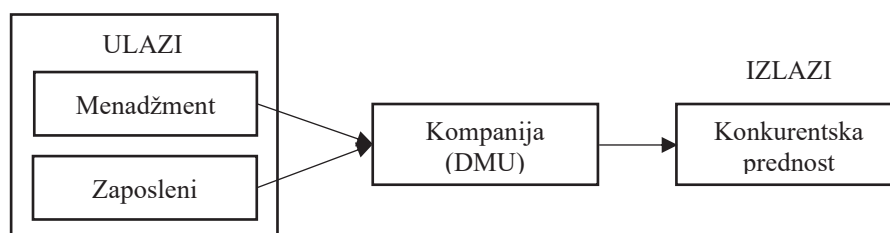
### 3. ANALIZA I REZULTATI

Ovim istraživanjem je analizirana efikasnost kompanija sistemom samoprocene, koja je ispitivana kroz uticaj različitih faktora unapređivanja procesa na konkurentsku prednost kompanije. U cilju prikupljanja podataka, sprovedeno je *online* istraživanje. Anketu su činila pitanja koja su bila deo šireg istraživanja o upravljanju poslovnim procesima (*Business Process Management* - BPM) i praksama lin transformacije. Pitanja o BPM praksama su podeljena u četiri kategorije: opšta pitanja o kompaniji i ispitanicima (7 pitanja), pitanja o zrelosti procesa, Indeksu performansi procesa (12 pitanja), BPM i BPI praksi (59 pitanja), i pitanja o ispunjenju ciljeva i konkurentskoj prednosti (22 pitanja).

U istraživanju je primenjen sistem samo-ocenjivanja, u kome su ispitanici odgovarali na anketna pitanja koja su uključivala jedan ili višestruki izbor odgovora, dok je nekoliko pitanja je sadržalo otvorenu formu kako bi se bolje sagledao stav ispitanika.

Efikasnost kompanija je merena posredstvom dva modela. Kako bi smo utvrdili efikasnost kompanije sistemom samoprocene, korišćeni su podaci prikupljeni kroz *online* anketu koja je obuhvatila zaposlene u vodećim domaćim i stranim kompanijama [12]. Anketa i pismo sa pozivom za učešće poslato je za 700 potencijalnih učesnika. U anketi je učestvovalo 82 ispitanika, što predstavlja stopu odgovora od 11,71%. Prema Wongu et al. [28], stopa odgovora u studijama iz ove oblasti varira od 6,7% do 20%, pa se stopa odgovora u ovom istraživanju može smatrati adekvatnom. Petostepena Likertova skala je služila za formiranje promenljivih. Međutim, čak 43 kompanije su isključene iz dalje analize zbog nedovoljnog zalaganja ispitanika prilikom davanja odgovora [10, 11]. Kriterijumi koji su uzeti u obzir prilikom isključivanja kompanije iz analize su broj nedostajućih podataka, korišćenje manje od tri različite vrednosti sa Likertove skale kao i uzastopno ponavljanje istog odgovora više od 10 puta [19]. U uzorku je ostala 21 kompanija koja se bavi proizvodnom delatnošću, 12 kompanija koje se bave uslužnom delatnošću, dok se šest kompanija bavi proizvodno-uslužnom delatnošću.

Dva SBM DEA modela [25] su sprovedena kako bi se izmerila i uporedila efikasnost kompanija – SBM DEA model sa konstantnim prinosom na obim (Model I) i SBM DEA model sa varijabilnim prinosom na obim (Model II). Podaci o ulazima i izlazima su dobijeni iz ankete formiranjem kompozitnih indikatora (Slika 1).



Slika 1: Model za utvrđivanje efikasnosti kompanija

Oba DEA modela imaju iste ulaze koji predstavljaju unutrašnju organizaciju kompanija (*Menadžment* i *Zaposleni*) i iste izlaze predstavljene kroz konkurentske prednosti (KP) (*Cenovna, Kvalitativna, Zadovoljstvo klijenata, Plasman proizvoda*). Prvi ulaz *Menadžment* je formiran na osnovu sedam pitanja, dok je drugi ulaz *Zaposleni* formiran na osnovu šest pitanja. Svi izlazi, osim *Cenovne* konkurentske prednosti koja je formirana na osnovu samo dva anketna pitanja, su formirani na osnovu tri pitanja. Pre formiranja kompozitnih indikatora sprovedena je ocena pouzdanosti merne skale merenjem unutrašnje saglasnosti upotrebom Kronbahovog koeficijenta alfa. Vrednosti koeficijenta su između 0,760 i 0,903 što je veće od preporučene minimalne granice od 0,7. Rezultati su dati u Tabeli 2.

**Tabela 2:** Rezultati Kronbahovog alfa koeficijenta prema grupama pitanja

	Unapređivanje procesa (BPI)		Konkurentska prednost (KP)			
	Menadžment	Zaposleni	Cena	Kvalitet	Zadovoljstvo klijenata	Plasman proizvoda
Broj stavki	6	7	2	3	3	3
Kronbahov alfa koef.	0,825	0,845	0,795	0,903	0,760	0,761

Prema Modelu I, samo su tri kompanije efikasne uz prosečnu efikasnost od 64,52%, dok je čak 14 kompanija efikasno prema Modelu II, sa prosečnom efikasnošću od 84,92%. Interesantno je da je prema Modelu I, najveća prosečna efikasnost u uslužnoj delatnosti (70,64%), dok je u Modelu II najveća prosečna efikasnost u proizvodnoj delatnosti (86,78%). Prema oba modela, najmanja prosečna efikasnost se pokazala za kompanije koje posluju u proizvodno-uslužnoj delatnosti.

Kompanije 5, 12 i 19 su jedine efikasne u oba modela. Kompanija 5 pripada proizvodnoj delatnosti, dok se kompanije 12 i 19 bave uslužnom delatnošću. U Modelu II, od 14 efikasnih kompanija, šest se bavi proizvodnom delatnošću, šest uslužnom delatnošću, dok se dve kompanije bave proizvodno-uslužnom delatnošću. Najneefikasnija kompanija u oba modela je Kompanija 35 (22,30% prema Modelu I i 31,37% prema Modelu II), koja pripada proizvodno-uslužnoj delatnosti. Njeni ulazi su na nivou prosečnih vrednosti za sve kompanije dok je prema konkurentskim prednostima ona najlošija u svakom izlazu. Rezultati efikasnosti po oba modela su dati u Tabeli 3.

**Tabela 3:** Efikasnost kompanija prema Modelima I i II

Kompanija	Efikasnost - model I	Efikasnost - model II	Kompanija	Efikasnost - model I	Efikasnost - model II
Kompanija 1	0,636140	0,873820	Kompanija 21	0,669508	1,000000
Kompanija 2	0,543261	0,857143	Kompanija 22	0,554450	1,000000
Kompanija 3	0,536349	0,748552	Kompanija 23	0,861075	1,000000
Kompanija 4	0,857964	1,000000	Kompanija 24	0,778756	1,000000
Kompanija 5	1,000000	1,000000	Kompanija 25	0,680999	0,940325
Kompanija 6	0,488486	0,713365	Kompanija 26	0,652625	0,848552
Kompanija 7	0,607206	0,773125	Kompanija 27	0,682853	1,000000
Kompanija 8	0,612357	0,794471	Kompanija 28	0,528141	0,663972
Kompanija 9	0,684899	0,889638	Kompanija 29	0,486992	0,495533
Kompanija 10	0,724438	0,814941	Kompanija 30	0,757294	1,000000
Kompanija 11	0,778678	1,000000	Kompanija 31	0,517647	0,720721
Kompanija 12	1,000000	1,000000	Kompanija 32	0,529115	0,733427
Kompanija 13	0,517348	0,776948	Kompanija 33	0,414975	0,721149
Kompanija 14	0,536726	0,701444	Kompanija 34	0,773327	0,839513
Kompanija 15	0,428729	0,533927	Kompanija 35	0,222970	0,313725
Kompanija 16	0,675769	1,000000	Kompanija 36	0,689881	1,000000
Kompanija 17	0,607736	0,871898	Kompanija 37	0,536397	0,924444
Kompanija 18	0,643972	0,797153	Kompanija 38	0,611461	0,897381
Kompanija 19	1,000000	1,000000	Kompanija 39	0,561999	0,873950
Kompanija 20	0,771982	1,000000			

### 3. ZAKLJUČAK

Ovim radom je ispitana efikasnost kompanija u pogledu primene faktora unapređivanja procesa na uspešnost konkurentske prednosti na tržištu. Efikasnost je merena uz pomoć DEA metode sistemom samoprocene, a podaci su dobijeni *online* anketom, kroz koju su ispitanici izrazili lični stav o kompaniji i njenom poslovanju. Prilikom analize su isključene kompanije zbog nedostajućih podataka, u slučaju kada su

korišćene manje od tri različite vrednosti sa Likertove skale, kao i za uzastopno ponavljanje istog odgovora više od 10 puta. Za dalju analizu je preostao uzorak od 21 kompanije sa proizvodnom delatnošću, 12 uslužnih kompanija i šest kompanija koje posluju u proizvodno-uslužnoj delatnosti. Korištena su dva SBM DEA modela, koja su sprovedena kako bi se izmerila i upoređila efikasnost kompanija. Oba DEA modela su imala iste ulaze koji predstavljaju unutrašnju organizaciju kompanija (*Menadžment* i *Zaposleni*) i iste izlaze predstavljene kroz konkurentske prednosti (KP) (*Cenovna*, *Kvalitativna*, *Zadovoljstvo klijenata*, *Plasman proizvoda*). Prvi model je izlazno orijentisani modeli sa konstantnim prinosom na obim, a najveća prosečna efikasnost je dobijena za kompanije koje posluju u uslužnoj delatnosti. Sa druge strane, prema modelu sa varijabilnim prinosom na obim, najveća prosečna efikasnost je dobijena za kompanije koje posluju u proizvodnoj delatnosti. Prema prvom modelu su samo tri kompanije efikasne uz prosečnu efikasnost od 64,52%, dok je prema drugom modelu efikasno mnogo više kompanija, čak 14 kompanija, sa većom prosečnom efikasnošću od 84,92%. Interesantno je što se primenom oba modela utvrdilo da je najmanja prosečna efikasnost dobijena za kompanije koje posluju u proizvodno-uslužnom sektoru. Razlog tome može biti šira sfera poslovanja u kojoj posluje kompanija ili veća konkurencija koja postoji u proizvodno-uslužnom sektoru. Dobijeni rezultati pružaju dobru osnovu za analizu trenda efikasnosti kompanija koje posluju u različitim sektorima i koje se bave različitim delatnostima.

Buduće istraživanje je planirano kao integracija sa SEM modelom (*Structural equation modeling* – SEM), čija bi primena omogućila dodeljivanje težinskih koeficijenata svakom faktoru, čime bi svaki ulaz i izlaz imao određeni značaj prilikom merenja efikasnosti. Takođe, ponavljanje istraživanja i poređenje rezultata bi omogućilo da se utvrdi da li postoje promene u efikasnosti kompanija u odnosu na delatnosti u kojima posluju, što bi se moglo dalje analizirati primenom drugih modela.

## LITERATURA

- [1] Abbas, S., & Hosein, D. (2010). Application of Analytic Network Process in Selection of Six Sigma Projects. *International Journal of Insutrial & Production Research*, 20(4), 157-164.
- [2] Abdolvand, N., Albadvi, A., & Ferdowsi, Z. (2008). Assessing readiness for business process reengineering. *Business Process Management Journal*, 14(4), 497-511.
- [3] Andersen, B. (2007). Business process improvement toolbox. Quality Press.
- [4] Bai, C., & Sarkis, J. (2014). A grey-based DEMATEL model for evaluating business process management critical success factors. *International Journal of Production Economics*, 146(1), 281-292.
- [5] Buh, B., Kovačić, A., & Indihar Štemberger, M. (2015). Critical success factors for different stages of business process management adoption—a case study. *Economic Research - Ekonomska Istraživanja*, 28(1), 243-258.
- [6] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429-444.
- [7] Chen, C.-H., & Lu, C.-L. (2011). Optimum profit model based on order quantity, product price, and process quality level. *Expert Systems with Applications*, 38(6), 7886–7893.
- [8] Čujić, M., Jovanović, M., Savić, G., & Jakšić, M. L. (2015). Measuring the efficiency of air navigation services system by using DEA method. *International Journal for Traffic and Transport Engineering*, 5(1), 36-44.
- [9] Färe, R., Grosskopf, S., & Lovell, K.C.A. (1985). The Measurement of Efficiency of Production. doi: 10.1007/978-94-015-7721-2
- [10] Hong, M., Steedle, J. T., & Cheng, Y. (2020). Methods of detecting insufficient effort responding: Comparisons and practical recommendations. *Educational and Psychological Measurement*, 80(2), 312-345.
- [11] Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, 27(1), 99-114.
- [12] Jovanović, I., Stojanović, D., Slović, D., Tomašević, I., Simeunović, B., & Maričić, M. (2021). Exploring the Key Factors of Process Improvement that Drive Competitive Advantage: A Case of Serbia. *International Conference on Business Process Management*, 179-191.
- [13] Kumar, V., Smart, P.A., Maddern, H., & Maull, R.S. (2008). Alternative perspectives on service quality and customer satisfaction: the role of BPM. *International Journal of Service Industry Management*, 19(2), 176-187.

- [14] Li, Y., Wang, Q., Wang, Z., & Chen, L. (2021). Improving business processes or human resources? The performance implications and contingencies. *Industrial Management & Data Systems*, 121(7), 1577-1598.
- [15] Lin, S.-M. (2010). Associating DEA with grey relational analysis for performance assessment of management competencies in logistic industry. *Journal of Statistics and Management Systems*, 13(5), 1117-1130.
- [16] Maddern, H., Maull, R., & Smart, P.A. (2007). Customer satisfaction and service quality in UK financial services. *International Journal of Operations & Production Management*, 27(9), 999-1019.
- [17] McCormack, K. et al. (2009). A global investigation of key turning points in business process maturity. *Business Process Management Journal*, 15(5), 792-815.
- [18] Nadarajah, D., & Kadir, A.S.L.S. (2014). A review of the importance of business process management in achieving sustainable competitive advantage. *The TQM Journal*, 26(5), 522-531.
- [19] Padilla, J. L., Castro, C., Doncel, P., & Taubman-Ben-Ari, O. (2020). Adaptation of the multidimensional driving styles inventory for Spanish drivers: Convergent and predictive validity evidence for detecting safe and unsafe driving styles. *Accident Analysis & Prevention*, 136, 105413.
- [20] Pastor, J.T., Ruiz, J.L., & Sirvent, I. (1999). An Enhanced DEA Russell Graph Efficiency Measure. *European Journal of Operational Research*, 115(3), 596-607.
- [21] Radojčić, M., Savić, G., Đoković, A., & Jeremić, V. (2017). Efikasnost i uspeh srednjih škola pri upisu na Fakultet. *XLIV International Symposium on Operational Research SYM-OP-IS 2017*, Zlatibor, Serbia, 48-53.
- [22] Savić, G., Radosavljević, M., & Ilievski, D. (2012). DEA Window Analysis Approach for Measuring the Efficiency of Serbian Banks Based on Panel Data. *Management: Journal for Theory and Practice Management*, 17(65).
- [23] Shirouyehzad, H., Lotfi, F.H., Aryanezhad, M.B., & Dabestani, R. (2012). A Data Envelopment Analysis Approach for Measuring Efficiency of Employees: a Case Study. *South African Journal of Industrial Engineering*, 23(1), 191-201.
- [24] Stojanović, D., Simeunović, B., Tomašević, I., & Jovanović, I. (2019). The Influence of Business Process Prioritization on Success of BPM adoption. *Smart Governments, Regions and Cities*, 235-258
- [25] Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European journal of operational research*, 130(3), 498-509.
- [26] Trihatmoko, R. A., Mulyani, R., & Lukviarman, N. (2018). Product placement strategy in the business market competition: studies of fast moving consumer goods. *Business and Management Horizon*, 6(1), 150-161.
- [27] Wolf, C., & Harmon, P. (2012). The State of business process management. Datum pristupa: 3.05.2022. [http://www.bptrends.com/bpt/wp-content/surveys/2012-\\_BPT%20SURVEY-3-12-12-CW-PH.pdf](http://www.bptrends.com/bpt/wp-content/surveys/2012-_BPT%20SURVEY-3-12-12-CW-PH.pdf).
- [28] Wong, W.P. (2013). Business-process management: a proposed framework for future research. *Total Quality Management and Business Excellence*, 24(5-6), 719-732.



## ANALIZA INDEKSA PERFORMANSI PROCESA U KOMPANIJAMA U SRBIJI

### ANALYSIS OF PROCESS PERFORMANCE INDEX IN COMPANIES IN SERBIA

DRAGANA STOJANOVIĆ<sup>1</sup>, IVAN TOMAŠEVIĆ<sup>2</sup>, BARBARA SIMEUNOVIĆ<sup>2</sup>, DRAGOSLAV SLOVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, dragana.stojanovic@fon.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, ivan.tomasevic@fon.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, barbara.simeunovic@fon.bg.ac.rs

<sup>4</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, dragoslav.slovic@fon.bg.ac.rs

**Rezime:** U ovom radu cilj je da se utvrdi u kojoj meri su kompanije u Srbiji uspešne u upravljanju poslovnim procesima, kao i da li kompanije koje su uspešnije u primeni procesnog pristupa ostvaruju bolje rezultate u pogledu povećanja konkurentnosti i sposobnosti kompanija. Korišćen je upitnik za prikupljanje podataka. Uspešnost u primeni procesnog pristupa je merena preko Indeksa performansi procesa (IPP) i Zrelosti procesa. Utvrđeno je da viši stepen IPP dovodi do većeg zadovoljenja potreba kupaca, kvalitetnijih proizvoda, pouzdanih isporuka i fleksibilnosti u ponudi proizvoda, što je u skladu sa fokusom na korisnike prilikom unapređivanja procesa. Procesni pristup ima i značajan uticaj na ostvarenje viših sposobnosti kompanije, a najviše na sticanje podrške zaposlenih za lin transformaciju, organizaciono učenje, sposobnost rešavanja problema i povećanja inovativnosti. Pored toga, istraživanje pokazuje i da manji opseg projekta unapređivanja procesa doprinosi uspešnijoj primeni procesnog pristupa. Rezultati u ovom radu daju stručnjacima iz prakse dalji pravac za razvoj primene procesnog pristupa u Srbiji.

**Ključne reči:** Upravljanje poslovnim procesima, Indeks performansi procesa, Unapređivanje procesa

**Abstract:** In this paper, the main goal is to determine success of BPM implementation in context of companies operating in Serbia, so as to determine whether the companies in Serbia with higher success of BPM implementation reach the better results in the companies competitiveness and capabilities. Questionnaire was used for data collection. BPM implementation was measured through Process Performance Index and Process Maturity. Results in this research shows Results showed that higher degree of PPI contribute to better customer satisfaction, better quality of products, reliability of delivery and flexibility in the product offer. Also, higher degree of BPM contribute to the higher companies capability, and most on the gaining employee support for lin transformation, organizational learning, problem solving ability and increasing innovation. In addition, research shows that the smaller scope of the process improvement project contributes to a more successful application of the process approach. Findings in this paper provide proposition for practitioners in companies operating in Serbia how to direct BPM practice for further development.

**Keywords:** Business Process Management, Process Performance Index, Process Improvement

## 1. UVOD

Sa visokim nivoom konkurentnosti na tržištu, kompanijama postaje sve važnije da poboljšaju performanse svojih procesa radi održavanja konkurentne prednosti [3]. Upravljanje poslovnim procesima (*Business Process Management* – BPM) je koncept koji se značajno razvio u poslednjih trideset godina i ima uticaj na povećanje efikasnosti i efektivnosti operacija u kompanijama [2]. Takođe, BPM može doprineti povećanju finansijskih i nefinansijskih performansi i smatra se dobrim rešenjem za kompanije u tranziciji [8]. Pored rastućeg trenda u primeni BPM-a, glavni izazov je kako implemetirati procesni pristup u operacije kompanija [9]. S obzirom da su konkurentna prednost i profit u fokusu zemalja u tranziciji, interesantno je analizirati uspeh primene BPM-a u kompanijama u Srbiji, kao odnos uspeha primene BPM-a i konkurentnosti i sposobnosti kompanija, kako bi se utvrdilo na koje elemente konkurentnosti i sposobnosti kompanija BPM najviše može da utiče.

## 2. METODOLOGIJA ISTRAŽIVANJA

U okviru ovog istraživanja postavljena su sledeća pitanja:

1) U kojoj meri su kompanije u Srbiji uspešne u upravljanju poslovnim procesima?

- 2) Da li kompanije koje su uspešnije u primeni procesnog pristupa ostvaruju bolje rezultate u pogledu povećanja konkurentnosti i sposobnosti kompanija?
- 3) Koji elementi konkurentnosti i sposobnosti su najviše poboljšani primenom BPM praksi?

Kako bi se odgovorilo na postavljena pitanja, potrebno je da se definiše način merenja uspeha primene BPM-a. U literaturi postoji nekoliko načina merenja uspešnosti primene BPM-a: modeli zrelosti procesa [9], Model zrelosti procesne orijentacije [5], i Indeks performansi procesa [7]. U ovom radu su za merenje uspeha primene BPM-a korišćena dva modela, i to Indeks performansi procesa i Zrelost procesa kroz Integrirani model zrelosti sposobnosti (*Capability Maturity Model Integration* - CMMI model) [10]. Obe mere su empirijski proverene, kvantitativne, i javno dostupne [1]. Kriterijumi i opis svakog kriterijuma koji čine Indeks performansi procesa su dati u tabeli 1.

**Tabela 1:** Kriterijumi i opis svakog kriterijuma koji čine Indeks performansi procesa [7]

Kriterijum	Opis
Usklađenost sa strategijom	Poslovni procesi su direktno povezani sa strategijom organizacije i kritičnim faktorima uspeha
Holistički pristup	Poslovni procesi su definisani pre iniciranja poboljšanja
Svesnost o procesnom pristupu od strane menadžmenta i zaposlenih	Ključni ljudi razumeju ulogu upravljanja procesima u poboljšanju performansi
Portfolio inicijativa poboljšanja procesa	Inicijative poboljšanja su prioritizovane prema „zdravlju“ procesa i vezom sa trenutnim izazovima
Metodologija poboljšanja procesa	Timovi za upravljanje poslovnim procesima koriste standardni pristup za analizu i projektovanje procesa
Metrika procesa	Performanse procesa se mere na pojedinačnom, nivou procesa i nivou kompanije
Fokus na korisnike	Napori unapređivanja poslovnih procesa su fokusirani na stvaranje vrednosti za korisnika
Upravljanje procesom	Menadžeri procesa svakodnevno prate performanse procesa i inicijative kontinualnog poboljšanja
Informacioni sistemi	Podrška koju pružaju automatizovane aplikacije je u skladu sa procesima koji su definisani u kompaniji Procesi su „gospodari“ a informacioni sistemi „sluge“
Upravljanje promenama	Prilikom promene poslovnih procesa efektivno se razmatraju problemi vezani za zaposlene i kulturu

Nivoi i opisi nivoa u CMMI modelu su dati u tabeli 2.

**Tabela 2:** Zrelost procesa prema CMMI modelu [10]

Nivo zrelosti	Opis nivoa zrelosti procesa
Nivo 1	Procesi su ad hoc, pojedine aktivnosti su eksplicitno definisane i uspeh zavisi od zalaganja pojedinca
Nivo 2	Definisane su procedure koje imaju za cilj održavanje postojeće dobre prakse, ima dovoljno discipline da se ponove raniji uspesi
Nivo 3	Procesi su definisani i dokumentovani sa velikim stepenom detaljnosti, i izvode se konzistentno u čitavom poslovnom sistemu, varijacije u izvršavanju procesa su dozvoljene uz poštovanje definisanih procedura
Nivo 4	Uspostavljene su performanse procesa na osnovu zahteva korisnika, i koriste se za upravljanje procesima u čitavoj kompaniji, odvijanje procesa je predvidivo i kontrolisano
Nivo 5	Procesi se kontinualno unapređuju na osnovu kvantitativnih performansi, kontinualno unapređenje je zadatak svakog zaposlenog u kompaniji

Kako bi se utvrdili elementi konkurentnosti i sposobnosti kompanija, korišćena su prethodna istraživanja, gde u varijable konkurentnosti i sposobnosti empirijski potvrđene u istraživanju [4] i date su u tabeli 3. Za promenljive koje se odnose na konkurentnosti i sposobnost kompanija koristila se Likertova skala.

Kako bi se odgovorilo na postavljena istraživačka pitanja, sprovedeno je istraživanje korišćenem upitnika. Upitnik je sadržao 28 pitanja koja su bila podeljena u tri grupe: (1) opšta pitanja o kompaniji i ispitanicima i razumevanju BPM-a, (2) pitanja o Zrelosti procesa i Indeksu performansi procesa koja su adaptirana prema BPTrends survey [11], (3) pitanja o konkurentnosti i sposobnosti kompanija nakon primene BPM-a preuzeta iz Process Excellence Network survey [6].

**Tabela 3:** Varijable konkurentnosti i sposobnosti kompanija [4]

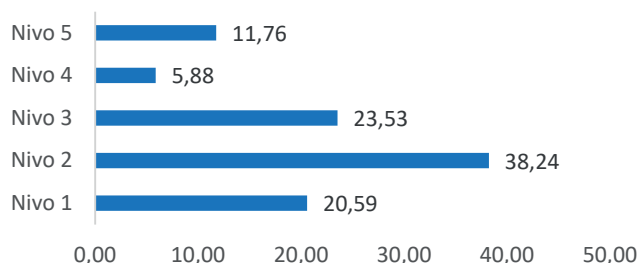
Varijabla konkurentnosti	Varijabla sposobnosti
Nudimo konkurentne cene	Sposobnost rešavanja problema
U mogućnosti smo da ponudimo niske cene ili cene niže od konkurencije	Sposobnost brzog reagovanja na nove prilike na tržištu
U mogućnosti smo da se takmičimo po kvalitetu sa konkurencijom	Sposobnost brzog reagovanja na izazove
Nudimo proizvode koji imaju visoku pouzdanost	Fleksibilnost usvajanja nepredviđene promene
Nudimo visokokvalitetne proizvode našim potrošačima	Sposobnost identifikovanja novih poslovnih prilika
Pravovremeno zadovoljavamo potrebe kupaca	Sposobnost prepoznavanja ključnih problema
Naše isporuke su puzdane po kvalitetu i kvantitetu	Sposobnost podsticanja organizacionog učenja
Prilagođavano ponudu naših proizvoda kako bismo zadovoljili potrebe klijenata	Sposobnost sticanja podrške zaposlenih za lin transformaciju
Uspešno odgovaramo na tražnju naši potrošača za novim proizvodima	Sposobnost povećanja inovativnosti
Prvi plasiramo nove proizvode na tržište	Sposobnost dostizanja konkurentne prednosti
Vreme plasiranja na tržište je kraće od proseka u industriji	

Istraživanje je uključivalo kompanije iz Srbije, a ispitanici su bili menadžeri, izvršni direktori ili analitičari procesa. Populacija je uključivala kompanije koje su primenile procesni pristup i bile sertifikovane po ISO standardu. Početna lista ispitanika je dobijena od strane Privredne komore Srbije, gde je bilo 500 kompanija za koje je bio dostupan mail kontakt. Dobijeno je ukupno 61 upotrebljivih odgovora, što predstavlja odziv od 12.2 procenata, koji je adekvatan za istraživanje ovog tipa [12].

Kako bi se analizirali rezultati, korišćen je paket SPSS, kao i *Spearman-Rho* test korelacije kako bi se utvrdilo da li postoji korelacija između Indeksa performansi procesa i varijabli konkurentnosti, kao i varijabli sposobnosti kompanija.

### 3. REZULTATI I DISKUSIJA

Proizvodni poslovni sistemi su činili 45,03% uzorka, uslužni 40,6% uzorka, a 14,1% uzorka su bili zastupljeni proizvodno – uslužni poslovni sistemi. Ocena zrelosti procesa po nivoima je prikazana na slici 1.

**Slika 1:** Zrelost procesa u [%]

Više od trećine kompanija je na drugom nivou zrelosti, a 23,53% je na trećem nivou zrelosti, dok se 11,76% ispitanika izjasnilo da su na najvišem nivou zrelosti procesa.

Ispitanicima je postavljeno pitanje vezano za projekte unapređivanja procesa, i to opseg projekata i dužinu projekata (tabela 4).

**Tabela 4:** Projekti unapređivanja procesa – opseg i dužina

Opseg projekata	%	Dužina projekata	%
Primenjuje se širom kompanije	35,29	Manje od 3 meseca	8,82
Primenjuje se u jednoj ili više poslovnih jedinica	38,24	3 - 6 meseci	20,59
Mali pilot projekti	8,82	6 - 12 meseci	32,35
Obučeni su menadžeri, ali nema formalnih programa	17,65	Više od 1 godine	11,76
		Nije mi poznato	26,47

Najveći broj kompanija u Srbiji radi na projektima unapređivanja procesa koji se primenjuju u jednoj ili više poslovnih jedinica (38,24%) ili širom kompanije (35,29%), gde projekti u najvećoj meri traju između 6-12 meseci (32,35%), odnosno 3-6 meseci (20,59%).

U nastavku su date analize vezane za Indeks performansi procesa. U tabeli 5. je prikazan broj kompanija koje su po kriterijumima za Indeks performansi procesa dale ocene od 1 do 5.

**Tabela 5:** Sumarna statistika Indeksa performansi procesa

Kriterijum	Opis	1	2	3	4	5
Usklađenost sa strategijom	Poslovni procesi su direktno povezani sa strategijom organizacije i kritičnim faktorima uspeha	0	11	12	24	14
Holistički pristup	Poslovni procesi su definisani pre iniciranja poboljšanja	0	11	16	26	8
Svesnost o procesnom pristupu od strane menadžmenta i zaposlenih	Ključni ljudi razumeju ulogu upravljanja procesima u poboljšanju performansi	2	13	20	20	6
Portfolio inicijativa poboljšanja procesa	Inicijative poboljšanja su prioritetizovane prema „zdravlju“ procesa i vezom sa trenutnim izazovima	1	8	13	30	9
Metodologija poboljšanja procesa	Timovi za upravljanje poslovnim procesima koriste standardni pristup za analizu i projektovanje procesa	2	11	16	24	8
Metrika procesa	Performanse procesa se mere na pojedinačnom, nivou procesa i nivou kompanije	0	3	16	22	20
Fokus na korisnike	Napori unapređivanja poslovnih procesa su fokusirani na stvaranje vrednosti za korisnika	1	6	20	27	7
Upravljanje procesom	Menadžeri procesa svakodnevno prate performanse procesa i inicijative kontinualnog poboljšanja	1	12	21	17	10
Informacioni sistemi	Procesi su „gospodari“ a informacioni sistemi „sluge“	2	7	18	21	13
Upravljanje promenama	Prilikom promene poslovnih procesa efektivno se razmatraju problemi vezani za zaposlene i kulturu	1	12	16	23	9

Indeks performansi procesa je izračunat na standardni način za svaku kompaniju kao suma ocena po svakom kriterijumu [7]. Vrednosti PPI se kreću od 10 do 50 za svaku kompaniju. Tabela 6. predstavlja deskriptivnu statistiku za PPI.

**Tabela 6:** Indeks performansi procesa – deskriptivna statistika

	N	Min	Max	Sredina	Std. devijacija
PPI	61	17	50	35,38	7,450

Frekvencija i procenat kompanija sa određenom vrednošću PPI je data u tabeli 7.

**Tabela 7:** Indeks performansi procesa – deskriptivna statistika

PPI vrednost	Frekvencija	Procenat	Kumulativni procesant
17	1	1,6	1,6
22	1	1,6	3,3
24	4	6,6	9,8
25	1	1,6	11,5
26	2	3,3	14,8
27	3	4,9	19,7
30	3	4,9	24,6
31	5	8,2	32,8
32	1	1,6	34,4
33	6	9,8	44,3
34	2	3,3	47,5
36	2	3,3	50,8
37	2	3,3	54,1
38	5	8,2	62,3
39	4	6,6	68,9
40	3	4,9	73,8
41	3	4,9	78,7
42	2	3,3	82,0
43	2	3,3	85,2
44	2	3,3	88,5
45	2	3,3	91,8
46	1	1,6	93,4
47	1	1,6	95,1
48	2	3,3	98,4
50	1	1,6	100,0
<b>Ukupno</b>	<b>61</b>	<b>100,0</b>	

Tabela 7. pokazuje da se vrednost PPI kreće između 30 i 45 za najveći broj kompanija, što ukazuje na srednje visoku vrednost PPI. Analizirana je korelacija između PPI i zrelosti procesa, kao i opsega i dužine projekata unapređenja procesa (tabela 8).

**Tabela 8:** Spearman's Rho korelacija između PPI, zrelosti procesa, opsega i dužine projekata

Varijabla	Koeficijent korelacije	Znač (2-tailed)
Zrelost procesa	0,643**	0,000
Opseg projekta unapređivanja procesa	-0,541**	0,000
Dužina projekta unapređivanja procesa	-0,274**	0,000
**Korelacija je značajna na nivou 0,01		

Visoka korelacija PPI i zrelosti procesa ukazuje na pouzdanost rezultata istraživanja. Pored toga, može se zaključiti da su kompanije sa manjim opsegom projekata bile uspešnije u primeni procesnog pristupa, dok jesa druge strane zaključeno da dužina projekata ne utiče značajno usvajanje procesnog pristupa. S obzirom da je  $r > 0,5$  predstavlja visok nivo korelacije, tako da postoji jaka, pozitivna korelacija između dve varijable [ $r = 0,643$ ,  $n = 61$ ,  $P = 0,000$ ], dok negativna korelacija pokazuje da postoji jaka, obrnuto proporcionalna korelacija između dve varijable [ $r = -0,541$ ,  $n = 61$ ,  $P = 0,000$ ], dok  $r = -0,274$  predstavlja nizak nivo korelacije, tako da ne postoji korelacija između dve varijable [ $r = -0,274$ ,  $n = 61$ ,  $P = 0,000$ ].

Analizirana je i korelacija između PPI i konkurentnosti i sposobnosti kompanije i rezultati su prikazani u tabelama 9. i 10.

**Tabela 9:** Spearman's Rho korelacija između PPI i varijabli konkurentnosti kompanije

Varijabla konkurentnosti	Sredina	Std. dev.	Koef. korelacije	Znač. (2-tailed)
Nudimo konkurentne cene	4,13	1,087	0,263*	0,041
U mogućnosti smo da ponudimo niske cene ili cene niže od konkurencije	4,03	1,278	0,289	0,064
U mogućnosti smo da se takmičimo po kvalitetu sa konkurencijom	4,36	0,857	0,503**	0,000
Nudimo proizvode koji imaju visoku pouzdanost	4,21	1,035	0,486**	0,000
Nudimo visokokvalitetne proizvode našim potrošačima	4,18	0,806	0,540**	0,000
Pravovremeno zadovoljavamo potrebe kupaca	3,93	0,998	0,543**	0,000
Naše isporuke su puzdane po kvalitetu i kvantitetu	4,13	0,957	0,511**	0,000
Prilagođavano ponudu naših proizvoda kako bismo zadovoljili potrebe klijenata	4,26	0,947	0,518**	0,000
Uspešno odgovaramo na tražnju naši potrošača za novim proizvodima	4,18	1,041	0,422**	0,001
Prvi plasiramo nove proizvode na tržište	3,75	1,362	0,123	0,346
Vreme plasiranja na tržište je kraće od proseka u industriji	3,87	1,408	0,142	0,275

Kompanije koje imaju viši PPI su i konkurentnije u pogledu Pravovremenog zadovoljavanja potreba kupaca, Ponude visokokvalitetnih proizvoda potrošačima, Prilagođavanja ponude proizvoda kako bi zadovoljili potrebe klijenata i Pouzdanosti isporuka po kvalitetu i kvantitetu.

**Tabela 10:** Spearman's Rho korelacija između PPI i varijabli sposobnosti kompanije

Varijabla sposobnosti	Sredina	Std. dev.	Koef. korelacije	Znač (2-tailed)
Sposobnost rešavanja problema	3,21	1,002	0,607**	0,000
Sposobnost brzog reagovanja na nove prilike na tržištu	3,15	0,963	0,591**	0,000
Sposobnost brzog reagovanja na izazove	3,36	1,001	0,517**	0,000
Fleksibilnost usvajanja nepredviđene promene	3,20	0,928	0,554**	0,000
Sposobnost identifikovanja novih poslovnih prilika	3,33	1,091	0,541**	0,000
Sposobnost prepoznavanja ključnih problema	3,51	0,924	0,579**	0,000
Sposobnost podsticanja organizacionog učenja	3,08	1,053	0,630**	0,000
Sposobnost sticanja podrške zaposlenih za lin transformaciju	2,80	1,077	0,660**	0,000
Sposobnost povećanja inovativnosti	3,25	1,075	0,600**	0,000
Sposobnost dostizanja konkurentne prednosti	3,30	1,116	0,558**	0,000

Na osnovu rezultata se može zaključiti da kompanije koje imaju viši PPI takođe ostvaruju i više sposobnosti, gde se najviše ističu Sposobnost sticanja podrške zaposlenih za lin transformaciju, Sposobnost podsticanja organizacionog učenja, Sposobnost rešavanja problema i Sposobnost povećanja inovativnosti.

#### 4. ZAKLJUČAK

Indeks performansi procesa u kompanijama u Srbiji ima srednje visoku vrednost i u potpunosti je u korelaciji sa Zrelošću procesa koja je merena prema CMMI modelu, što ukazuje da kompanije u Srbiji najviše rade na dokumentovanju procesa i izradi procedura, prioritizaciji procesa prema „zdravlju“ procesa i vezom sa trenutnim izazovima, kao i unapređenjima fokusiranim na korisnike. Pri realizaciji projekata unapređivanja procesa, utvrđeno je da su kompanije sa manjim opsegom projekata bile uspešnije u primeni procesnog pristupa, dok je sa druge strane zaključeno da dužina projekata ne utiče značajno usvajanje procesnog pristupa. Pored navedenog, utvrđeno je da primena i viši stepen usvajanja procesnog pristupa dovodi do većeg zadovoljenja potreba kupaca, kvalitetnijih proizvoda, pouzdanih isporuka i fleksibilnosti u ponudi proizvoda, što je u skladu sa fokusom na korisnike prilikom unapređivanja procesa. Procesni pristup ima i značajan uticaj na ostvarenje viših sposobnosti kompanije i to na: Sposobnost sticanja podrške zaposlenih za lin transformaciju, Sposobnost podsticanja organizacionog učenja, Sposobnost rešavanja problema i Sposobnost povećanja inovativnosti.

#### LITERATURA

- [1] Hribar, B., & Medling, J. (2014). The Correlation of organizational culture & BPM adoption success. *In Twenty Second European Conference on Information Systems* (pp. 1- 16). Recanati Business School Tel Aviv University, Tel Aviv. [http://dx.doi.org/10.1207/s15327752jpa5701\\_17](http://dx.doi.org/10.1207/s15327752jpa5701_17)
- [2] Indihar Štemberger, M., Buh, B., Milanović Glavan, L., & Mendling, J. (2013) Propositions on the interaction of organizational culture with other factors in the context of BPM adoption. *Business Process Management Journal* 24(2), 425-445. <https://doi.org/10.1108/BPMJ-02-2017-0023>
- [3] Jovanović, I., Stojanović, D., Tomašević, I., Simeunović, B. & Slović, D. (2021). Uticaj BPM-a na sposobnost kompanija u Srbiji. *XIII Skup privrednika i naučnika SPIN '21 – Industrija 4.0 – mogućnosti, izazovi i rešenja za digitalnu transformaciju privrede. Zbornik radova*, Fakultet organizacionih nauka, Beograd, str. 84-91. ISBN 978-86-7680-396-5
- [4] Marshall, D.A. (2014). Lean Transformation: Overcoming The Challenges, Managing Performance, And Sustaining Success. *Theses and Dissertations--Marketing and Supply Chain*
- [5] McCormack, K., & Johnson, W. (2001). Business process orientation: gaining the ebusiness competitive advantage. St. Lucie Press, Florida.
- [6] Process Excellence Network (2012). Trends and Success Factors in Business Process Excellence. internet]. Available at:< [http://www. processexcellencenetwork. com/lean-six-sigma-business-transformation/white-papers/trendsand-success-factors-in-business-process](http://www.processexcellencenetwork.com/lean-six-sigma-business-transformation/white-papers/trendsand-success-factors-in-business-process)>[Accessed: May 10, 2013].
- [7] Rummler-Brache Group (2004), <https://www.rummlerbrache.com/process-performance-index> last accessed: 2018/11/27.
- [8] Stojanović, D., Tomašević, I., Slović, D., Gošnik, D., Suklan, J., Kavčič, K. (2017) Bpm In Transition Economies: Joint Empirical Experience Of Slovenia And Serbia, *Economic Research-Ekonomska Istraživanja*, 30 (1), 1237-1256. <https://doi.org/10.1080/1331677X.2017.1355256>
- [9] Škrinjar, R., & Trkman, P. (2013). Increasing process orientation with business process management: Critical practices. *International Journal of Information Management*, 33(1), 48-60. <https://doi.org/10.1016/j.ijinfomgt.2012.05.011>
- [10] Team, C. P. (2002). Capability maturity model® integration (CMMI SM), version 1.1. CMMI for systems engineering, software engineering, integrated product and process development, and supplier sourcing (CMMI-SE/SW/IPPD/SS, V1. 1), 2.
- [11] Wolf, C., & Harmon, P. The State of business process management (2012), . BPTrends HomePage, [http://www.bptrends.com/bpt/wp-content/surveys/2012-\\_BPT%20SURVEY-3-12-12-CW-PH.pdf](http://www.bptrends.com/bpt/wp-content/surveys/2012-_BPT%20SURVEY-3-12-12-CW-PH.pdf) last accessed: 2018/12/26
- [12] Wong, W., Ahmad, N., Nasurdin, A., & Mohamad, M.: The impact on external environmental on business process management and organizational performance. *Service Business* 8 (4), 559-586 (2014). <https://doi.org/10.1007/s11628-013-0207-9>

## ANALIZA POKAZATELJA PERFORMANSI UPRAVLJANJA PRIHODIMA U RESTORANIMA

### THE ANALYSIS OF THE REVENUE MANAGEMENT PERFORMANCE INDICATORS IN THE RESTAURANTS

BRANISLAVA HRISTOV STANČIĆ<sup>1</sup>, BOJAN ZEČEVIĆ<sup>2</sup>, IGOR KOVAČEVIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, branislava.hristov@ekof.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, bojan.zecevic@ekof.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Ekonomski fakultet, igor.kovacevic@ekof.bg.ac.rs

**Rezime:** Restoratersku industriju karakteriše upravljanje relativno fiksnim kapacitetom, čija se upotreba ne može odložiti, kao i varijabilnom tražnjom koja se jasno može segmentisati, što ovu industriju čini odličnim kandidatom za primenu koncepta upravljanja prihodima. Menadžeri restorana se svakodnevno suočavaju sa pitanjem maksimalnog iskorišćavanja kapaciteta prilikom pokušaja da povećaju prihode, pri čemu strategije upravljanja prihodima mogu imati veoma bitnu ulogu. Upravljanje prihodima u restoranima ima za cilj da odredi najefikasnije načine uravnoteženja tražnje i ponude restorana tako da se prihod maksimizira, ali bez pojave nezadovoljnih kupaca. Rezultati sprovedenih strategija upravljanja prihodima se mogu oceniti kroz analizu specifičnih operativnih pokazatelja performansi restorana. Najznačajniji pokazatelji performansi za merenje operativne uspešnosti rada restorana su: prosečan račun, prihod po raspoloživom mestu po satu, profit po raspoloživom mestu po satu i profit po raspoloživom kvadratnom metru. Cilj ovog rada jeste da se proceni koliko uspešno određeni pokazatelji performansi upravljanja prihodima doprinose analizi poslovanja restorana. Rezultati sprovedene komparativne analize prihoda po raspoloživom mestu po satu i profita po raspoloživom mestu po satu na bazi hipotetičkog primera restorana, pokazuju da profit po raspoloživom mestu po satu daje preciznije podatke uspešnosti rada restorana. Međutim, da bi se sprovela sveobuhvatna analiza uspešnosti primene upravljanja prihodima u restoranima, potrebno je sve dostupne pokazatelje posmatrati integralno. Kako svi oni imaju određene nedostatke, očekuje se da će se u budućnosti naučna literatura detaljnije baviti ovim pitanjem.

**Ključne reči:** Restorani, Upravljanje prihodima, Pokazatelji operativnih performansi

**Abstract:** The business characteristics of restaurants such as relatively fixed capacity, product perishability, as well as unstable demand that can be forecasted and segmented, makes this industry an excellent candidate for the application of the revenue management concept. Restaurant managers face the challenge how to maximize daily occupancy rate in order to increase revenue of the restaurant. Therefore, revenue management strategies play a very important role. Restaurant revenue management aims to determine the most efficient management tools to balance restaurant demand and supply, in order to maximize revenue without the appearance of dissatisfied customers. The results of the implemented revenue management strategies can be measured through the analysis of specific restaurant operational performance indicators. The most significant performance indicators for measuring the operational performance of restaurants are: average check, revenue per available seat hour, profit per available seat hour and profit per available square meter. The aim of this paper is to evaluate how successfully certain revenue management performance indicators contribute to the analysis of restaurant business performance. The results of the comparative analysis of revenue per available seat hour and profit per available seat hour based on a hypothetical example of a restaurant, determined that the profit per available seat hour provides more accurate information on the success of the restaurant. However, a comprehensive analysis of successful application of restaurant revenue management can be achieved if all these indicators are observed integrally. As they all have certain shortcomings, it is expected that in the future the scientific literature in more detail will deal with this topic.

**Keywords:** Restaurants, Revenue management, Operational performance indicators

## 1. UVOD

Upravljanje prihodima je koncept koji je zaživeo nakon deregulacije avio saobraćaja 80-ih godina 20. veka. Od tada počinje njegova primena u mnogim industrijama koje se pre svega suočavaju sa ograničenim

kapacitetima, visokim fiksnim troškovima, predvidivom i promenljivom tražnjom, kao i cenovno osetljivim potrošačima. Značaj koncepta upravljanja prihodima u restoraterskoj industriji uočen je krajem 20. veka [7], ali restoranski menadžeri do danas nisu u potpunosti prihvatili njegovu primenu.

Sektor za pružanje usluge hrane i pića (*engl. food and baverage-F&B*) u okviru hotela dugo je ostvarivao manje zarade od željenih. Maksimizacija prihoda ovog sektora postignuta je putem franšize, jer je upravljanje restoranima prepušteno profesionalnim timovima. Drugo potencijalno rešenje predstavljala je primena koncepta upravljanja prihodima [10].

Merenje uspešnosti rada sektora hrane i pića zahtevalo je definisanje specifičnih pokazatelja performansi, što predstavlja prvi korak u procesu napretka i ostvarivanja boljih poslovnih rezultata.

U tom smislu, cilj ovog rada jeste analiza preciznosti merenja uspešnosti rada restorana na bazi odgovarajućih pokazatelja performansi upravljanja prihodima. Iz cilja rada proističe glavno istraživačko pitanje koje će biti obrađivano u radu, a koje glasi:

Da li se prilikom analize uspešnosti rada restorana, menadžeri prihoda mogu oslanjati samo na jedan odabrani pokazatelj upravljanja prihodima ili je potrebno da istovremeno posmatraju veći broj ovih pokazatelja?

Da bi se dobio odgovor na ovo pitanje u radu je sprovedena komparativna analiza dva pokazatelja performansi upravljanja prihodima, prihoda po raspoloživom mestu po satu i profita po raspoloživom mestu po satu. U radu su korišćeni hipotetički podaci za restoran X kako bi se na ilustrativan način prikazale razlike između ova dva pokazatelja.

## 2. UPRAVLJANJE PRIHODIMA U RESTORANIMA

U naučnoj literaturi tema upravljanja prihodima u restoranima je postala aktuelna tek krajem 20. veka [5,7]. Restorani se svakodnevno suočavaju sa pitanjem maksimalnog iskorišćavanja kapaciteta prilikom pokušaja da povećaju prihode, jer nezauzeta sedišta restorana predstavljaju izgubljeni prihod [1]. Principi upravljanja prihodima mogu se primeniti u restoranima s obzirom na to da je prodajna jedinica u restoranima vreme potrebno za pružanje usluge, a ne samo sam obrok [5]. Iako restorani imaju veći procenat varijabilnih troškova od avio-kompanija i hotela, potencijalni rast prihoda usled primene koncepta upravljanja prihodima, može biti veoma visok [8]. Kao i kod drugih industrija, industrija restorana ima svoje posebne karakteristike, pa bi stoga pokazatelji koje se koriste za procenu njihove operativne uspešnosti trebalo da odražavaju specifična pitanja s kojima se suočavaju. U prošlosti su mnogi restorani ili definisali visoke stope zauzetosti stolova ili visoke prosečne račune, kao svoje primarne poslovne ciljeve [2]. Međutim, menadžeri restorana bi trebalo da imaju na umu da cilj upravljanja prihodima jeste istovremeno postizanje najvećeg mogućeg doprinosa za restoran i najveće vrednosti ili korisnosti za korisnika usluga [9]. Uspešna strategija upravljanja prihodima zasniva se na efikasnoj kontroli tražnje kupaca. Takođe, menadžeri restorana mogu manipulirati sa dve glavne strateške poluge za upravljanje prihodima, a to su cena i trajanje obroka [6]. Cena je prilično očigledan način za manipulaciju. Mnogi restorani nude usluge po promotivnim cenama u periodima niske tražnje, dok su u periodima visoke tražnje cene znatno više, kako bi se kupci koji su cenovno osetljiviji preusmerili na periode kada se kapaciteti teže popunjavaju. Promotivne cene važe za određene stavke menija u tačno određenom delu dana. Sofisticiranije manipulacije cenama uključuju različite cene u toku dana, kao i dana u sedmici, odobravanje popusta za određenu poziciju i veličinu stola.

## 3. KLJUČNI INDIKATORI PERFORMANSI UPRAVLJANJA PRIHODIMA U RESTORANIMA

Indikatori performansi igraju ključnu ulogu kako u evaluaciji ostvarenih rezultata, tako i u planiranju budućih odluka. I pored toga, naučna literatura se do sada nedovoljno bavila diskusijama o merenju uspešnosti rada restorana. U prošlosti restorani su najčešće koristili pokazatelj prosečnog računa (*engl. average check*) kako bi izmerili ostvarenost postavljenih ciljeva. Prosečan račun pokazuje kolika je prosečna potrošnja svakog gosta u restoranu. Ovaj pokazatelj se može izračunati po sledećoj formuli:

$$\text{Prosečan račun} = \frac{\text{Ukupan prihod restorana}}{\text{Ukupan broj posluženih gostiju}} \quad (1)$$

Pružanje usluge u restoranima zavisi od veličine restorana, tj. od površine na kojoj se ta usluga realizuje, te je opravdano uvođenje pokazatelja profita po raspoloživom kvadratnom metru (*engl. profit per available square meter- ProPASM*). Menadžeri restorana teže da maksimizuju profit svakog kvadratnog metra za svaki vremenski period u kojem je prostor dostupan. ProPASM pokazuje koliko uspešno restoran upravlja prostorom



iz profitne perspektive. Ovaj pokazatelj se može izračunati stavljanjem u odnos ukupnog profita svakog sata sa ukupno raspoloživim prostorom merenim u kvadratnim metrima u istom tom periodu [2].

Računanje ovog pokazatelja može se izvršiti po sledećoj formuli:

$$\begin{aligned} \text{Profit po raspoloživom kvadratnom metru} &= \frac{\sum(\text{Cena jela} - \text{Trošak jela})}{\text{Ukupno raspoloživ prostor po kvadratnom metru}} & (2) \\ &= \frac{\sum \text{Ukupna kontribuciona marža}}{\text{Ukupno raspoloživ prostor po kvadratnom metru}} \end{aligned}$$

Do sada, ovaj pokazatelj nije dovoljno istraživan u naučnoj literaturi, pa se očekuju dodatne studije na tu temu.

U naučnoj literaturi kao i današnjoj praksi, najčešće diskutovan i korišćen pokazatelj uspešnosti rada restorana je prihod po raspoloživom mestu po satu (*engl. revenue per available seat hour- RevPASH*) [4,5,7]. Ovaj pokazatelj se izračunava prema sledećoj formuli:

$$\text{Prihod po raspoloživom mestu po satu} = \frac{\text{Ukupan restoranski prihod}}{\text{Broj raspoloživih mesta}} \quad (3)$$

RevPASH pokazuje kako stopa iskorišćenosti kapaciteta doprinosi ostvarivanju prihoda restorana. Rast ovog pokazatelja obezbeđuje se kroz porast obrta stolova, tj. kroz rast broja gostiju koji se smene za istim stolom, kao i kroz smanjenje vremena trajanja obroka. Tačnije, što je kraće vreme trajanja obroka, više klijenata može da bude posluženo za istim stolom tokom određenog vremenskog perioda [2]. Takođe, njegov rast obezbeđen je i kroz rast prosečnog računa po gostu [4]. RevPASH nudi uvid u to koliko je korisnika restorana opsluženo, kao i koliko brzo im je pružena usluga, pa stoga ovaj pokazatelj pomaže u proceni ne samo efikasnosti, već i efektivnosti pružene usluge [2]. Međutim, precizna analiza uspešnosti rada restorana u domenu profitabilnosti ne može se vršiti na osnovu ovog pokazatelja.

Dobar pokazatelj operativne profitabilnosti restorana je profit po raspoloživom mestu po satu (*engl. profit per available seat hour- ProPASH*). Ovaj pokazatelj uzima u obzir maržu doprinosa svake stavke menija, a ne ukupno generisani prihod. Bitno je naglasiti da se fiksni trošak ne posmatra, jer se pretpostavlja njegova ravnomerna raspodela po svakom raspoloživom kapacitetu [2].

Računanje ovog pokazatelja može se izvršiti po sledećoj formuli:

$$\begin{aligned} \text{Profit po raspoloživom mestu po satu} &= \frac{\sum(\text{Cena jela} - \text{Trošak jela})}{\text{Broj raspoloživih mesta}} & (4) \\ &= \frac{\sum \text{Ukupna kontribuciona marža}}{\text{Broj raspoloživih mesta}} \end{aligned}$$

Kako bi bolje razumeli razliku između RevPASH-a i ProPASH-a obrađen je hipotetički primer restorana X koji raspolaže sa ukupno 200 mesta. U obzir su uzeta samo ova dva pokazatelja, kako su oni međusobno uporedivi, jer u imeniocu uzimaju istu jedinicu posmatranja u smislu broja raspoloživih mesta [3].

U tabeli 1 i tabeli 2 prikazane su sledeće veličine: ukupan prihod (TR), stopa popunjenosti stolova (Occ %), prosečan račun (AC), cena (P), količina (Q), trošak (C), kontribuciona marža (CM), ukupna kontribuciona marža (TCM), RevPASH i ProPASH.

**Tabela 1.** Analiza pokazatelja RevPASH u restoranu X

<i>Vreme</i>	<i>Ponedeljak</i>			<i>Subota</i>			<i>RevPASH</i>
	<i>TR</i>	<i>Occ %</i>	<i>AC</i>	<i>TR</i>	<i>Occ %</i>	<i>AC</i>	
10-11h	1200\$	40%	15\$	1200\$	30%	20\$	6\$
11-12h	2500\$	50%	25\$	2500\$	62,5%	20\$	12,5\$
12-13h	1200\$	30%	20\$	1200\$	24%	25\$	6\$
13-14h	4800\$	80%	30\$	4800\$	60%	40\$	24\$
<i>Ukupno/Prosek</i>	9700\$	50%	22,5\$	9700\$	44,125%	26,25\$	48,5\$

Izvor: Hristov Stančić, B. (2021). *Primena koncepta upravljanja prihodima u analizi poslovanja hotela sa stanovišta operativnih performansi*, Doktorska disertacija, Ekonomski fakultet, Univerzitet u Beogradu.

U tabeli 1 je napravljena komparacija poslovanja restorana X u ponedeljak i subotu, tako da je i jednog i drugog dana ostvaren identičan ukupan prihod u iznosu od 9700\$. Takođe, radi boljeg razumevanja prikazano je da je restoran X i jednog i drugog dana u svakom vremenskom intervalu generisao isti iznos prihoda, ali su se popunjenost kapaciteta, kao i prosečna potrošnja po svakom gostu tj. prosečan račun, razlikovali za svaki interval vremena [3]. Iako [2] tvrdi da RevPASH nudi uvid u to koliko je korisnika restoranske usluge opsluženo, kao i koliko brzo im je pružena usluga, te da stoga ovaj pokazatelj omogućava da se evaluiira efikasnost i efektivnost rada restorana, na osnovu datog primera mi vidimo da on uzima identične vrednosti za oba dana u istom vremenskom intervalu, jer se prilikom njegovog izračunavanja u obzir uzimaju ukupni prihodi i broj raspoloživih mesta u restoranu, bez obzira na dinamiku popunjavanja kapaciteta, količinu novca koju svaki gost potroši i sadržaj usluge koju konzumira. Kako bi se prevazišli ovi nedostaci, menadžeri prihoda restorana ne bi trebalo da ograniče svoju analizu samo na pomenuti pokazatelj [3].

**Tabela 2.** Komparacija pokazatelja RevPASH i ProPASH u restoranu X

<i>vreme</i> (10-11h)		<i>P</i>	<i>Q</i>	<i>C</i>	<i>CM</i>	<i>TR</i>	<i>TCM</i>	<i>Occ %</i>	<i>AC</i>	<i>RevPASH</i>	<i>ProPASH</i>
<i>Ponedeljak</i>	Supa	10\$	20	2\$	8\$	200\$	160\$	40%	15\$	6\$	4,675\$
	Piletina	20\$	25	5\$	15\$	500\$	375\$				
	Riba	25\$	16	5\$	20\$	400\$	320\$				
	Kafa	5\$	20	1\$	4\$	100\$	80\$				
	<i>Ukupno</i>		81			1200\$	935\$				
<i>Subota</i>	Piletina	20\$	15	5\$	15\$	300\$	225\$	30%	20\$	6\$	4,425\$
	Biftek	30\$	15	8\$	22\$	450\$	330\$				
	Torta	10\$	15	4\$	6\$	150\$	90\$				
	Kafa	5\$	60	1\$	4\$	300\$	240\$				
	<i>Ukupno</i>		105			1200\$	885\$				

Izvor: Hristov Stančić, B. (2021). *Primena koncepta upravljanja prihodima u analizi poslovanja hotela sa stanovišta operativnih performansi*, Doktorska disertacija, Ekonomski fakultet, Univerzitet u Beogradu.

U tabeli 2 napravljeno je poređenje pokazatelja RevPASH-a i ProPASH-a. Restoran X generiše identičan prihod u dva različita dana, ali se taj prihod generiše na različite načine tj. stavke menija se razlikuju. U tom slučaju RevPASH će i dalje biti identičan, dok će se ProPASH razlikovati. Tačnije, restoran X ostvaruje manji iznos profita po raspoloživom mestu po satu u subotu nego u ponedeljak, upravo zbog različite strukture potrošnje što implicira i različitu strukturu troškova [3]. Može se zaključiti da veći RevPASH ne znači uvek i

veći profit [2]. Prethodna analiza pokazala je da ProPASH pokriva određene nedostatke koje sa sobom nosi merenje na bazi RevPASH-a, pa je preporuka da se u analizi u obzir uzmu oba pokazatelja [3].

U naučnoj literaturi su trenutno dostupna samo prethodno pomenuta četiri pokazatelja upravljanja prihodima u restoranima. Kako bi se dobila što preciznija slika uspešnosti rada restorana, trebalo bi ih posmatrati i analizirati integralno.

#### 4. ZAKLJUČAK

Upravljanje prihodima u restoranima ima za cilj da maksimizira pokazatelje performansi upravljanja prihodima, ali uz uvažavanje percepcije pravičnosti. Menadžeri restorana suočavaju se sa stalnim izazovima u domenu maksimizacije stope popunjenosti kapaciteta, jer ukoliko mesta ostanu nepopunjena, ona za restoran predstavljaju čist gubitak [1]. Jedan od problema restorana potiče i od strukture potrošača, kao i dužine trajanja njihovog boravka u restoranu. Restoranski menadžeri se dakle suočavaju i sa izazovom postavljanja strategija upravljanja prihoda na način da se ne omogućava samo maksimizacija njihovih prihoda, već i stvaranje vrednosti za kupce.

Upravljanje prihodima u restoraterskoj industriji može se definisati kao prodaja pravog mesta, pravom kupcu po pravoj ceni, koji će konzumirati uslugu u optimalnom vremenu trajanja. Definisane termina „pravog“ podrazumeva postizanje istovremeno najviše koristi za restoran, ali i kreiranje najveće vrednosti za potrošača. Bez te ravnoteže, praksa upravljanja prihodima dugoročno bi dovela do otuđivanja nezadovoljnih kupca [5].

Merenje uspešnosti definisanih strategija upravljanja prihodima postiže se kroz specifične pokazatelje operativnih restoraterskih performansi. U naučnoj literaturi analizirani su pokazatelji prosečnog računa, prihoda po raspoloživom mestu po satu, profita po raspoloživom mestu po satu i profita po raspoloživom kvadratnom metru. U ovom radu, na osnovu hipotetičkog primera, izvršena je analiza uspešnosti rada restorana na osnovu dva pokazatelja performansi upravljanja prihodima – prihoda po raspoloživom mestu po satu i profita po raspoloživom mestu po satu. Prihod po raspoloživom mestu po satu vrlo uspešno pokazuje koliko će stepen iskorišćenosti raspoloživih kapaciteta doprineti stvaranju prihoda restorana, i takođe, doprinosi analizi efikasnosti i efektivnosti rada restorana. Međutim, kako je za vlasnike, investitore i druge stejkholdere ostvareni profit važnije merilo performansi, u upotrebi je i drugi pokazatelj, profit po raspoloživom mestu po satu. U radu je izvršena komparativna analiza ova dva pokazatelja, kako bi se utvrdilo koliko dobro ukazuju na uspešnost rada restorana. Rezultati sprovedene analize su pokazali da, iako generiše identične prihode po različitim danima u nedelji, restoran nužno ne mora biti jednako profitabilan u tim istim vremenskim intervalima upravo zbog strukture troškova menija. Stoga, kako bi se dobila što preciznija slika u domenu uspešnosti rada restorana na bazi pokazatelja performansi upravljanja prihodima, poželjno je istovremeno posmatrati oba navedena pokazatelja. Makimizacija profita restorana stoga zahteva od menadžera da sprovedu detaljnu analizu svog menija, kako bi znali koliko je svaka pojedinačna stavka profitabilna. U tom smislu, upravljanje menijem bi bilo značajno poboljšano. Takođe, izuzetno je važno da se period služenja obroka skрати, kako bi se povećao obrt (broj usluženih) klijenata po stolu, a samim tim bi se povećali i prihodi restorana.

Sveobuhvatna analiza operativne uspešnosti rada restorana može se postići ukoliko se svi pokazatelji performansi upravljanja prihodima posmatraju integralno. Kako svi oni imaju određene nedostatke, očekuje se da će se u budućnosti naučna literatura detaljnije baviti ovim pitanjem.

#### LITERATURA

- [1] Heo, C.Y. (2016). Exploring group-buying platforms for restaurant revenue management. *International Journal of Hospitality Management*, 52, 154–159. <http://dx.doi.org/10.1016/j.ijhm.2015.07.009>
- [2] Heo, C. Y. (2017). New performance indicators for restaurant revenue management: ProPASH and ProPASM. *International Journal of Hospitality Management*, 61, 1-3. <https://doi.org/10.1016/j.ijhm.2016.10.005>
- [3] Hristov Stančić, B. (2021). *Primena koncepta upravljanja prihodima u analizi poslovanja hotela sa stanovišta operativnih performansi*, Doktorska disertacija, Ekonomski fakultet, Univerzitet u Beogradu.
- [4] Ivanov, S. (2014). *Hotel revenue management: From theory to practice*. Zangador.
- [5] Kimes, S. E. (1999). Implementing restaurant revenue management: A five-step approach. *Cornell Hotel and Restaurant Administration Quarterly*, 40(3), 16-21. <https://doi.org/10.1177/001088049904000315>
- [6] Kimes, S. E., & Chase, R. B. (1998). The strategic levers of yield management. *Journal of service research*, 1(2), 156-166. <https://doi.org/10.1177/109467059800100205>

- [7] Kimes, S. E., Chase, R. B., Choi, S., Lee, P. Y., & Ngonzi, E. N. (1998). Restaurant revenue management: Applying yield management to the restaurant industry. *Cornell Hotel and Restaurant Administration Quarterly*, 39(3), 32-39. <https://doi.org/10.1177/001088049803900308>
- [8] Kimes, S. E., & Thompson, G. M. (2004). Restaurant revenue management at Chevys: determining the best table mix. *Decision Sciences*, 35(3), 371-392. <https://doi.org/10.1111/j.0011-7315.2004.02531.x>
- [9] Kimes, S. E., & Wirtz, J. (2003). Perceived fairness of revenue management in the US golf industry. *Journal of Revenue and Pricing Management*, 1(4), 332-344. <https://doi.org/10.1057/palgrave.rpm.5170037>
- [10] Westering, J. V. (1994). Yield management: the case for food and beverage operations. *Progress in tourism, recreation and hospitality management*, 6, 139-147.

## ODREĐIVANJE NAJVAŽNIJIH KRITERIJUMA PRILIKOM SELEKCIJE KANDIDATA U IT INDUSTRIJI PRIMENOM MACBETH METODE

### DETERMINATION OF THE MOST IMPORTANT CRITERIA CANDIDATES SELECTION IN THE IT INDUSTRY USING THE MACBETH METHOD

ANĐELA MRDAK<sup>1</sup>, TIJANA NANUŠEVSKI<sup>1</sup>, BISERA ANDRIĆ GUŠAVAC<sup>1</sup>, MILENA POPOVIĆ<sup>1</sup>

<sup>1</sup>Univerzitet u Beogradu – Fakultet organizacionih nauka, Beograd, am20180088@student.fon.bg.ac.rs; nanusevski.tijana@gmail.com; bisera@fon.bg.ac.rs; milena.popovic@fon.bg.ac.rs;

**Rezime:** Cilj ovog rada je da oceni važnosti kriterijuma za selekciju kandidata u IT idustriji. U tu svrhu korišćena je multiatributivna metoda za određivanje vrednosti težinskih koeficijenata kriterijuma MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique). U radu je sprovedeno istraživanje među IT regruterima koji su zaposleni u HR sektoru i učestvuju u selekciji kadrova. Ispitane su značajnosti pet grupa kriterijuma specifičnih za IT industriju i osnovne komunikacione i organizacione veštine kandidata, a za svaki kriterijum definisani su i podkriterijumi. Rezultatima istraživanja određene su preferencije HR sektora u IT industriji koje ukazuju i potencijalnim kandidatima koje veštine i sposobnosti bi trebalo da poseduju ili usavršavaju kako bi prošli selekciju i zaposlili se u IT industriji.

**Ključne reči:** MACBETH metoda, težinski koeficijenti, linearno programiranje, selekcija kandidata.

**Abstract:** The aim of this paper is to determine the importance of criteria for the selection of candidates in the IT industry. For this purpose, a multi-attribute method MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique) was used to determine the values of the weighting coefficients of selected criteria. The research was conducted among IT experts who are employed in the HR sector and participate in the selection of candidates, professors and students of the Faculty of Organizational Sciences. The significance of five groups of criteria specific to the IT industry and the basic communication and organizational skills of the candidates were examined, and sub-criteria were given for each criterion. The results of the research determined the preferences of the HR sector in the IT industry, which indicate to potential candidates what skills and abilities they should possess or improve in order to pass the selection and get a job in the IT industry.

**Keywords:** Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH), weights, linear programming, selection of candidates.

## 1. UVOD

Proces regrutacije i selekcije kandidata za zapošljavanje u IT kompaniji obuhvata niz stručnih postupaka koji se preduzimaju sa ciljem proveravanja znanja, radnih sposobnosti, veština i iskustava kandidata. Na osnovu toga šta se od njega traži i koje su to veštine i sposobnosti koje bi bilo poželjno da poseduje, svaki kandidat će uložiti trud i napor kako bi zadovoljio i svoje i potrebe poslodavca. S druge strane cilj IT regrutera je da precizno definiše radno mesto, ključne zahteve radnog mesta i način na koji se mogu proveriti ovi zahtevi.

U teorijskim naučnim istraživanjima, za identifikaciju ključnih kompetencija kandidata korišćene su brojne metode multiatributivnog odlučivanja. U ovom radu, predložena je metoda merenja atraktivnosti kroz kategorije zasnovana na evaluativnim tehnikama - MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique). U MACBETH metodi kreira se stablo odlučivanja kriterijuma pri čemu se određuje važnost svakog kriterijuma i njihovo rangiranje na osnovu mišljenja zainteresovanih strana.

Na osnovu sprovedenog empirijskog istraživanja određene su preferencije IT regrutera prema osobinama i veštinama kandidata koji se prijavljuju na konkurs za posao u IT kompanijama. Izdvojeno je pet grupa kriterijuma specifičnih za IT industriju i osnovne komunikacione i organizacione veštine koje bi kandidat koji konkuriše na IT poziciju trebalo da poseduje. Za svaki kriterijum dati su i podriterijumi, čije su važnosti i rang dobijeni. Na ovaj način, u procesu selekcije kandidata za zapošljavanje mogu se potencijalno smanjiti troškovi selekcije kandidata, unaprediti proces razmatranja hiljade biografija i sprovođenja na stotine intervjuja, a time će se i IT regruteri fokusirati na manji broj obećavajućih kandidata.

Rad je organizovan na sledeći način: u drugom poglavlju dat je pregled dosadašnje primene MACBETH metode u oblasti selekcije kandidata za zapošljavanje, u trećem poglavlju prikazani su osnovni koncepti i metodologija izvođenja metode. U četvrtom poglavlju prikazana je studija slučaja u okviru koje su definisani kriterijumi, objašnjen postupak istraživanja i dat prikaz rezultata. U poslednjem poglavlju izložena su zaključna razmatranja.

## 2. MACBETH U OBLASTI SELEKCIJE KANDIDATA

MACBETH metoda je korišćena kao jedan od pristupa za upravljanje ljudskim resursima, odnosno kao instrument kojim bi mogao da se proceni učinak državnih službenika [1]. Snaga multatributivnih metoda koje su primenjene u pomenutom radu ogleda se u složenim situacijama, gde postoji kombinacija različitih aktera, svaki sa svojim sopstvenim različitim vrednostima, percepcijama, ciljevima, odnosno, situacija koja uključuje ne samo jedan kriterijum, već i višestruke i konfliktne kriterijume [1].

Interesantna je studija [2] u kojoj se predlaže primena MACBETH metode za procenu studenata koji traže međunarodnu stipendiju za mobilnost, gde su kao kriterijumi korišćeni: opterećenje studenta, prihodi, trenutni nivo studija, iskustvo, itd.

Prema saznanjima autora ovog rada, ne postoji puno radova koji se bave primenom MACBETH metode u oblasti selekcije, tako da je rad u ovoj oblasti i primena MACBETH metode interesantna i za buduća istraživanja.

## 3. MACBETH METODA

Metoda merenja atraktivnosti kroz kategorije zasnovana na evaluativnim tehnikama - MACBETH (*Measuring Attractiveness by a Categorical Based Evaluation Technique*) je subjektivna metoda za definisanje težinskih koeficijenta na osnovu kvalitativnih procena eksperata [3].

### 3.1. Algoritam MACBETH metode

Pretpostavimo da je definisano  $n$  kriterijuma  $C_j, j=1,2,\dots,n$  koji su korišćeni za evaluaciju  $m$  alternativa  $A=\{A_1, A_2, \dots, A_m\}$  [4]. Kriterijumi iz skupa  $C_j=\{C_1, C_2, \dots, C_n\}$  rangiraju se prema njihovoj značajnosti, tako da se na prvoj poziciji nalazi najuticajniji kriterijum, dok se na poslednjoj poziciji nalazi najmanje uticajan kriterijum. Formira se matrica poređenja  $M = [\xi_{ij}]_{n \times n}$  u kojoj se vrši poređenje u parovima kriterijuma. U matrici su kriterijumi poređani prema značajnosti s leva na desno i od vrha prema dnu matrice. U matrici  $M$  poređenje se vrši samo iznad dijagonale matrice, dok se na dijagonali matrice upisuju vrednosti nula. Za poređenje u parovima koristi se semantička skala [0-6] [4]. Svaka semantička kategorija predstavlja nivo atraktivnosti među kriterijumima: (0) indiferentno, (1) veoma slabo, (2) slabo, (3) umereno, (4) jako, (5) veoma jako i (6) ekstremno.

Određivanje vrednosti težinskih koeficijenata kriterijuma se dobija rešavanjem MACBETH linearnih modela. Za svakog eksperta (ukupno broj eksperata je  $e$ ) koji učestvuje u istraživanju rešava se po jedan model koji se dobije na osnovu matrice poređenja  $M^b = [\xi_{ij}^b]_{n \times n}$ ,  $1 \leq b \leq e$ . Potrebno je odrediti vrednosti težinskih koeficijenata kriterijuma za svakog eksperta  $w_j^b = (w_1^b, w_2^b, \dots, w_n^b)^T$ ,  $1 \leq b \leq e$ .

MACBETH model formira se na osnovu vrednosti koje su dobijene u agregiranim ekspertskim matricama  $M^b = [\xi_{ij}^b]_{n \times n}$  i prikazan je u nastavku:

$$\min \phi(W_1) \quad (1)$$

p.o.

$$\forall W_i, W_j, i, j \in \{1, 2, \dots, n\} : W_i > W_j \Rightarrow \phi(W_i) \geq \phi(W_j) + \xi(i, j) \quad (2)$$

$$\forall W_i, W_j, i, j, k, l \in \{1, 2, \dots, n\} : \phi(W_i) - \phi(W_j) \geq \phi(W_k) - \phi(W_l) + \xi(i, j, k, l) \quad (3)$$

$$\phi(W_n) = 1 \quad (4)$$

Funkcija cilja data je u (1), gde  $W_1$  predstavlja vrednost najuticajnijeg kriterijuma. Prva grupa ograničenja u modelu su ordinalna ograničenja (2), gde  $\xi(i, j)$  predstavlja razliku nivoa preferencija između  $W_i$  i  $W_j$ . Semantička ograničenja (3) su druga grupa ograničenja, gde je vrednost  $\xi(i, j, k, l)$  definisana kao razlika između  $\xi(i, j)$  i  $\xi(k, l)$ . Vrednost  $W_n$  u ograničenju (4) predstavlja vrednost najmanje uticajnog

kriterijuma. Rešavanjem prethodno definisanog linearnog modela (1)-(4) dobijaju se vrednosti težinskih koeficijenata kriterijuma. S obzirom da je potrebno normalizovati dobijene vrednosti težinskih koeficijenata tako da ispunjavaju uslov da je  $\sum_{j=1}^n w_j = 1$ , neophodno je primeniti izraz (5):

$$w_j = \frac{\phi(W_j)}{\sum_{j=1}^n \phi(W_j)} \quad (5)$$

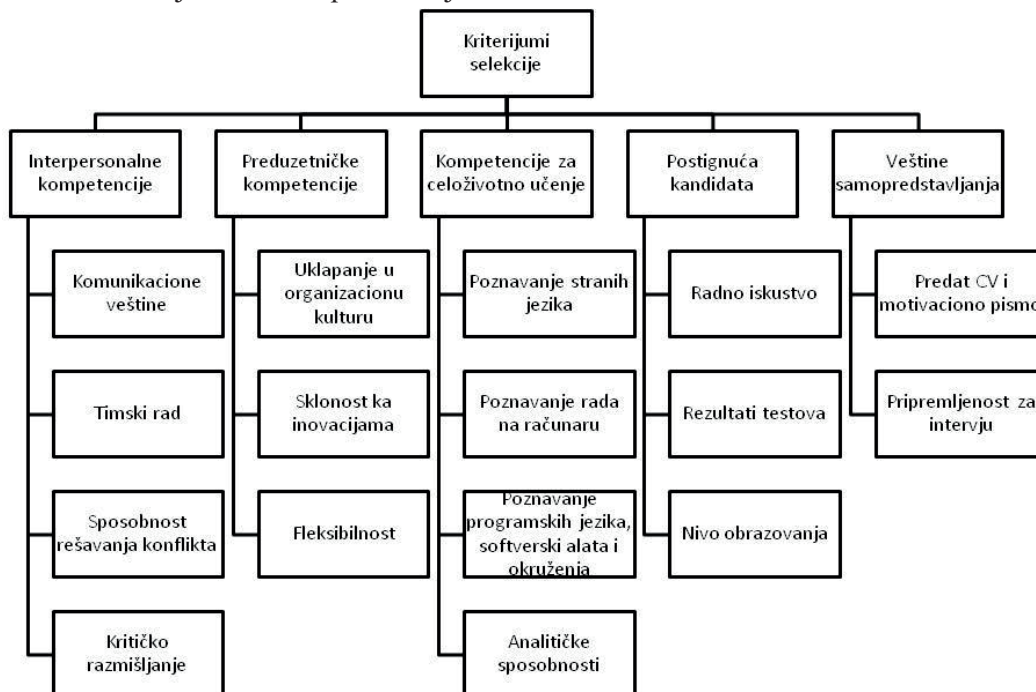
gde  $\phi(W_j)$ ; ( $j=1,2,\dots,n$ ), predstavlja značajnost  $j$ -tog kriterijuma koji je dobijen kao optimalno rešenje MACBETH modela, dok  $w_j$  predstavlja normalizovane vrednosti težinskih koeficijenata.

### 3. PRIMENA MACBETH METODE U PROCESU SELEKCIJE KANDIDATA: STUDIJA SLUČAJA

Selekcija kandidata predstavlja proces tokom kog se, među kandidatima koji su poslali prijave za slobodno radno mesto, vrši izbor kandidata za kog se procenjuje da može na najbolji način da zadovolji zahteve radnog mesta i da ostvari najbolje rezultate [5]. Proces selekcije čini nekoliko faza, a to su [6]: analiza podnetih prijava i pratećih dokumenata i izbor kandidata koji zadovoljavaju uslove, obavljanje preliminarnih intervju sa kandidatima, testiranje kandidata, obavljanje dijagnostičkog intervju sa kandidatima, provera referenci koje su kandidati podneli tokom konkurisanja za posao i davanje ponude kandidatu sa uslovima njegovog zaposlenja. Imajući u vidu da u poslednjih par decenija digitalizacija menja način poslovanja u svakoj oblasti, a samim tim na tržištu rada raste tražnja za stručnjacima u IT oblasti, cilj ovog istraživanja je da se uvide preferencije IT regrutera prema određenim kriterijumima koje kandidat u većoj ili manjoj meri treba da poseduje kako bi odgovorio na zahteve određenog radnog mesta u IT industriji.

#### 3.1 Opis kriterijuma i podkriterijuma

Kao rezultat empirijskih istraživanja izdvojeno je pet grupa kriterijuma specifičnih za IT industriju i osnovne komunikacione i organizacione veštine koje bi kandidat koji konkuriše na IT poziciju trebalo da poseduje (Slika 1). Za svaki kriterijum dati su i podkriterijumi.



Slika 1: Stablo odlučivanja kriterijuma

U okviru *Interpersonalnih kompetencija* razmatrane su komunikacione veštine kandidata, koje mogu biti izuzetno važne u situacijama komunikacije sa zainteresovanim stranama kompanije, potom sposobnost saradnje sa ostalim zaposlenima u kompaniji kako bi poslovi u kompaniji bili obavljeni na što efikasniji način, i rešavanje konflikta nastalih između kolega, što je važno za održavanje prijatne radne atmosfere, a i

za samo poslovanje kompanije. Kritičko razmišljanje se odnosi na sposobnost rešavanja problema, kao i na donošenje odluka, bilo operativnih, taktičkih ili strateških.

Da bi zaposleni obavljao posao na što kvalitetniji način, neophodno je da, pre svega, prihvati i razume organizacionu kulturu same organizacije i samim tim bude motivisan da svojim radom doprinosi poslovanju organizacije. Proaktivno reagovanje zaposlenih i inovacije pružaju mogućnost organizaciji da odgovori na bilo koje izazove okruženja i prihvati promene u bilo kom trenutku poslovanja na tržištu. Fleksibilnost se odnosi na mogućnost da kandidat obavlja poslove na drugim radnim mestima kao i na spremnost kandidata da se adaptira na promene u organizaciji, koje ista uvodi kako bi postigla konkurentsku prednost na tržištu.

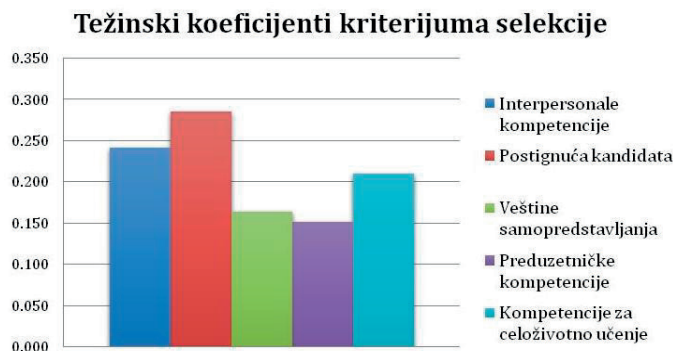
*Kompetencije za celoživotno učenje* tiču se veština koje je kandidat usavršavao u svom obrazovanju i spreman je da ih i dalje usavršava, i time sebe izdvoji u odnosu na druge kandidate i buduće kolege.

Nivo obrazovanja, prethodno radno iskustvo kandidata, mišljenje bivših poslodavaca o kandidatu, kao i postignuti rezultati na testovima inteligencije, znanja i izvršenja su veoma bitni faktori na osnovu kojih se može delimično proceniti način ponašanja kandidata u različitim situacijama na budućem radnom mestu.

U okviru *veština samopredstavljanja i upravljanja utiskom* razmatrana su 2 podkriterijuma. Prvi se odnosi na predat CV, biografiju, motivaciono pismo, kao i na izgled i strukturu *LinkedIn* profila, koji se sve češće analizira u procesu selekcije, a sa druge strane ova društvena mreža postaje medijum komunikacije pomoću koga poslodavci mogu pronaći potencijalne kandidate za posao, a isto tako i kandidati mogu putem ove društvene mreže pronaći aktuelne oglase za posao u različitim oblastima. Drugi podkriterijum je pripremljenost za intervju koji se odnosi na ponašanje kandidata tokom samog intervjua, njegove manire, odevnu kombinaciju, kao i ocenu odgovora na situaciona i bihevioralna pitanja regrutera.

### 3.1 Rezultati MACBETH analize

Kako bi se odredili težinski koeficijenti kriterijuma i, na taj način, zaključili i koje su to kompetencije koje je potrebno da kandidat poseduje da bi bio izabran za određenu poziciju u IT industriji, primenjena je MACBETH metoda. Anketirano je 20 IT regrutera koji su u svojim firmama učestvovali u selekciji kandidata za neku od IT pozicija. Svako od ispitanike je na osnovu svojih ličnih preferencija rangirao kriterijume i podkriterijume, nakon čega je na osnovu definisanih prioriteta, poredio međusobno kriterijume i podkriterijume po parovima, davajući im ocene iz semantičke skale [4]. Formirano je 120 matrica poređenja  $M^b = [\epsilon_{ij}^b]_{n \times n}$ ,  $1 \leq b \leq 20$ , na osnovu kojih je formulisano 120 linearnih modela. Rešavanjem modela dobijeni su težinski koeficijenti kriterijuma i podkriterijuma. Vrednosti težinskih koeficijenata se normalizuju primenom (5). Konačne vrednosti težinskih koeficijenata svih pet kriterijuma dobijeni su izračunavanjem aritmetičke sredine težinskih koeficijenata dobijenih od svakog regrutera (Slika 2).



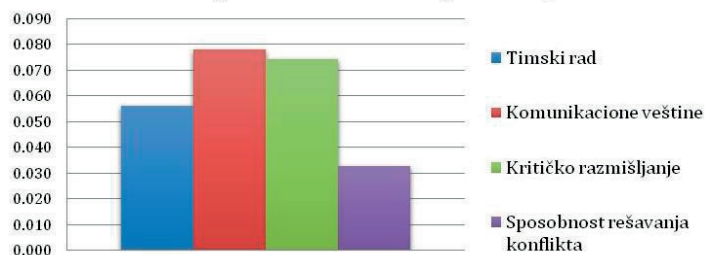
**Slika 2:** Težinski koeficijenti kriterijuma

Regruteri su procenili da su *Prethodna postignuća kandidata* najznačajni kriterijum prilikom selekcije kandidata za određenu poziciju, dok sa druge strane najmanje značajan kriterijum predstavljaju *Preduzetničke kompetencije*. Uočava se da je u uzorku od 20 regrutera, za čak 58.82% dobijeno da *Veštine samopredstavljanja kandidata* nisu značajne prilikom njihove selekcije, tačnije da je težina ovog kriterijuma 0. Sa druge strane, nijedan od regrutera nije na taj način ocenio *Interpersonalne kompetencije i prethodna postignuća kandidata*. Najveći raspon vrednosti ima kriterijum *Veštine samopredstavljanja*, s obzirom da je minimalna vrednost 0, a maksimalna čak 0,98, dok su se regruteri više složili u proceni *Preduzetničkih kompetencija i prethodnih postignuća kandidata*, čiji raspon vrednosti iznosi samo 0,289.

Kada je reč o *Interpersonalnim kompetencijama* uočava se da su *Komunikacione veštine kandidata* najvažnije za regrutere, a odmah nakon njih *Kritičko razmišljanje* (Slika 3). Kao najmanje važan kriterijum ocenjena je *Sposobnost rešavanj konflikata*. Upravo oko ovog kriterijuma, regruteri su bili najslošniji, tako da se vrednosti težine ovog kriterijuma kod svih regrutera kreću od 0,004 do 0,115.



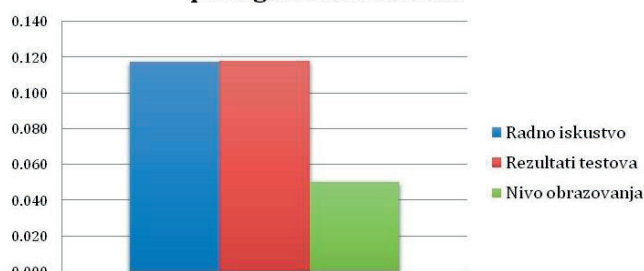
### Težinski koeficijenti interpersonalnih kompetencija



Slika 3: Težinski koeficijenti interpersonalnih kompetencija

Sa veoma malom razlikom, regruteri su respektivno ocenili *Radno iskustvo*, a potom i *Rezultate testova* kandidata tokom selekcije kao najvažnije kriterijume, dok *Nivo obrazovanja* ne smatraju indikatorom koji ukazuje koliko će dobro zaposleni obavljati svoj posao na svom radnom mestu. Ovaj kriterijum je u 70,59% slučajeva bio poslednji prilikom rangiranja prioriteta (Slika 4).

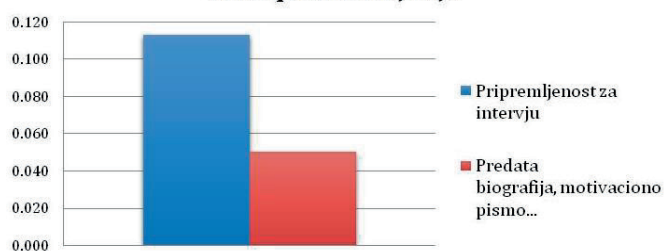
### Težinski koeficijenti prethodnih postignuća kandidata



Slika 4: Težinski koeficijenti prethodnih postignuća kandidata

Čak 82,35% regrutera smatra da ukoliko je kandidat došao pripremljen na intervju i na istom ostavio dobar utisak, njegove reference, poput *CV-a i motivacionog pisma*, mogu se zanemariti. Najčešća ocena koju su regruteri davali u korist *Pripremljenosti za intervju* je 3, što po semantičkoj skali znači da je *Pripremljenost za intervju* umereno atraktivna u odnosu na *CV-a i motivaciono pismo* (Slika 5).

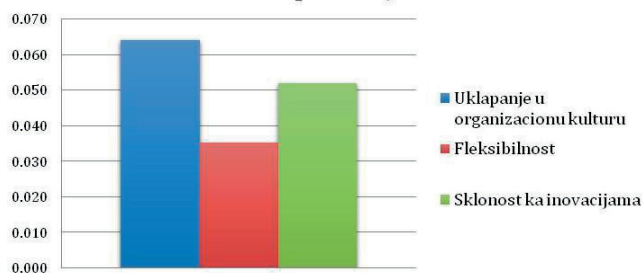
### Težinski koeficijenti veština samopredstavljanja



Slika 5: Težinski koeficijenti veština samopredstavljanja

Među preduzetničkim preferencijama, regruteri preferiraju kriterijum *Uklapanje u organizacionu kulturu*, pa potom da iskazuje *Sklonost ka inovacijama* i kao najmanje važan faktor selekcije uzimaju u obzir *Fleksibilnost* kandidata (Slika 6). *Pripremljenost za intervju* ima veliki raspon vrednosti, najmanja vrednost težinskog koeficijenta ovog kriterijuma za neke od regrutera iznosi 0, a najveća 0,784.

### Težinski koeficijenti preduzetničkih kompetencija



Slika 6: Težinski koeficijenti preduzetničkih kompetencija

Kriterijum *Poznavanje programskih jezika, softverskih alata i okruženja* regruteri smatraju najvažnijim, obzirom da je to i samo sredstvo rada u ovoj industriji, a od njih ne razdvajaju *Osnovna poznavanja rada na računaru* i vrlo često ih rangiraju podjednako. Sa druge strane, ubedljivo najmanje važan faktor prilikom odabira kandidata za posao je *Poznavanje stranih jezika* (Slika 7).



Slika 7: Težinski koeficijenti kompetencija za celoživotno učenje

## 5. ZAKLJUČAK

Jaka konkurencija na tržištu rada IT industrije uslovljava slobodan i brz protok ljudskih resursa i time daje vitalnost lokalnoj ekonomiji i omogućava joj da se efikasno prilagođava svetskom tržištu. Porast značaja IT industrije je posledica toga što kompanije sve više pokušavaju da integrišu IT rešenja u svoje poslovne procese. Predviđen je porast zanimanja u ovoj oblasti do čak 11% do 2029. godine [7], ili preciznije rečeno 531.000 novih poslova, što je znatno više u odnosu na ostale oblasti. Samim tim tržište rada sve više teži da se prilagodi novim načinima poslovanja, te tako sve više mladih ljudi širom sveta teži obrazovanju u ovoj oblasti. Ovo otežava i sam posao IT regruterima, koji su u kompanijama zaduženi da za određenu poziciju u pravom trenutku nađu pravog kandidata, a s obzirom da je konkurencija sve veća, neophodno je da se kandidat po određenim kriterijumima izdvaja od drugih, kako bi upravo on bio zaposlen na željenoj poziciji. Imajući u vidu da je selekcija kandidata delom subjektivni proces i da često mladim ljudima pri završetku studija nije najjasnije šta se od njih očekuje, u ovom radu je urađeno istraživanje sa ciljem da se primenom MACBETH metode odrede preferencije IT regrutera. S druge strane na ovaj način će se i kandidatima koji apliciraju za posao ukazati na veštine i sposobnosti koje bi trebalo da poseduju i usavršavaju.

U sprovedenom istraživanju, posmatrano je pet kriterijuma, gde se kao najvažniji izdvojio *Prethodna postignuća kandidata*, ali ne i kao apsolutno dominantan, jer ga blisko prati kriterijum *Interpersonalne kompetencije*. Najmanje važan je kriterijum *Preduzetničke kompetencije*. Razlog za to može biti taj da se smisao za organizaciju, vođenje projekata i uvođenje nekih novih proizvoda ili tehnologija vezuje za iskusnije zaposlene, a ne za kandidate koji traže posao. Najvažniji podkriterijumi na nivou celog uzorka su *Rezultati testova* i *Radno iskustvo*. IT kompanije pri zapošljavanju često traže kandidate koji poseduju radno iskustvo zbog niza prednosti kojih ovaj podkriterijum ima. Jedan je svakako i skraćivanje vremena samog procesa obuke. Regruteri su dodali da se na osnovu ova dva objektivna podkriterijuma najčešće formira prva preliminarna lista kandidata koji potom ulaze u uži krug selekcije. Interesantno je da je *Nivo obrazovanja* najmanje važan podkriterijum regruterima za IT poziciju, što se slaže sa onim što susrećemo u praksi.

## LITERATURA

- [1] Ensslin, L., Dutra, A., & Ensslin, R. (2000). MCDA: a constructivist approach to the management of human resources at a governmental agency. *International transactions in operational Research*, 7(1), 79-100.
- [2] La Torre, D. (2017). Preface: Multiple criteria optimization and goal programming in science, engineering, and social sciences. *Annals of Operations Research*, 251(1), 1-5.
- [3] E Costa, C. A. B., & Vansnick, J. C. (1994). MACBETH—An interactive path towards the construction of cardinal value functions. *International transactions in operational Research*, 1(4), 489-500.
- [4] E Costa, C. A. B., Vieira, A. C., Nóbrega, M., Quintino, A., Oliveira, M. D., & e Costa, J. B. (2019). Collaborative Value Modelling in corporate contexts with MACBETH. *Procedia Computer Science*, 162, 786-794.
- [5] Ristić Ž: Menadžment ljudskih resursa. Beograd – Etnostil, 2012.
- [6] Todosić, A. (2021). Regrutovanje i selekcija kandidata, analiza praksi srpskih kompanija i iskustva kandidata. *Železnice*, 2021(6), 22-32.
- [7] <https://www.statista.com/statistics/507389/united-states-it-market-share-breakdown/> (pristupano 2.6.2022.)

## PRODUCTIVITY OF LEADING R&D ENTERPRISES IN THE WORLD-LEADING ECONOMIES: A META-FRONTIER APPROACH

ALEKSANDER ARISTOVNIK<sup>1</sup>, GUOLIANG YANG<sup>2</sup>, YAOYAO SONG<sup>3</sup>, DEJAN RAVŠELJ<sup>4</sup>

<sup>1</sup> Faculty of Public Administration, University of Ljubljana, Slovenia, [aleksander.aristovnik@fu.uni-lj.si](mailto:aleksander.aristovnik@fu.uni-lj.si)

<sup>2</sup> Institutes of Science and Development, Chinese Academy of Sciences, China, [glyang@casipm.ac.cn](mailto:glyang@casipm.ac.cn)

<sup>3</sup> School of Economics, Capital University of Economics and Business, China, [songyaoyao@cueb.edu.cn](mailto:songyaoyao@cueb.edu.cn)

<sup>4</sup> Faculty of Public Administration, University of Ljubljana, Slovenia, [dejan.ravselj@fu.uni-lj.si](mailto:dejan.ravselj@fu.uni-lj.si)

---

**Abstract:** *The R&D investment is considered one of the important driving forces of economic development in all countries. Therefore, the paper aims to examine and compare the productivity of leading R&D enterprises between world-leading economies and six different industries during the period 2016-2020. The empirical analysis is performed in the framework of the DEA method, employing the meta-frontier approach and the Malmquist productivity index, utilized on a sample of 1155 companies from the European Union, the United States, Japan, and China operating in 6 different industries. The results reveal that the R&D productivity of all industrial enterprises rose slightly during our examined period, with China exhibiting the worst productivity levels due to an obvious technology gap between China and other regions. Moreover, the Consumer Goods & Services and ICT Goods & Services achieved good R&D performance among the six industries but have not consistently outperformed the average during the investigated period. The findings of the paper can guide improvements in R&D efficiency at both the economy and industry levels.*

**Keywords:** *R&D productivity, DEA, Meta-frontier, Malmquist productivity index, Comparative analysis*





# EKOLOŠKI MENADŽMENT I UPRAVLJANJE PRIRODNIM RESURSIMA

---

## ENVIRONMENTAL MANAGEMENT AND NATURAL RESOURCES MANAGEMENT





## GOING GREEN IN DIGITAL BANKING: A CASE STUDY

JOVAN TRTICA<sup>1</sup>, JELENA ANDREJA RADAKOVIĆ<sup>2</sup>

<sup>1</sup> University of Belgrade – Faculty of Mechanical Engineering jtrtica92@gmail.com

<sup>2</sup> University of Belgrade – Faculty of Organizational Sciences jelenaandreja.radakovic@fon.bg.ac.rs

**Abstract:** *Climate-friendly banking is a critical component of a long-term civilization. A large majority of digital banking users are entrepreneurs, with many of them opening their first bank account. As a result, digital banks play an important role in ensuring that businesses employ a long-term banking partner from the start. At the same time, the corporation wants to meet its climate goals in a consistent and effective manner. To achieve this, several digital banks have chosen to follow the ESG (Environmental, Social, and Governance) model to assure environmental protection, consideration of people and relationships, and corporate governance requirements.*

**Keywords:** *ESG, Digital Banking, Sustainable Banking, Sustainable Development*

### 1. INTRODUCTION

Banks of all sizes are beginning to incorporate Environmental, Social, and Corporate Governance (ESG) into their business strategy as sustainability has emerged as a business-critical endeavor in banking and financial services [4]. The term "environmental, social, and corporate governance" refers to a method of assessing how far a company goes to advance social objectives beyond its basic obligation to make profits for its owners [3]. Commonly, the social objectives endorsed by an ESG perspective involve striving towards a certain set of environmental objectives, a set of objectives involving supporting specific social movements, and a third set of objectives involving determining whether the corporation is run in a manner that is consistent with the objectives of the diversity, equity, and inclusion movement.



Figure 1. ESG model

Sustainable banking has many advantages that are obvious. Banks and other financial institutions can gain a variety of advantages by effectively leveraging sustainability, including cost savings, improved operational efficiency, increased environmental friendliness, and the capacity to both retain and attract new consumers [1].

The process of globalization and domestic and international trade both rely on financial institutions. As intermediate organizations, banks move funds from savers to borrowers to encourage investments and company expansion [2].

The Principles for Responsible Investment (PRI) were developed by the United Nations (UN) in 2005 and they focused on ESG issues [9]. The 17 Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development were further declared by the UN in 2015 [10]. The UN's 2005 PRI initiative demonstrated how ESG factors affect the success of investment portfolios. The appraisal of the companies was then gradually guided by principles for corporate governance, social responsibility, and environmental stewardship.

## 2. ESG CRITERIA

„The three elements of ESG investing are environmental, social, and corporate governance, and each one comprises a range of elements that may be considered by both socially aware investors and companies wishing to adopt a more ESG-friendly approach to operations“ [4]. There are several initiatives underway to produce more accurate, unbiased assessments of a company's ESG policies and practices.



**Figure 2.** Visual representation of three ESG criteria

### 2.1. ESG – Environmental

The usage of renewable energy sources, a company's waste management plan, how it handles any air or water pollution issues brought on by its activities, deforestation issues (if applicable), and its views and actions on climate change issues are all environmental considerations. The source of raw materials (does the business employ fair trade suppliers and organic products, for example?) and other environmental issues are other potential issues. and if a business upholds biodiversity standards on land it owns or controls.

### 2.2. ESG – Social

Social factors encompass a wide variety of possible issues. Even though ESG has many different social aspects, they are all primarily focused on interpersonal relationships. Many socially conscious investors view a company's connection with its employees as one of its most important ties. The following is a list of the possible considerations when evaluating how a company maintains its social relationships [4]:

- When compared to salaries for corresponding jobs or positions throughout the sector, is employee remuneration reasonable or even generous? What retirement plans are available to employees? Does the business make contributions to the retirement plans of its employees?
- What additional perks or bonuses are provided to employees above their base compensation or salary? If a firm provides advantages that aren't common at all businesses, like an on-site fitness center, or does things like providing a free buffet lunch to every employee every Friday, it may have a big impact on how ESG-conscious investors see the company.
- Workplace policies that support diversity, inclusivity, and the prevention of sexual harassment are frequently taken into consideration.
- What opportunities are there for employees to learn new job skills at the organization that would qualify them for positions with greater salaries? For instance, does the company provide financial aid for employees who want to continue their studies or flexible work schedules?
- What percentage of the employees participates in management? How much influence do employees have over decisions that are made in their specific departments?
- The rate of employee turnover.
- What is the mission statement of the business? Is it societally beneficial and relevant?
- How skillfully are consumer relations handled? Does the business interact with its clients on social media? How effective and quick-witted is the personnel providing customer service? Does the business have a troublesome track record with concerns about consumer protection, such as product recalls?



- Does the business publicly or politically support causes related to human rights? Does it make donations to nonprofit organizations?

### 2.3. ESG – Governance

Simply put, governance in the context of ESG refers to how a firm is led by those in positions of authority. How well do the board of directors and top management represent the diverse interests of the company's consumers, shareholders, and suppliers? Does the business support the area where it is based? Successful corporate governance is usually considered as requiring openness in finances and accounting, as well as thorough and truthful financial reporting. The board members' real fiduciary obligations to investors and their careful avoidance of any conflicts of interest are equally important. Are the management and board members of the organization a friendly and diverse group? Successful corporate governance is usually considered as requiring openness in finances and accounting, as well as thorough and truthful financial reporting. The board members' real fiduciary obligations to investors and their careful avoidance of any conflicts of interest are equally important. Are the management and board members of the organization a friendly and diverse group? The issue of CEO remuneration is a major worry for many ESG investors, and they, for example, often do not support executive bonuses of millions of dollars when the business maintains a salary freeze for all other employees. Is the increase in CEO compensation correctly correlated with the enhancement of the business's long-term value, viability, and profitability?

### 3. CASE STUDY OF GREEN BANKING: PENTA

Penta is an IT company that helps business— with its digital features, innovative tools and support. With Penta, users manage their financial activities from a single banking platform, including basic payments, accounting tasks, cost management, and credit management [7]. In order to empower creators and businesses for a sustainable and prosperous future, Penta's founding principles included rethinking business banking and making a constructive contribution to climate change. As part of its own sustainability initiatives, the business collaborates with Planetly to learn more about its carbon emissions and to choose the best course of action for long-term reduction [8].

Climate-friendly banking is a critical component of a long-term civilization. Startups make up the majority of Penta's clientele, and many of them utilize it to open their initial bank accounts. Therefore, Penta may be quite useful in assisting companies in making an immediate decision on a sustainable banking provider. The company wants to achieve its climate targets in a consistent and efficient manner at the same time. In order to accomplish this, Penta has decided to work with Planetly in order to understand, reduce, and offset its CO<sub>2</sub> emissions.

It won't just be a combination of traditional and internet banking that defines the bank of the future. A sustainable, environmentally sensitive digital lifestyle will be supported by and enabled by the bank of the future. Banks will be able to develop new revenue streams as more clients decide to bank with institutions that place a high priority on sustainability.

#### 3.1. Penta's carbon footprint

Planetly closely examined all pertinent actions throughout Penta's business in order to determine the carbon footprint. This includes emissions that are created by acquired goods and services, other emissions that take place outside the organization, as well as direct emissions and emissions connected to energy. To make them addable and comparative, the various types of emissions are estimated as CO<sub>2</sub> equivalents, or CO<sub>2</sub>e. In 2020, Penta had an overall carbon footprint of 862.97 t CO<sub>2</sub>e [8].

The largest sources of emissions for Penta are the company's external service providers, such as the advertising, consulting, and marketing services they hire, followed by employee and building emissions. Customers that use online applications generate relatively few emissions. The results of Planetly's analysis of Penta carbon emissions can be found in Table 1.

**Table 1:** Analysis Results of Penta

Total carbon emissions	863t
Procurement emissions	680.88t CO <sub>2</sub> e
Employee emissions	109.97t CO <sub>2</sub> e
Employee emissions	68.84t CO <sub>2</sub> e

Source [8]

Penta's initial objective was to reach CO<sub>2</sub> neutrality no later than 2022. First they wanted to know more about the composition of their CO<sub>2</sub> profile – so they calculated their CO<sub>2</sub> footprint. The next stage was to cut CO<sub>2</sub> emissions as much as they could. To offset their remaining CO<sub>2</sub> emissions, they additionally supported initiatives for climate protection that have Verified Carbon Standard or are CDM Gold Standard certified. As part of their commitment to their intended activities, their Chief Product Officer has also joined the Leaders for Climate Action (LFCA) group. They emitted 862.97 tons of greenhouse gases in 2020, which is nearly equal to the average annual emissions of 176 persons [5].

Planetly divided Penta's CO<sub>2</sub>e emissions into three scopes in order to be able to determine the reasons of these emissions shown in Table 2.

**Table 2:** Scopes of Penta's CO<sub>2</sub>e emissions

Scope	Causes	Share of footprint in 2020
1	Fugitive emissions	0.1%
2	Electricity	6.9%
3	Purchased products & services, commuting, capital goods, travel	93.0%

Source [6]

They also came up with further reduction strategies by segmenting their CO<sub>2</sub>e footprint into four emission fields [6]:

- Procurement & Product: 78.9%
- Employees: 12.6%
- Building & Office: 8.0%
- Customers: 0.5%

The Penta Green Team, which explores reduction initiatives inside the organization and is overseen by the Penta climate officer, was formed to continue working on this issue and consists of eight individuals from various departments. Every aspect of the company, from office administrative services to advertising, merchandise, and technology, has been improved by the Green Team in order to get Penta closer to its net-zero goal [6]:

- Recycled Penta Cards with packaging composed of recyclable grass paper and made entirely of reusable PET-G plastic;
- A cloud and web hosting company that is climate-neutral;
- The Berlin office uses green energy;
- A combined work arrangement to reduce commute;
- Quarterly report to track their CO<sub>2</sub>e reductions from working from home;
- Use of public transportation is free;

Since internal improvements take time to implement, they also chose a unique portfolio of climate protection initiatives that allowed them to retrospectively offset our 2020 emissions. All 17 UN Sustainable Development Goals (SDGs) are positively impacted by the high-impact portfolio they selected. This means that their offsets support social and economic development objectives in addition to having a positive influence on ecological projects. If the emission reductions attained by an offset project would not have happened without funding from the sale of offset credits, then the project meets the additionality requirement [6].

By doing all of this and more, Penta has been functioning as a carbon-neutral business since 2020 and has done so retroactively thanks to extensive compensation measures [5]. Banks that are sustainable exhibit resource-saving practices and make financial decisions that don't harm people or the environment. As a finance company that is rapidly expanding, Penta's duty is to provide their clients with future goods and services that will have a positive influence on their ability to reduce emissions. Their customer deposits are invested in accordance with the strictest sustainability guidelines, and no funds are provided to organizations that harm the environment.

#### 4. CONCLUSION

The bank of the future won't just be one that combines the experiences of traditional and online banking. The bank of the future will support a sustainable, eco-conscious digital lifestyle and be an enabler of it. As more customers choose to bank with companies that prioritize sustainability, this will allow banks to create new

revenue sources. Companies that embrace a sustainable business model receive a competitive edge in addition to helping society and the environment. Naturally, this is assuming that the subject is taken seriously and that their efforts are openly reported on and shared.

## REFERENCES

- [1] Chang, H.-Y., Liang, L.-W., & Liu, Y.-L. (2021). Using Environmental, Social, Governance (ESG) and Financial Indicators to Measure Bank Cost Efficiency in Asia. *Sustainability*, 13(20), 11139. MDPI AG. <http://dx.doi.org/10.3390/su132011139>
- [2] Ghosh, A. (2018). What drives banking industry competition in developing countries? *Journal of Economic Development*, 43(4), 1-20.
- [3] Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance*, 66, 101889.
- [4] Corporate Finance Institute. (2015-2022). *ESG (Environmental, Social and Governance)*. Retrieved on 20<sup>th</sup> of June 2022 from: <https://corporatefinanceinstitute.com/resources/knowledge/other/esg-environmental-social-governance/>
- [5] Get Penta. (2021a). *Penta Is Climate-Neutral: Another Step Towards Sustainable Banking*. Retrieved on 18th of June 2022 from: <https://getpenta.com/en/blog/penta-is-carbon-neutral/>
- [6] Get Penta. (2021b). *Sustainable banking: Penta's journey to becoming a carbon-neutral company*. Retrieved on 18th of June 2022 from: <https://getpenta.com/en/blog/sustainable-banking/>
- [7] Get Penta. (2022). *Revolutionising digital banking is what we do*. Retrieved on 18th of June from: <https://getpenta.com/en/company/why-penta/>
- [8] Planetly. (2021). *Penta Revolutionises the Digital Banking World*. Retrieved on 18th of June from: <https://www.planetly.com/case-studies/penta?fbclid=IwAR3YmeIf9JxDW4vPUeO2nqA941zE65F1izjyX69vI9RdUuJsn-BVCZ7PJRk>
- [9] United Nations - UN. (2020). *Principles for Responsible Investment*. 2020. Retrieved on 21st of June from: <https://www.unpri.org/sustainability-issues>
- [10] United Nations – UN. (2022). *Do you know all 17 SDGs?*. Retrieved on 21st of June from: <https://sdgs.un.org/goals>



## PRISTUPI, TEHNIKE I METODE OPERACIONIH ISTRAŽIVANJA U SMANJIVANJU RIZIKA OD KATASTROFA

### APPROACHES, TECHNIQUES AND METHODS OF OPERATIONAL RESEARCH IN DISASTER RISK REDUCTION

JELENA ANDREJA RADAKOVIĆ<sup>1</sup>, DRAGANA MAKAJIĆ-NIKOLIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, jelenaandreja.radakovic@fon.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, dragana.makajic-nikolic@fon.bg.ac.rs

**Rezime:** U 21. veku, uprkos svim naprecima u tehnologiji, neminovna je činjenica da se broj katastrofa povećava. Bile to prirodne, antropogene ili hibridne katastrofe, njihova frekvencija je učestalija. Katastrofe svake godine odnesu veliki broj ljudskih života, devastiraju životnu sredinu i prave značajne ekonomske gubitke. Uloga operacionih istraživanja u proučavanju i minimiziranju rizika od katastrofa je krucijalna i sve bitnija. Smanjenje rizika od katastrofa, udruženo sa naporima zasnovanim na odgovoru i oporavku, čini osnovu savremenog upravljanja rizikom od katastrofa. Iz ovih razloga primenjuju se različiti pristupi tehnike i metode operacionih istraživanja. U ovom radu su posebno prikazani Analiza stabla neispravnosti i Analiza stabla događaja.

**Ključne reči:** Operaciona istraživanja, Katastrofe, Rizik od katastrofa, Smanjivanje rizika od katastrofa

**Abstract:** In the 21<sup>st</sup> century, despite all the advances in technology, the fact that the number of catastrophes is increasing is inevitable. Be it natural, anthropogenic or hybrid disasters, their frequency is raising. Every year, catastrophes take a large number of human lives, devastate the environment and cause significant economic losses. The role of operational research in studying and minimizing disaster risk is crucial and increasingly important. Disaster risk reduction, coupled with response and recovery-based efforts, forms the basis of modern disaster risk management. For these reasons, different approaches to techniques and methods of operational research are applied. In this paper, Fault Tree Analysis and Event Tree Analysis are presented separately.

**Keywords:** Operational Research, Disasters, Disaster Risk, Disaster Risk Reduction

## 1. UVOD

Katastrofe su veliki nerešivi problemi koji ispituju sposobnost zajednica i nacija da efikasno zaštite stanovništva i infrastrukture, kako bi se smanjio i gubitak ljudi i imovine, a sistem brzo oporavio. Prividna slučajnost uticaja i problema i jedinstvenost incidenata zahtevaju dinamična, efikasna i ekonomična rešenja u realnom vremenu, što katastrofe čini veoma „pogodnim” za primenu operacionih istraživanja, tako da je sve više istaknuta potreba za proučavanjem i primenom različitih pristupa, tehnika i metoda operacionih istraživanja u upravljanju katastrofama.

Većina istraživanja u upravljanju katastrofama odnosi se na društvene nauke. Ova vrsta istraživanja usredsređena je na rezultate katastrofa, sociološke uticaje na zajednice, psihološke efekte na preživeli i spasilačke timove, kao i na organizacioni dizajn i probleme u komunikaciji.

S druge strane, definicija operacionih istraživanja nije unificirana. Churchman i ostali [4] definišu operaciona istraživanja kao primenu naučnih metoda, tehnika i alata na probleme koji uključuju rad sistema, kako bi onima koji upravljaju operacijama pružili optimalna rešenja problema. Winston [21] ih definiše kao naučni pristup donošenju odluka, koji nastoji da odredi kako najbolje dizajnirati i upravljati sistemom, obično pod uslovima koji zahtevaju alokaciju oskudnih resursa. Asocijacija evropskih društava za operaciona istraživanja (*Association of European Operational Research Societies – EURO*) definiše operaciona istraživanja kao naučni pristup rešavanju problema u upravljanju složenim sistemima u promenljivom okruženju što po samoj svojoj definiciji implicira primenljivost operacionih istraživanja u smanjivanju rizika od katastrofa i upravljanju katastrofama.

## 2. HAZARDI I KATASTROFE

Hazardni su preduslov svakog rizičnog događaja i predstavljaju stvarno ili potencijalno stanje koje može da izazove eventualne povrede, bolesti ili smrti, štetu po životnu sredinu, štetu ili potpuni gubitak sistema, opreme ili imovine [5, 9]. Samim tim, hazard je i preduslov pojavljivanja hazardnog (neželjenog) događaja. Ovakvi događaji možda se nikada ne bi ni desili kada ljudi ne bi bili izloženi hazardima ili kada bi sistemске barijere efikasno sprečavale hazardne događaje [14]. Hazardni događaj prvi je u nizu događaja koji ukoliko nisu kontrolisani vode do neželjenih posledica [13]. Ovakav događaj može predstavljati momenat nakon koga je kontrola hazarda izgubljena i nakon čega je bilo kakav uticaj na njega moguć samo u vidu sanacije njegovih rezultujućih negativnih posledica. U literaturi hazardni događaji zovu se takođe i: nesreće, nesrećni događaji, neželjeni događaji, devijacije, neispravnosti, početni događaji, itd.

Prema Kancelariji ujedinjenih nacija za smanjivanje rizika od katastrofa (*United Nations Office for Disaster Risk Reduction – UNDRR*) [20] hazardni su potencijalno štetni fizički događaji, fenomeni i/ili ljudske aktivnosti koji mogu da prouzrokuju gubitak ljudskih života ili povrede, oštećenje imovine, društvena i ekonomska ometanja ili ekološku degradaciju. Oni mogu biti pojedinačni, sekvencijalni ili kombinovani u odnosu na svoje poreklo i efekte.

Katastrofe su prouzrokovane hazardima. Hazarde treba razlikovati od događaja i rizika. U literaturi koja obrađuje ovu tematiku, neki autori ograničavaju hazarde na prirodno izazvane događaje, dok drugi uključuju i događaje izazvane ljudskim aktivnostima.

Iako nijedna definicija katastrofa nije generalno prihvaćena i njihova formulacija zavisi od oblasti istraživanja [17], definicija UNDRR-a je jedna od najopsežnijih. Prema UNDRR [18], katastrofa je ozbiljan poremećaj u funkcionisanju zajednice ili društva koji uključuje kako ljudske, tako i materijalne, ekonomske ili gubitke životne sredine, kao i negativne uticaje, prevazilazeći sposobnost pogođene zajednice ili društva da se nosi sa posledicama sopstvenim resursima. Pored toga, Centar za istraživanje epidemiologije katastrofa (*Center for Research on the Epidemiology of Disasters – CRED*), koji održava Bazu podataka o hitnim događajima (*Emergency Events Database – EM-DAT*), definiše katastrofe kao događaje gde je prijavljeno deset ili više poginulih ljudi; 100 ljudi je prijavljeno kao pogođeno; upućen je poziv za međunarodnu pomoć; proglašeno je vanredno stanje [10].

Prema autoru Fritz [6], katastrofa je događaj koji je vremenski i prostorno ograničen, u kome nastaju fizičke štete i oštećenja nekih suštinskih funkcija društva. Fizička oštećenja i socijalni poremećaji (sada se češće nazivaju fizičkim i socijalnim uticajima) nastaju zato što događaj prevazilazi uobičajenu zaštitu [7]. Zahtev da se događaj koncentriše na vreme i prostor od suštinske je važnosti za razlikovanje smrtnih slučajeva u zemljotresima koji bi mogli da broje i 50 smrtnih slučajeva u roku od nekoliko minuta od smrtnih slučajeva sudara automobila koji u Sjedinjenim Američkim Državama, na primer, broje približno 40.000 godišnje.

Prema Ujedinjenim nacijama pod pojmom katastrofe podrazumevaju se rezultati hazarda uz nedovoljnu mogućnost izbegavanja ili smanjenja nastalih negativnih posledica [12], dok prema Zakonu o vanrednim situacijama Republike Srbije katastrofa je elementarna nepogoda ili druga „nesreća i događaj koji veličinom, intenzitetom i neočekivanošću ugrožava zdravlje i živote većeg broja ljudi, materijalna dobra i životnu sredinu, a čiji nastanak nije moguće sprečiti ili otkloniti redovnim delovanjem subjekata odbrane” [23].

## 3. METODE I TEHNIKE ZA PROCENU I SMANJIVANJE RIZIKA OD KATASTROFA

Smanjivanje rizika od katastrofa (*Disaster Risk Reduction - DRR*) ima za cilj prevenciju nastanka novih rizika od katastrofa kao i smanjenje već nastalih, čime se jača društveno-ekonomska otpornost na katastrofe, smanjuju negativni posledični uticaji na životnu sredinu, te tako ostvaruju i ciljevi održivog razvoja [19]. Tako, se već u Jokahama strategiji i planu za akciju za bezbedniji svet (*Yokohama Strategy and Plan of Action for a Safer World*) [22] kao prvom međunarodnom okviru za DRR, prepoznaje veza između održivog razvoja i DRR-a, da bi se nastavilo sa jačanjem ove povezanosti kako u Milenijumskim ciljevima razvoja (*Millennium Development Goals - MDGs*), tako i u Planu implementacije iz Johanesburga (*Johannesburg Plan of Implementation*), „Hjogo okviru za akciju (2005-2015)” (*Hyogo Framework for Action (2005-2015)*”), dokumentu „Budućnosti koju želimo” (*Future We Want*”), Sendai okviru za DRR (*Sendai Framework for DRR*) i 2030 Agendi za održivi razvoj (*2030 Agenda for Sustainable Development*).

DRR zahteva pristup usmeren na ljude i više sektora, izgradnju otpornosti na višestruke, kaskadne i interaktivne hazarde i stvaranje kulture prevencije i otpornosti. Stoga uključuje strategije dizajnirane da:

- izbegavaju stvaranja novih rizika,
- se bave postojećim rizicima,

- bi sprečile da gubici od katastrofa budu apsorbovani u drugim razvojnim ishodima uz stvaranje dodatnog siromaštva.

Smanjenje rizika od katastrofa predstavlja osnovu savremenog upravljanja rizikom od katastrofa (*Disaster Risk Management – DRM*), a zasniva se na hitnim odgovorima i oporavcima od katastrofalnih događaja [9].

Iako DRM uključuje aktivnosti pripremljenosti i reagovanja na katastrofe, predstavlja mnogo više od upravljanja katastrofama [18].

Kada se govori o upravljanju pouzdanošću, rizikom i sigurnošću sistema, treba naglasiti da se u ovoj oblasti koristi preko 700 metoda i tehnika koje mogu da budu specijalizovane ili opšte [8]. Razvoju ovako velikog broja metoda i tehnika je doprineo razvoj računara, posebno kada je u pitanju prikupljanje i obrada podataka.

Faza procene i analize rizika se smatra najkompleksnijom i zahtevnom fazom u upravljanju pouzdanošću i rizikom. Neke od najpoznatijih metoda u fazi procene i analizi rizika su [5]:

- Analiza stabla događaja,
- Analiza stabla neispravnosti,
- Analiza načina i efekata otkaza (*Failure Mode and Effect Analysis – FMEA*),
- Binarni dijagrami odlučivanja (*Binary Decision Diagram – BDD*),
- Modeli Markova (*Markov Analysis*),
- Petrijeve mreže (*Petri Nets – PN*),
- Preliminarna analiza hazarda (*Preliminary Hazard Analysis – PHA*),
- ...

#### 4. PRIMENE METODA ZA SMANJIVANJE RIZIKA OD KATASTROFA: ANALIZA STABLA NEISPRAVNOSTI I ANALIZA STABLA DOGAĐAJA

Donošenje odluka u svrhu smanjivanja rizika od katastrofa uključuje više zainteresovanih strana - stejkholdera, regiona i sektora, kao i više ciljeva koji se odnose na upotrebu resursa i koristi. U slučaju prilagođavanja ovim ekstremnim događajima, treba takođe sprovesti i modeliranje puteva udara i posledica od katastrofa [15]. Konsultuju se stručnjaci koji predstavljaju potencijalno pogođene infrastrukturne usluge o upotrebljivosti metoda kao što su Analiza stabla neispravnosti (*Fault Tree Analysis – FTA*) i Analiza stabla događaja (*Event Tree Analysis – ETA*) za pružanje strukturiranih informacija o scenarijima katastrofa, uticajima i posledicama sistema, rizicima i protivmerama. Glavni korisnici rezultata analize su vlasnici imovine i lokalni donosioci odluka čiji su zajednički naporu obično potrebni za finansiranje i određivanje prioriteta takvih mera prilagođavanja.

FTA prema autorima Badida i ostali [2] je važna tehnika procene rizika koja verovatnoće otkaza komponentata tretira kao tačne vrednosti za procenu verovatnoće pojave vršnog događaja. Zbog nedostatka istorijskih podataka za izračunavanje stope otkaza cevovoda usled prirodnih hazarda, njihova studija imala je za cilj da analizira verovatnoću kvara cevovoda korišćenjem Fazi analize stabla neispravnosti (*Fuzzy Fault Tree Analysis – FFTA*) uz stručno navođenje. Za rangiranje primarnih događaja korišćena je *Fussel-Vesely* mera značajnosti. Predloženi okvir FFTA korišćen je za analizu pojave najvažnijih događaja čak i u odsustvu istorijskih podataka o verovatnoći.

U još jednom zanimljivom primeru primene FTA na katastrofe, autori Baum i ostali [3] u svom radu procenjuju verovatnoće izbijanja nuklearnog rata. Ova verovatnoća je glavni faktor u mnogim važnim političkim pitanjima, ali je privukla malo pažnje naučnika. Ovaj rad predstavlja model za izračunavanje ukupne verovatnoće nuklearnog rata. Model se zasniva na 14 međusobno povezanih scenarija kako može izbiti nuklearni rat, pokrivajući možda čitav niz scenarija nuklearnog rata. Scenariji se razlikuju u zavisnosti od faktora, uključujući da li država namerava da izvrši prvi udarni napad, da li nuklearnom napadu prethodi konvencionalni rat ili neratna kriza, da li je eskalacija namerna ili nenamerna, prisustvo lažnih alarma različitih vrsta, kao i prisustvo neratnih nuklearnih detonacija poput nuklearnog terorizma. Kao prvi korak ka kvantifikovanju verovatnoće nuklearnog rata pomoću modela, rad takođe uključuje i niz istorijskih incidenata koji bi mogli da pređu u nuklearni rat. Uključeno je 60 istorijskih incidenata, što ga čini možda trenutno najvećim dostupnim skupom podataka u ovoj oblasti.

U radu autora Santos i ostali [16], strategije upravljanja rizikom od nastajanja suša uzimaju u obzir i direktne i indirektne posledice poremećaja u snabdevanju vodom. Autori predlažu metodologiju za procenu strategija upravljanja sušom kombinovanjem ekonomskog input-output modeliranja sa ETA. Metodologiju primenjuju na simulirani scenario suše koji utiče na region glavnog grada Sjedinjenih Američkih Država – Vašington. Rezultati su pokazali da iako upravljanje potrošnjom vode donosi najmanji kumulativni ekonomski gubitak u regionu, smanjenje početnog nivoa poremećaja vodosnabdevanja i prioritet

zavisnostima korišćenja vode, rezultiraju manjom neoperabilnošću kritičnih sektora. Ova otkrića pružaju uvid donosiocima odluka u identifikovanje kritičnih sektora i formulisanje pravovremenih strategija intervencija koje umanjuju ukupne efekte suša na ekonomske sisteme.

Autori Neri i ostali [11] u svom radu istražuju primenljivost ETA na erupciji vulkana Vezuv. Autori ističu da da bi se postigla koherentnost podataka, postoji mnogo različitih vrsta dokaza objedinjenih unutar formalizovane strukture povezane stručnim znanjem. U tu svrhu, stablo događaja koje su kreirali ima zadatak da rezimira u numeričko-grafičkom obliku, na različitim nivoima relativne verovatnoće koje se odnose na genezu i stil erupcije vulkana, razvoj i prirodu vulkanskih hazarda i verovatnoće pojave različitih vulkanskih rizika. Formulacija stabla događaja pruža logičan put koji povezuje generičku verovatnoću procene hazarda sa kvantitativnom procenom rizika. Za ovaj sveobuhvatni pristup, bili su im potrebni iscrpni modeli rizika, kvantifikovani raspodelom nesigurnosti za sve uključene faktore. Prikazana je struktura stabla događaja erupcije Vezuva, a stablo događaja se proteže od hazarda od erupcije vulkana sve do ljudskog uticaja i posledica za infrastrukturu, formirajući kvantitativni sinoptički okvir za sveobuhvatnu procenu hazarda i mapiranje uticaja rizika. Organizacija stabla događaja na ovaj način omogućava autorima lako ažuriranje, kada i ako nove informacije postaju dostupne.

U radu autora Rosqvist i ostali [15] se ekspertima i državnim službenicima odgovornim za uzimanje u obzir mogućih efekata klimatskih promena u njihovim sektorima pokazuje korisnost metode ETA za procenu uticaja i podršku odlukama javnog sektora za upravljanje rizikom od poplava. Direktne posledice širokog spektra uticaja se kategorišu prema ekonomskim, humanitarnim i ekološkim posledicama. U skladu sa osnovnom metodom ETA, autori su razvili model stabla događaja, nakon čega je on predstavljen ekspertima za poplave i menadžerima sektora u oblastima zgrada/nekretnina, vodoprivrede, električne i telekomunikacione mreže. Uloga eksperata bila je da daju stručne povratne informacije o pristupu ETA i pruže informacije za specifikaciju modela, posebno o funkcijama barijera i povezanim verovatnoćama uspeha/neuspeha. Ovaj rad demonstrira pristup u odnosu na tekuće inicijative za zaštitu od poplava grada Pori u Finskoj. Na osnovu povratnih informacija eksperata, za najperspektivnije područje primene pronađeno je upravljanje na nivou sektora lokalnih elektroenergetskih radova, upravljanje stambenim područjima, upravljanje oborinskim vodama itd.

Iako su se stručnjaci složili da je ETA korisna za donošenje odluka o zaštiti od poplava, primetili su da postoje neki izazovi i ograničenja u primeni pristupa: Opšti zaključci su bili da se ranjivost sistema i sektora neprestano menja zbog autonomne adaptacije. Uvedena generička struktura stabla događaja smatrana je adekvatnom za demonstraciju i diskusiju o ETA pristupu. Međutim, neki eksperti su predvideli situacije gde se redosled naglašavanja barijera u nekim slučajevima može razlikovati. Presude o verovatnoći grananja zahtevaju operativno iskustvo, dok generisanje alternativa za zaštitu od poplava i procena prihvatljivog rezidualnog rizika (ranjivosti) zahteva upravljačko iskustvo. Uslovljavanje verovatnoće grananja na trajanje poplave eksperti su ocenili zanimljivim, ali i donekle teškim za razumevanje zbog naprednijeg računa verovatnoće. Eksperti su lako razumeli rezultate u smislu godišnjeg rezidualnog rizika i troškova i koristi od ulaganja u novu kontrolu poplava (gde je korist bila smanjenje očekivanih direktnih troškova zbog štete). Numeričke analize osetljivosti smatrale su se neophodnim da bi se proverila robusnost rezultata. Uprkos gorenavedenim izazovima, ETA metoda je ocenjena kao korisna za strukturiranje problema procene uticaja i sistematizaciju procesa procene rizika u upravljanju rizikom od poplave, čak i ako bi numerički rezultati bili vrlo sirovi [15].

## 5. ZAKLJUČAK

Sa pojavom sve većih šteta i sve većeg broja žrtava od katastrofa pojačanog intenziteta i učestalosti, te sve manjeg uticaja lokalnih zajednica koje nemaju kapacitete da same odgovore na posledice katastrofa, javilo se interesovanje za procene rizika od katastrofa u naučnoj zajednici.

Kod primene operacionih istraživanja u proceni rizika od katastrofa i upravljanju katastrofama treba istaći da su autori Altay i Green III [1] uradili opsežni pregled literature objavljene u ovoj oblasti. Naučne radove su kategorisali na osnovu njihovog doprinosa znanju iz oblasti upravljanja katastrofama i na osnovu razvoja modela, razvoja teorije i razvoja primene (alata). Radovi koji razvijaju analitički model za rešavanje i analizu problema ili za procenu ishoda pripadaju prvoj grupi, istraživanje koje testira hipoteze, istražuje ponašanje sistema ili pruža okvir i unapređuje razumevanje nekih pojava na terenu svrstava se u klasu razvoja teorije, a istraživanje u kojem se proizvodi računarski alat ili razvija prototip spada u grupu za razvoj aplikacija.

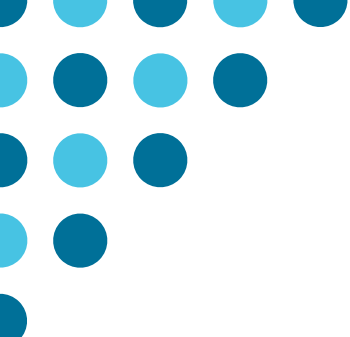
Dalje u oblasti primenljivosti metoda ETA i FTA objavljeni radovi na ovu temu pokazuju da je moguće primeniti ove dve analize na procenu rizika, verovatnoća hazarda i posledica katastrofa. Takođe, pokazuju i da postoji prostor za proširenje istraživanja u ovoj oblasti.



## LITERATURA

- [1] Altay, N., & Green III, W. G. (2006). OR/MS research in disaster operations management. *European journal of operational research*, 175(1), 475-493.
- [2] Badida, P., Balasubramaniam, Y., & Jayaprakash, J. (2019). Risk evaluation of oil and natural gas pipelines due to natural hazards using fuzzy fault tree analysis. *Journal of Natural Gas Science and Engineering*, 66, 284-292. <https://doi.org/10.1016/j.jngse.2019.04.010>
- [3] Baum, S., de Neufville, R., & Barrett, A. (2018). A model for the probability of nuclear war. *Global Catastrophic Risk Institute Working Paper*, 18-1.
- [4] Churchman, C. W., Ackoff, R. L., & Arnoff, E. L. (1957). *Introduction to operations research*. John Wiley & Sons.
- [5] Ericson II, C. A. (2005). *Hazard Analysis technique for System Safety*. John Wiley & Sons.
- [6] Fritz, C. E. (1961). Disaster. In R. K. Merton & R. A. Nisbet (Eds.) *Contemporary Social Problems*. Harcourt, Brace and World, 651-694.
- [7] Kreps, G. (1984). Response to social crisis and disaster. *Annual Review of Sociology*, 10, 309-30.
- [8] Makajić-Nikolić, D. (2012). *Novi pristup analizi pouzdanosti sistema primenom inverznih Petrijevih mreža* [Nepublikovana doktorska disertacija]. Univerzitet u Beogradu – Fakultet organizacionih nauka.
- [9] Makajić-Nikolić, D. (2020). Disaster Risk Reduction. In: Leal Filho W., Azeiteiro U., Azul A., Brandli L., Özuyar P., Wall T. (Eds.), *Climate Action. Encyclopedia of the UN Sustainable Development Goals*. Springer, Cham. [https://doi.org/10.1007/978-3-319-71063-1\\_65-1](https://doi.org/10.1007/978-3-319-71063-1_65-1)
- [10] Mohamed Shaluf, I. (2007). Disaster types. *Disaster Prevention and Management: An International Journal*, 16(5), 704-717. <https://doi.org/10.1108/09653560710837019>
- [11] Neri, A., Aspinall, W. P., Cioni, R., Bertagnini, A., Baxter, P. J., Zuccaro, G., Andronico, D., Barsotti, S., Cole, P. D., Esposti Ongaro, T., Hincks, T. K., Macedonio, G., Papale, P., Rosi, M., Santacroce, R., & Woo, G. (2008). Developing an event tree for probabilistic hazard and risk assessment at Vesuvius. *Journal of volcanology and geothermal research*, 178(3), 397-415. <https://doi.org/10.1016/j.jvolgeores.2008.05.014>
- [12] Radaković, J. A., Drašković, B., Makajić-Nikolić, D., & Petrović, N. (2021 Septembar 20-23). Uloga nacionalnih platformi za smanjenje rizika od katastrofa u jačanju otpornosti na poplave i degradaciju životne sredine. U D. Urošević, M. Dražić & Z. Stanimirović (Eds.), *Zbornik radova XLVIII Simpozijuma o operacionim istraživanjima - SYM-OP-IS 2021*, (pp. 47-52), Banja Koviljača. Matematički fakultet Univerziteta u Beogradu: Beograd.
- [13] Rausand, M. (2013). *Risk assessment: theory, methods, and applications* (Vol. 115). John Wiley & Sons.
- [14] Renn, O. (2008). *Risk governance: coping with uncertainty in a complex world*. Earthscan.
- [15] Rosqvist, T., Molarius, R., Virta, H., & Perrels, A. (2013). Event tree analysis for flood protection - An exploratory study in Finland. *Reliability Engineering & System Safety*, 112, 1-7. <https://doi.org/10.1016/j.res.2012.11.013>
- [16] Santos, J. R., Pagsuyoin, S. T., Herrera, L. C., Tan, R. R., & Krista, D. Y. (2014). Analysis of drought risk management strategies using dynamic inoperability input-output modeling and event tree analysis. *Environment Systems and Decisions*, 34(4), 492-506
- [17] Shaluf, I. M., Ahmadun, F. L., & Mat Said, A. (2003). A review of disaster and crisis. *Disaster Prevention and Management: An International Journal*, 12(1), 24-32.
- [18] United Nations Office for Disaster Risk Reduction – UNDRR. (2015). Sendai Framework for Disaster Risk Reduction 2015-2030. Pristupljeno 27. maja 2022 sa [https://www.unisdr.org/files/43291\\_sendaiframeworkfordrren.pdf](https://www.unisdr.org/files/43291_sendaiframeworkfordrren.pdf). Accessed 16 July 2018
- [19] United Nations Office for Disaster Risk Reduction - UNDRR. (2020-2022). Disaster risk reduction. Pristupljeno 26. maja 2022 sa <https://www.undrr.org/terminology/disaster-risk-reduction#:~:text=Disaster%20risk%20reduction%20is%20aimed,the%20achievement%20of%20sustainable%20development>
- [20] United Nations Office for Disaster Risk Reduction - UNISDR (2009). *UNISDR terminology on disaster risk reduction*. United Nations.
- [21] Winston, W. L. (1994). *Operations research-Applications and algorithms*. Wadsworth

- [22] Yokohama Strategy and Plan of Action for a Safer World. (1994 May 23-27). World Conference on Natural Disaster Reduction, Yokohama, Japan. Pristupljeno 27. maja 2022 sa <https://www.eird.org/fulltext/Yokohama-strategy/Yokohama%20Strategy%20and%20Plan%20of%20Action%20for%20a%20Safer%20World.pdf>
- [23] Zakon o vanrednim situacijama Republike Srbije. (2012). „Sl. glasnik RS“, br. 111/2009, 92/2011 i 93/2012. Pristupljeno 21 maja 2022, sa [http://eupravnik.rs/zakoni/zakon\\_o\\_vanrednim\\_situacijama.pdf](http://eupravnik.rs/zakoni/zakon_o_vanrednim_situacijama.pdf)



# EKONOMSKI MODELI I EKONOMETRIJA

## ECONOMIC MODELS AND ECONOMETRICS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
**Faculty of Economics  
and Business**



## BAJESOV PRISTUP U ANALIZI EFEKATA REGULACIONOG FOKUSA NA ODLUČIVANJE POTROŠAČA

### BAYESIAN APPROACH IN ANALYSIS OF THE EFFECTS OF REGULATORY FOCUS ON CONSUMER DECISION-MAKING

LAZAR ČOLIĆ<sup>1</sup>, IVANA PRICA<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, lazar.colic@ekof.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, ivanapopov.prica@gmail.com

**Rezime:** U ovom radu razvijen je model zasnovan na kombinovanju neuromarketinškog instrumenta i Bajesovog metoda u ocenjivanju preferencija potrošača. Primenjen je hijerarhijski Bajesov model izbora uz korišćenje metrika merenih elektroencefalografom, kako bi se u model uključilo trenutno psihološko stanje potrošača. Dobijenim modelom analiziran je uticaj muzike na cenovnu osetljivost tražnje. Rezultati sugerišu da energičnost muzike kroz regulaciono stanje značajno smanjuje cenovnu osetljivost tražnje. Metod razvijen u ovom radu može se, osim muzike, primeniti na različite kontekstualne faktore koji primuju regulaciona stanja potrošača.

**Ključne reči:** Neuromarketing, Bajesova analiza, Regulacioni fokus, Modeliranje izbora

**Abstract:** In this paper, we have developed a model based on neuromarketing and the Bayesian method in assessing consumer preferences. A hierarchical Bayesian choice model and metrics obtained using electroencephalograph were applied in empirical research to model consumer choices and take into account their current psychological state. The resulting model analyses the influence of music on the price sensitivity of demand. The results suggest that the arousing effect of music via the regulatory state significantly reduces the price sensitivity of demand. Except for music, the method developed in this paper could also be applied to other contextual factors that affect the regulatory states of consumers.

**Keywords:** Neuromarketing, Bayesian analysis, Regulatory focus, Choice modelling

#### 1. UVOD

Predmet istraživanja ovog rada je analiza efekata regulacionog fokusa na odlučivanje potrošača prilikom kupovine proizvoda. Regulacioni fokus potrošača je psihološko stanje gde potrošač pridaje veću važnost minimalnim ciljevima, ako je u prevencionom fokusu, odnosno maksimalnim ciljevima, ako je u promociionom fokusu. Promocioni fokus karakteriše pojačana osetljivost potrošača na pozitivne stimuluse, dok prevencioni fokus karakteriše pojačana osetljivost na negativne stimuluse [7], [10], [11]. Muzika primuje regulaciona stanja potrošača u zavisnosti od tempa i energičnosti konkretne numere [4], [5], [13]. Energična muzika izaziva uzbuđenje kod potrošača i dovodi ga u promociiono stanje [6]. U ovom radu je sprovedeno istraživanje kojim se želi objasniti veza između muzike kojoj je potrošač izložen i regulacionog stanja na odlučivanje potrošača. Analiza efekata muzike koja se čuje u pozadini (pozadinska muzika), na psihološko stanje ispitanika biće merena putem elektroencefalografa (EEG). Cena, kao komponenta proizvoda, predstavlja gubitak koji je potrošač spreman da žrtvuje u zamenu za proizvod (ili uslugu). Kako su osobe sa promociionim fokusom manje osetljive na gubitke, sledi da bi potrošači trebalo da manje obraćaju pažnju na cenu (kao negativnu komponentu) kada su izloženi energičnijem muzičkom stimulusu. Promociiono regulaciono stanje karakterišu intenzivnije emocije, odnosno potrošači primovani energičnijom muzikom biće u stanju veće emocionalne pobuđenosti [6]. Iz ovoga sledi prva hipoteza: emocionalna pobuđenost ispitanika izazvana pozadinskom muzikom smanjuje cenovnu osetljivost tražnje ispitanika. Sa druge strane, potrošači usklađuju svoje ponašanje sa regulacionim stanjem samo ako nisu u stanju prevelike involviranosti tokom izvođenja zadatka [1], [18]. Efekat muzike, kao periferni izvor informacija o proizvodu će imati uticaja ako je involviranost potrošača tokom kupovine niska [5]. Druga hipoteza glasi: sa većom involviranošću potrošača prilikom donošenja odluka efekat emocionalne pobuđenosti na cenovnu osetljivost tražnje nestaje.

## 2. METOD

### 2.1. Uzorak

Izvučen je uzorak 32 studenata Ekonomskog fakulteta Univerziteta u Beogradu (u decembru 2021). Uslov za učešće u istraživanju je bio da ispitanik konzumira redovno espresso kafu i da je ženskog pola (kako bi se povećala homogenost uzorka). Prosečna starost ispitanika je bila 20,3 godina, u intervalu od imala 18 do 26.

### 2.2. Procedura istraživanja i eksperimentalni dizajn

Na osnovu informacija dobijenih tokom predistraživanja formirana je matrica atributa i njihovih nivoa (Tabela 1).

**Tabela 1:** Eksperimentalni dizajn sa atributima i njihovim nivoima

Brend	Tip posluživanja*	Zemlja proizvodnje	Cena
Lavazza	Sa slikom	Srbija	105 rsd
Luccari	Bez slike	Italija	125 rsd
Eight o'clock		Brazil	145 rsd
			165 rsd

\*Tip posluživanja sa slikom se odnosi na espresso kafu posluženu sa artistskom slikom napravljenom od pene.

Rezultati istraživanja su mereni metodom ličnog intervjuisanja uz pomoć kompjutera (engl. computer aided personal interviewing, CAPI). D-optimizacija je vršena tako što je korišćen nehijerarhijski model izbora sa parametrima ocenjenim u predistraživanju. Napravljeno je 10 upitnika koji su podeljeni u dve korpe: pet za merenje baznih preferencija i pet za merenje preferencija tokom izloženosti muzičkom stimulusu. Svaki ispitanik je dobijao nasumično dve verzije upitnika za ocenjivanje preferencija, jednu pre i jednu posle odlučanog muzičkog stimulusa. Tokom popunjavanja upitnika meren je neurosignal putem EEG-a. Svi ispitanici su bili podeljeni u dve eksperimentalne grupe, gde je svaka grupa imala različit muzički stimulus. Korišćene su dve instrumentalne verzije numere Wake me up - Avicci, brzog i sporog ritma, čije su validnosti potvrđene u fazi predistraživanja.

### 2.3. Varijable i merni instrumenti

Za potrebe istraživanja korišćen je EEG uređaj marke EMOTIV EPOC+ koji je dokazano pouzdan instrument kad su u pitanju istraživanja tipa kao u ovoj disertaciji [9], [14], [17]. Signal dobijen putem EEG-a je prebačen u frekvencioni domen korišćenjem brzih Furijeovih transformacija (FFT algoritam). Za potrebe ovog istraživanja korišćeni su teta i alfa frekvencioni pojasevi, neophodni za računanje neuroloških metrika za merenje psiholoških stanja, emocionalne valence i emocionalne pobuđenosti. Za formiranje emocionalne valence je korišćen standardan pristup frontalne asimetrije koji se koristi u neuronaučnim radovima, kao i u praktičnim istraživanjima [2], [12], [16]. Frontalna teta asimetrija (u nastavku FA) je računata prema metodu koji se tipično koristi u praktičnim i naučnim istraživanjima [12]. Za merenje emocionalne pobuđenosti korišćen je pristup merenja emocionalnog intenziteta prema kom niska aktivacija alfa talasa u frontalnom korteksu odražava visok nivo emocionalne pobuđenosti [15]. Cenovna osetljivost tražnje je merena preko ocenjene parcijalne korisnosti hijerarhijskog Bajesovog modela izbora na osnovu upitnika koji su ispitanici dobijali pre i nakon izloženosti muzičkom stimulusu.

### 2.4. Specifikacija modela i metoda

Osoba (potrošač)  $i \in I: I$  bira između ponuđenih opcija  $j \in J: J$  u zadatku izbora  $t \in T: T$ . Svaka opcija generiše korisnost u skladu sa linearnom kombinacijom atributa. Parcijalne korisnosti  $\beta$  variraju među pojedincima. Model korisnosti prema kome ispitanici biraju ponuđene opcije je dat izrazom:

$$U_{itj} = [BezIzbora_{tj}; Brend_{tj}; ZemljaPorekla_{tj}; Cena_{tj}; Cena_{tj}\lambda_{tj}] \beta_i \quad (1)$$

Svaki atribut u uglastoj zagradi je matrica sa kolonama koje predstavljaju varijable u modelu. Sve varijable su kodirane kao varijable efekata (engl. *effect coding*), dok je varijabla cena tretirana kao standardizovana metrička varijabla. Struktura varijabli modela je data u sledećoj tabeli:

**Tabela 2:** Varijable korišćene u modelu

Atribut <sup>1</sup>	Varijable u modelu <sup>2</sup>
Bez izbora	Bez izbora
Brend	Eight o'clock
Brend	Luccari
Tip posluživanja	Bez slike
Zemlja proizvodnje	Brazil
Zemlja proizvodnje	Italija
Cena	Metrička varijabla
Cena* $\lambda$	Cena*FA*EP
Cena* $\lambda$	Cena*FA
Cena* $\lambda$	Cena*EP

<sup>1</sup> Sve nemetričke varijable su kodirane kao varijable efekata (engl. *effect coding*), dok je cena standardizovana.

<sup>2</sup> FA – frontalna asimetrija; EP – emocionalna pobudenost.

Verovatnoća izbora je definisana u skladu sa logit modelom koristeći softmax funkciju veze korisnosti i verovatnoće izbora.

$$P(U_{itj} = \max(U_{it})) = \frac{\exp(U_{itj})}{\sum_{j=1}^4 \exp(U_{itj})} \quad (2)$$

Preferencije potrošača su definisane kroz parcijalne korisnosti koje prate apriornu multivarijacionu normalnu raspodelu sa vektorskim parametrom lokacije  $\beta$ . Kovarijaciona matrica je definisana kroz parametrizaciju koja podrazumeva korelacionu matricu  $\Omega$  i vektorske parametre skaliranja  $\tau$ . Ovakva parametrizacija omogućava lakše ocenjivanje hijerarhijskih modela, naročito u slučaju izražene heterogene strukture individualnih preferencija [3].

$$\beta_i \sim N(\beta, \text{diag}(\tau)\Omega\text{diag}(\tau)) \text{ za } i = 1, 2, \dots, I \quad (3)$$

Parametar lokacije  $\beta$  prati apriornu studentovu raspodelu sa tri stepena sloboda. Ovaj izbor je pravdan činjenicom da deblji repovi omogućavaju fleksibilnije ocenjivanje, ali da ipak ne budu previše debeli kako ne ne bi došlo do problema prilikom ocenjivanja (kao što je to slučaj kod problema separacije) [8]. Korelaciona matrica prati LKJ apriornu raspodelu (engl. *Lewandowski-Kurowicka-Joe distribution*) sa parametrom oblika jednakim dva. U slučaju potrebe za neinformativnom raspodelom bira se parametar 1 koji daje uniformnu raspodelu u dozvoljenom prostoru korelacionih matrica. Veće vrednosti favorizuju jediničnu korelacionu matricu. Vektorski parametar skaliranja  $\tau$  prati pozitivno odsečenu studentovu raspodelu po istom principu kao i parametar  $\beta$ .

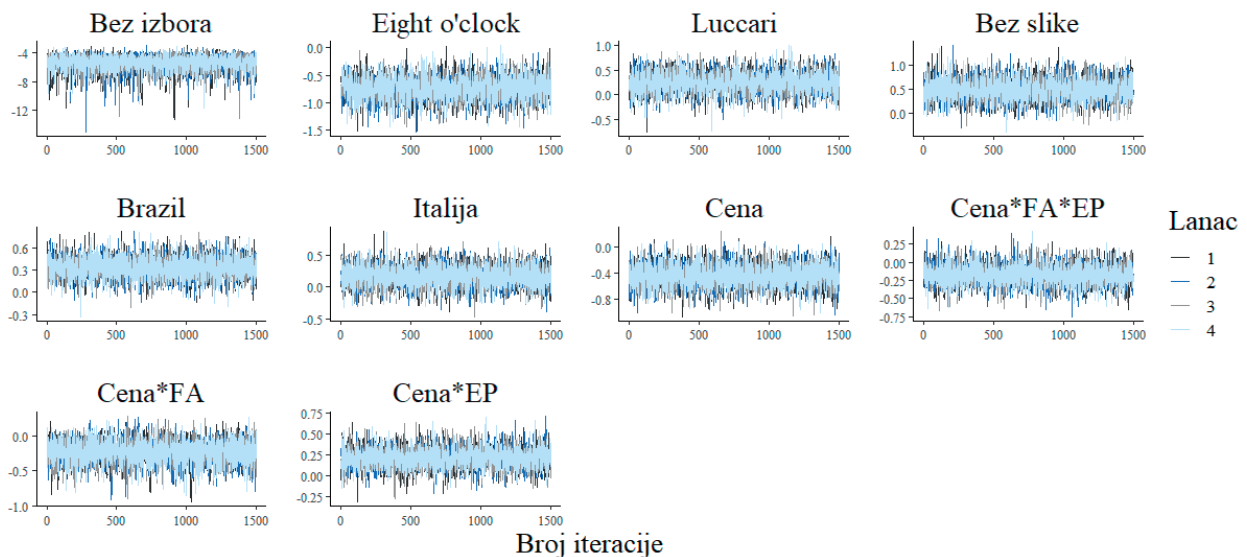
$$\beta \sim T(3, \mu, \text{diag}(1)), \quad \Omega \sim \text{LKJcorr}(2), \quad \tau \sim T^+(3, \mathbf{0}, \text{diag}(1)) \quad (4)$$

Hiperparametarski vektor  $\beta$  ima definisanu lokaciju ocenjenu u fazi predistraživanja:

$$\mu = [-4,06; -0,44; -0,08; -0,64; -0,21; 0,41; -0,58; 0; 0; 0] \quad (5)$$

## 2.5. Provera validnosti modela i testiranje hipoteza

Broj iteracija koji je bio potreban da se stabilizuju dobijene ocene je iznosio 12.000. Polovina iteracija je korišćena za period zagrevanja modela (engl. *warm up period*), dok je ostatak korišćen za ocenjivanje aposteriornih raspodela. Posmatranjem dijagrama traga, osim atributa "Bez izbora", svi hiperparametri parcijalnih korisnosti naliče na oblik gusenice što sugeriše dobro mešanje lanaca. Gelman-Rubinoва statistika je manja od 1,05 za sve parametre (na individualnom nivou) i hiperparametre parcijalnih korisnosti, što sugeriše da se ocene stabilne.



Slika 1: Dijagram traga za beta hiperparametre. Izvor: kalkulacija autora.

## 2.6. Analiza osetljivosti modela

Prilikom analize osetljivosti korišćena su dva alternativna modela. Glavni model (M1) je poređen sa informativnijim (M2) i fleksibilnijim (M3) modelom. M2 koristi restriktivnije apriorne raspodele, dok M3 koristi fleksibilnije apriorne raspodele. Rezultati iz Tabele 3 sugerišu da je glavni model robustan u odnosu na alternativne modele. Svi hiperparametri imaju logičan predznak što sugeriše da postoji logička validnost modela (engl. *face validity*).

Tabela 3: Analiza osetljivosti hiperparametara parcijalnih korisnosti.

Atribut	M1	M2	% M1-M2	M3	% M1-M3
Bez izbora	-5,46	-4,90	10,26%	-7,04	28,94%
Eight o'clock	-0,73	-0,74	1,37%	-0,76	4,11%
Luccari	0,25	0,25	0%	0,25	0%
Bez slike	0,49	0,49	0%	0,55	12,24%
Brazil	0,34	0,34	0%	0,35	2,94%
Italija	0,17	0,17	0%	0,17	0%
Cena	-0,44	-0,45	2,27%	-0,44	0%
Cena*FA*EP	-0,17	-0,17	0%	-0,19	11,76%
Cena*FA	-0,24	-0,24	0%	-0,26	8,33%
Cena*EP	0,21	0,21	0%	0,22	4,76%

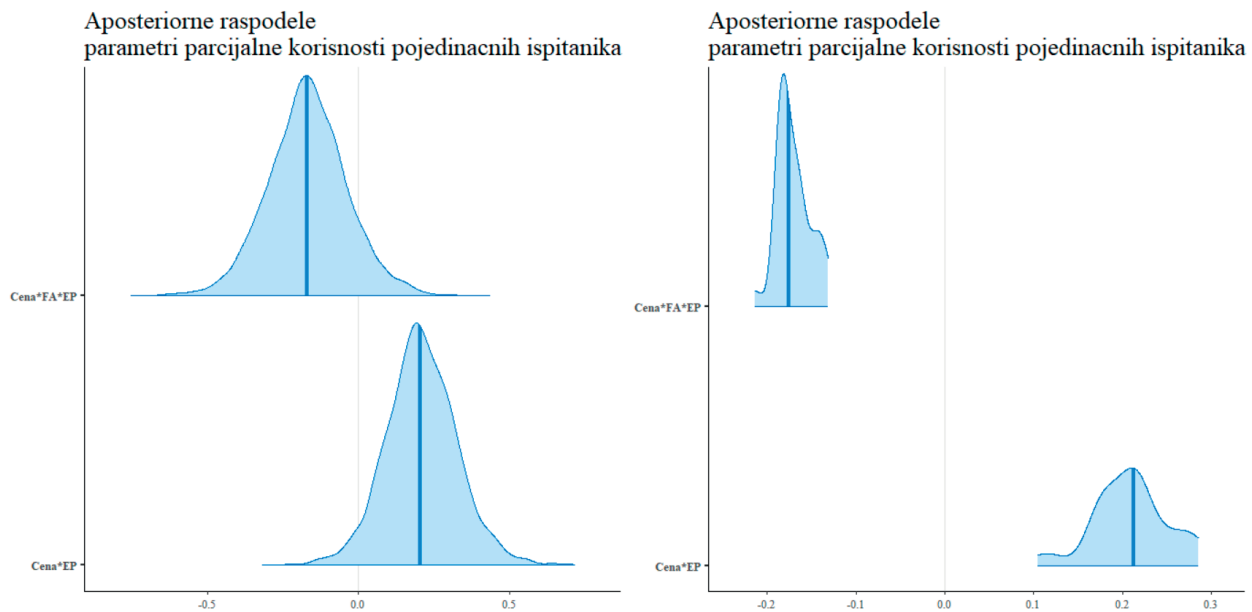
M1 – osnovni model; M2 – informativniji model; M3 – neinformativniji model.

Izvor: kalkulacija autora.

## 3. Rezultati i diskusija

Prva hipoteza pretpostavlja da će cenovna osetljivost biti smanjena prilikom veće emocionalne pobuđenosti. Efekat prve hipoteze će biti anuliran sa većom involviranošću ispitanika prilikom odlučivanja o izboru, što predstavlja drugu hipotezu ovog rada. Prilikom testiranja analizirane su raspodele parametara uz Cena\*EP, odnosno interakcije cene i emocionalne pobuđenosti, kao i parametra uz Cena\*FA\*EP, odnosno interakcije cene i emocionalne pobuđenosti i frontalne asimetrije.





**Slika 2:** Aposteriorne raspodele parcijalnih korisnosti. U prvom redu se nalaze gustine raspodele koeficijenta interakcije cene, frontalne asimetrije i emocionalne pobuđenosti, a u drugom interakcije cene i emocionalne pobuđenosti. Izvor: kalkulacija autora.

Veće vrednosti frontalne asimetrije znače veću involviranost ispitanika prilikom izbora espresso kafa. Pozitivne vrednosti raspodele parametra uz Cena\*EP govore da sa povećanjem emocionalne pobuđenosti smanjuje se cenovna osetljivost tražnje, uz prosečnu involviranost ispitanika (nulta vrednost FA). Ovaj rezultat je izražen i kod raspodele hiperparametra (96%) i kod raspodele parcijalnih korisnosti medijana pojedinačnih ispitanika (100%). Negativna vrednost raspodele parametra uz Cena\*FA\*EP govori da sa povećanjima involviranosti ispitanika (veće FA), efekat emocionalne pobuđenost na cenovnu osetljivost se anulira. Ovaj rezultat je takođe pouzdan, odnosno 91% vrednosti je negativno kod raspodele hiperparametra i 100% ispitanika je imalo negativnu medijanu za ovaj parametar.

#### 4. ZAKLJUČAK

Dobijeni rezultati sugerišu značajnu vezu muzike i regulacionog stanja potrošača, kao i efekte interakcije ovih faktora na cenovnu osetljivost tražnje. Sprovedeno istraživanje pokazuje da energičnost i tempo muzike koja se pušta u prodavnicama predstavlja bitan instrument za donošenje odluka o ceni koju je potrošač spreman da plati, budući da rezultati ukazuju da izloženost energičnijoj muzici smanjuje cenovnu osetljivost tražnje. Bolje razumevanje efekata kontekstualnih faktora, poput muzike, na regulaciono stanje potrošača i, posledično, na njegovo ponašanje je jedan od doprinosa ovog istraživanja. Drugi bitan doprinos odnosi se na oblast ekonometrijskog modeliranja izbora potrošača. Razvijen je metod koji kombinuje neuromarketing pristup merenja psihološkog stanja pojedinca i Bajesov hijerarhijski metod u modeliranju preferencija potrošača. Na ovaj način ekonometrijski model uzima u obzir heterogenu prirodu potrošača na tržištu, kao i uticaj kontekstualnih faktora poput muzike na izbor potrošača. Metod koji smo ovde razvili se može koristiti da se testira uticaj i drugih kontekstualne faktori, poput mirisa, na preferencije potrošača. Takođe je moguće primeniti i druge neuromarketing instrumente, poput uređaja za praćenje pogleda ispitanika, čime bi merenje psihološkog stanja bilo sveobuhvatnije.

#### LITERATURA

- [1] Aaker, J. L., & Lee, A. Y. (2006). Understanding Regulatory Fit. *Journal of Marketing Research*, 43(1), 15–19.
- [2] Allen, J. J. B., Coan, J. A., & Nazarian, M. (2004). Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. *Biological Psychology*, 67(1), 183–218. <https://doi.org/10.1016/j.biopsycho.2004.03.007>
- [3] Betancourt, M. J., & Girolami, M. (2013). Hamiltonian Monte Carlo for Hierarchical Models. *ArXiv:1312.0906 [Stat]*. <http://arxiv.org/abs/1312.0906>

- [4] Caldwell, C., & Hibbert, S. A. (1999). Play That One Again: The Effect of Music Tempo on Consumer Behaviour in a Restaurant. *ACR European Advances, E-04*. <https://www.acrwebsite.org/volumes/11116/volumes/e04/E-04/full>
- [5] Chebat, J.-C., Chebat, C. G., & Vaillant, D. (2001). Environmental background music and in-store selling. *Journal of Business Research, 54*(2), 115–123. [https://doi.org/10.1016/S0148-2963\(99\)00089-2](https://doi.org/10.1016/S0148-2963(99)00089-2)
- [6] Das, G., Roy, R., & Spence, M. T. (2020). The mitigating effect of matching regulatory focus with arousal-inducing stimuli in service failure situations. *Psychology & Marketing, 37*(10), 1420–1432. <https://doi.org/10.1002/mar.21390>
- [7] Forster, J., Higgins, E. T., & Idson, L. C. (1998). *Approach and Avoidance Strength During Goal Attainment: Regulatory Focus and the “Goal Looms Larger” Effect*. 17.
- [8] Ghosh, J., Li, Y., & Mitra, R. (2018). On the use of Cauchy prior distributions for Bayesian logistic regression. *Bayesian Analysis, 13*(2), 359–383.
- [9] Giraldo, S., & Ramirez, R. (2013). *BRAIN-ACTIVITY-DRIVEN REAL-TIME MUSIC EMOTIVE CONTROL*. 6.
- [10] Higgins, E. T. (1998). Promotion and Prevention: Regulatory Focus as A Motivational Principle. In *Advances in Experimental Social Psychology* (Vol. 30, pp. 1–46). Elsevier. [https://doi.org/10.1016/S0065-2601\(08\)60381-0](https://doi.org/10.1016/S0065-2601(08)60381-0)
- [11] Higgins, E. T., Shah, J., & Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology, 72*(3), 515–525. <https://doi.org/10.1037/0022-3514.72.3.515>
- [12] iMotions. (2019). *Electroencephalography: The compete pocket guide*. <https://imotions.com/guides/electroencephalography-eeg/>
- [13] Milliman, R. E. (1982). Using Background Music to Affect the Behavior of Supermarket Shoppers. *Journal of Marketing, 46*(3), 86–91. <https://doi.org/10.2307/1251706>
- [14] Ramele, R., Villar, A. J., & Santos, J. M. (2021). Report: EPOC Emotiv EEG Basics. *TECHNICAL REPORT*, 15.
- [15] Schmidt, L. A., & Trainor, L. J. (2001). Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cognition and Emotion, 15*(4), 487–500. <https://doi.org/10.1080/0269993004200187>
- [16] Smith, E. E., Reznik, S. J., Stewart, J. L., & Allen, J. J. B. (2017). Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. *International Journal of Psychophysiology, 111*, 98–114. <https://doi.org/10.1016/j.ijpsycho.2016.11.005>
- [17] Strmiska, M., & Koudelkova, Z. (2018). Analysis of Performance Metrics Using *Emotiv EPOC +*. *MATEC Web of Conferences, 210*, 04046. <https://doi.org/10.1051/mateconf/201821004046>
- [18] Wang, J., & Lee, A. Y. (2006). The role of regulatory focus in preference construction. *Journal of Marketing Research, 43*(1), 28–38.

## DETERMINANTE STOPE VIŠKA SMRTNOSTI PO ZEMLJAMA U PERIODU PANDEMIJE COVID-19<sup>1</sup>

### DETERMINANTS OF EXCESS MORTALITY RATES DURING THE COVID-19 PANDEMIC ACROSS COUNTRIES

ALEKSANDRA NOJKOVIĆ<sup>1</sup>, MILOJKO ARSIĆ<sup>2</sup>, EMILIJA MAKSIMOVIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, [aleksandra.nojkovic@ekof.bg.ac.rs](mailto:aleksandra.nojkovic@ekof.bg.ac.rs)

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, [milojko.arsic@ekof.bg.ac.rs](mailto:milojko.arsic@ekof.bg.ac.rs)

<sup>3</sup> Univerzitet u Beogradu – Ekonomski fakultet, [emilija.maksimovic@ekof.bg.ac.rs](mailto:emilija.maksimovic@ekof.bg.ac.rs)

**Rezime:** Stopa viška smrtnosti predstavlja jedan od najvažnijih pokazatelja uspešnosti određenog društva u suočavanju sa pandemijom COVID-19. Uspešnost suočavanja sa zdravstvenim posledicama pandemije zavisi od velikog broja faktora, kao što su vrsta epidemioloških mera i efikasnost u njihovoj primeni, kvalitet zdravstvenog sistema i institucija, karakteristike stanovništva i prosečna temperatura. Primenom ekonometrijskog modela panel podataka, na kombinovane mesečne i godišnje podatke za preko 60 zemalja, dobijeno je da je povećanje stope smrtnosti bilo manje u zemljama koje su češće varirale intenzitet mera, kao i u zemljama koje su imale kvalitetniji zdravstveni sistem, veću stopu vakcinacije, kvalitetnije institucije, mlađe stanovništvo i u kojima je stepen kolektivizma manji.

**Ključne reči:** stopa viška smrtnosti, panel model, pandemija COVID-19

**Abstract:** The excess mortality rate during COVID-19 pandemic is one of the most important indicators of the success of a particular society in coping with the virus. The success in curbing health consequences of a pandemic depends on a number of factors, such as the type of epidemiological measures and the effectiveness of their implementation, the quality of the health system and institutions, the characteristics of the population and the average temperature. By applying the econometric panel data model, with combined monthly and annual data for over 60 countries, it was obtained that the increase in mortality rate was smaller in countries that more often varied the intensity of measures, as well as in countries with better health system, higher vaccination rate, better institutions, younger population and in which the degree of collectivism is lower.

**Keywords:** excess mortality rate, panel model, pandemic COVID-19

#### 1. UVOD

Višak smrtnosti predstavlja sintetički pokazatelj uspešnosti određenog društva u ublažavanju negativnih zdravstvenih posledica pandemije. Višak smrtnosti predstavlja obuhvatniji pokazatelj od zvaničnog broja umrlih od COVID-19 koji obuhvata preminule od COVID-19, kao i povećanje broja umrlih od drugih bolesti u periodu pandemije. Podaci o višku smrtnosti su pouzdaniji nego podaci o broju umrlih od COVID-19, jer su greške merenja i podsticaji za manipulaciju podacima u slučaju viška smrtnosti manji.

Višak smrtnosti pokazuje povećanje broja umrlih u određenom periodu (mesec, kvartal, godina) tokom pandemije u odnosu na prosečan broj umrlih u istom vremenskom intervalu pre pandemije. Za međunarodna poređenja adekvatniji pokazatelj je stopa viška smrtnosti koja pokazuje koliko je povećana smrtnost u nekoj zemlji u određenom periodu u odnosu na 100.000 stanovnika. Kumulativna stopa viška smrtnosti u 2020. i 2021. godini na uzorku od 77 zemalja u proseku je iznosila 257, ali ona pokazuje velike varijacije po zemljama i kreće se u intervalu od -50 u slučaju Novog Zelanda do 881 u slučaju Bugarske.

Potencijalne determinante stope viška smrtnosti klasifikovane su u nekoliko grupa: epidemiološka ograničenja, stopa vakcinacije, nivo razvijenosti zemalja, kvalitet zdravstvenog sistema i institucija, karakteristike stanovništva i prosečna temperatura ([5], [6], [7]).

<sup>1</sup> Projekat: *Macroeconomic implications of COVID-19 and effectiveness of policy response in Europe: Empirical evidence and econometric modelling* finansiran od strane Fonda za nauku Republike Srbije.

Epidemiološka ograničenja su merena pomoći dva pokazatelja - indeksa strogosti epidemioloških mera i varijacije epidemioloških mera. Očekuje se da strožije epidemiološke mere smanjuju zaražavanje, a time i smrtnost od COVID-19, ali strožije epidemiološke mere mogu da pogoršaju zdravlje ljudi i smrtnost od drugih bolesti, što može da umanja njihove pozitivne efekte na smrtnost od COVIDa-19 [1]. Varijacija epidemioloških mera u periodu pandemije odražava prilagođavanje mera vlade promeni epidemiološke situacije, pa se stoga očekuje da veće varijacije impliciraju manju smrtnost. Uticaj stope vakcinacije na stopu smrtnosti je nedvosmislen - veća stopa vakcinacije, pod ostalim nepromenjenim okolnostima, utiče na manju stopu viška smrtnosti [2].

Nivo ekonomske razvijenosti zemlje, meren veličinom GDP po stanovniku u dolarima jednake kupovne snage, predstavlja sintetički pokazatelj veličine njenih ekonomskih resursa za suočavanje sa pandemijom. Viši nivo razvijenosti implicira po pravilu razvijeniji zdravstveni sistem, bolje institucije, masovniju vakcinaciju i dr. što utiče na smanjenje stope smrtnosti. Na drugoj strani, viši nivo razvijenosti implicira starije stanovništvo, viši stepen urbanizacije, veću učestalost putovanja i dr. što može da utiče na veću smrtnost. Stoga teorijski nije izvesno kako nivo razvijenosti zemlje utiče na višak smrtnosti.

Uspešnost suočavanja sa pandemijom ali i uspešnost u lečenju od drugih bolesti u periodu pandemije zavisi od karakteristika zdravstvenog sistema, kao što su visina ulaganja u zdravstvo i kvalitet zdravstvenih usluga [6]. Očekujemo da veća ulaganja u javni zdravstveni sistem u godinama koje su prethodile pandemiji kao i povećanje ulaganja u periodu pandemije utiče na manji rast stope smrtnosti. Osim visine ulaganja, očekuje se da stopa viška smrtnosti zavisi od kvaliteta zdravstvenih usluga u godinama koje su prethodile pandemiji, koji je meren stopom smrtnosti odojčadi u predpandemijskom periodu. Konačno, kao pokazatelj spremnosti zdravstvenih sistema različitih zemalja za suočavanje sa velikim zdravstvenim izazovima, kao što je pandemija, korišćen je opšti indeks zdravstvene bezbednosti (Global health security index).

Karakteristike stanovništva utiču na verovatnoću zaražavanja u periodu pandemije, kao i na ishode razboljevanja. Na osnovu empirijskih istraživanja sledi da je starije stanovništvo naročito izloženo visokom riziku od smrtnosti u slučaju razboljevanja od COVID-19, pa se stoga očekuje da veće učešće starijeg stanovništva utiče na veći porast smrtnosti. Snažniji kolektivizam može da indicira veću solidarnost između ljudi, što poboljšava zdravstvene ishode, ali i slabije poštovanje epidemioloških ograničenja jer su u kolektivističkim društvima neformalna pravila ponašanja (običaji) važniji od zakonskih propisa [4], što implicira veću verovatnoću zaražavanja.

Većina mera za ublažavanje zdravstvenih posledica pandemije nalazi se u nadležnosti države. U periodu pandemije država uvodi epidemiološka ograničenja, obezbeđuje dodatna sredstva za zdravstvo, organizuje medijsku kampanju kojom postiže ponašanje kojim se ublažava pandemija, obezbeđuje vakcine i podstiče ljude na vakcinaciju i dr. Takođe, državne politike u predpandemijskom periodu su ključno uticale na nivo razvijenosti zemlje i njenog zdravstvenog sistema, a u određenoj meri uticale su i na sistem vrednosti građana. Stoga kvalitet upravljanja društvom od strane države u predpandemijskom periodu predstavlja važan indikator njene sposobnosti za suočavanje sa pandemijom. Očekujemo da država u kojoj se pravila usvajaju na osnovu najbolje međunarodne prakse i racionalnih argumenata (kvalitet regulative), a pri tome se formalna pravila striktnije poštuju (vladavina prava, efektivnost države), a gde je korupcija manja i dr. bude uspešnija u ublažavanju negativnih efekata pandemije, uključujući i smrtnost ([3], [6]).

Brzina širenja virusa, a time i broj umrlih zavisi od prosečne temperature jer se sa povećanjem temperature smanjuju učestalost boravka većeg broj ljudi u zatvorenim prostorima, čime opada verovatnoća zaražavanja [6]. Stoga je kao jedna od objašnjavajućih promenljivih u modelu korišćena prosečna mesečna temperatura u glavnim gradovima zemalja.

## 2. PODACI I EKONOMETRIJSKI MODEL

Podaci o višku smrtnosti za uzorak od 77 zemalja u periodu mart 2020 - decembar 2021. godine preuzeti su iz baze koju objavljuje *The Economist's tracker*. Formirana je stopa viška smrtnosti pojedinačnih zemalja koja pokazuje koliko je povećana smrtnost na 100.000 stanovnika u periodu pandemije na mesečnom nivou, u odnosu na uobičajeni broj umrlih u tom mesecu pre pandemije. Dodatno, formirana je i kumulativna stopa viška smrtnosti koja meri koliko je povećana smrtnost od početka pandemije do kraja određenog meseca u odnosu na uobičajeni broj umrlih u predpandemijskom periodu. U okviru epidemioloških ograničenja, na osnovu dnevnih podataka o indeksu striktnosti (*Oxford Stringency index - OSI*) koje objavljuje Univerzitet u Oksfordu, formirana su dva pokazatelja koja obuhvataju intenzitet epidemioloških mera (mesečni proseki) i varijacije epidemioloških mera (standardna devijacija intenziteta epidemioloških mera od početka pandemije do kraja određenog meseca).

Podaci o kumulativnoj stopi vakcinacije od početka procesa vakcinisanja u pojedinačnoj zemlji do kraja određenog meseca preuzeti su sa sajta *Our World in Data*. Baza podataka sadrži i druge pokazatelje u okviru

karakteristika zdravstvenog sistema kao što su prosečni rashodi za zdravstvo kao % BDP u tri godine koje su prethodile pandemiji, dodatni rashodi za zdravstvo u toku pandemije kao % BDP, smrtnost odojčadi i opšti indeks zdravstvene bezbednosti. Izvori podataka su IMF, Svetska banka i Centar za zdravstvenu bezbednost Džons Hopkins. Takođe, u analizi su korišćeni pokazatelji koji obuhvataju karakteristike pojedinačnih zemalja, kao i njihovih populacija (BDP po stanovniku u dolarima jednake kupovne snage, indikatori kvaliteta institucija (WGI<sup>2</sup>), prosečna mesečna temperatura u glavnom gradu, medijska starost i mera kolektivizma). Glavni izvori podataka su Svetska banka i sajt *Our World in Data*, dok je kao mera kolektivizma korišćen indeks koji je formiran u radu [8].

Problem koji se pojavio pri ocenjivanju značaja kvaliteta institucija je visoka multikolinearnost između različitih pokazatelja, te smo koristeći šest pokazatelja kvaliteta primenom metoda glavnih komponenta (*Principal Component Analysis, PCA*) formirali jednu linearnu kombinaciju – glavnu komponentu *pc1* (karakteristično rešenje veće od 1). Pri tome, svi korišćeni pokazatelji kvaliteta institucija ulaze sa pozitivnim predznakom te veća vrednost pokazatelja *pc1* označava kvalitetnije institucije u posmatranoj zemlji.

Ocenjeni su sledeći modeli:

$$edrate_{i,t} = \alpha_{it} + \beta_1 avgosi_{i,t} + \beta_2 stdosi_{i,t} + \beta_3 pfv_{i,t-1} + \theta amt_{i,t} + \gamma_1 che_i + \gamma_2 ash_i + \gamma_3 inm_i + \gamma_4 glsi_i + \delta_1 mage_i + \delta_2 cm_i + \mu_1 yppp2019_i + \mu_2 pc1_i + \epsilon_{i,t} \quad (1)$$

$$ed_{cumrate}_{i,t} = \alpha'_{it} + \beta'_1 avgosi_{i,t} + \beta'_2 stdosi_{i,t} + \beta'_3 pfv_{i,t-1} + \theta' amt_{i,t} + \gamma'_1 che_i + \gamma'_2 ash_i + \gamma'_3 inm_i + \gamma'_4 glsi_i + \delta'_1 mage_i + \delta'_2 cm_i + \mu'_1 yppp2019_i + \mu'_2 pc1_i + \epsilon_{i,t} \quad (2)$$

Sa ekonometrijskog stanovišta relevantno je da su zavisna promenljiva, kao i objašnjavajuće promenljive koje se odnose na epidemiološke mere i proces vakcinacije (uz parametre  $\beta$ ) mereni na mesečnom nivou i variraju i po zemljama i po vremenu. Dakle, promenljive pod ključnim uticajem vlade, kao i prosečna mesečna temperatura (uz parametar  $\theta$ ) su varirale tokom pandemije, dok promenljive koje se odnose na karakteristike stanovništva, zdravstva i države ne variraju po vremenu, već samo po zemljama. Specifičnost modela je da su promenljive koje ne variraju u vremenu (uz parametre  $\gamma$ ,  $\delta$  i  $\mu$ ) tretirane kao fiksni efekti, dok se  $\alpha_0$  odnosi na konstantu, a  $\epsilon_{it}$  na klasičnu slučajnu grešku modela panel podataka. Opredelili smo se Price-Winsten postupak ocenjivanja modela panel podataka, koji tretira potencijalno prisustvo heteroskedastilnosti i autokorelaciju u modelu. Sprovedeni testovi specifikacije ukazuju na zadovoljavajući kvalitet ocenjenih modela.

### 3. EKONOMETRIJSKI REZULTATI

Na osnovu ekonometrijskog panel Modela 1, ocenjene su determinante viška smrtnosti po mesecima i zemljama (*edrate*). Prema ocenama Modela 1 (kolone 1-3 u Tabeli 1) sledi da je stopa viška smrtnosti manja u zemljama koje u intenzivnijoj menjale intenzitet epidemioških ograničenja, dok intenzitet ograničenja uglavnom nije statistički značajan ili ima pogrešan znak<sup>3</sup>. Stopa viška smrtnosti je manja u zemljama koje su u predpandemijskom periodu imale manju smrtnost odojčadi (*inm*) i koje su više izdvajale za zdravstvo (*che*), što ukazuje da je bolji zdravstveni sistem uticao na manji rast ukupne smrtnosti u periodu pandemije. Osim toga, vanredno povećanje izdataka za zdravstvo u periodu pandemije (*ash*) uticalo je na manju stopu smrtnosti. Zbog relativno visoke pozitivne linearne zavisnosti između *che* i *ash* (koeficijent korelacije 0,51) uključivanje obe promenljive u model ima za posledicu da jedna od njih ima pogrešan znak ili da nije statistički značajna. Iz istog razloga opšti indeks zdravstvene bezbednosti (*ghli*) ima očekivani znak i statistički je značajan samo ako se izdvajanja za zdravstvo iz predpandemijskog perioda (*che*) isključe iz modela. Bolji kvalitet upravljanja od strane države (*pc1*), odnosno bolje institucije u predpandemijskom periodu, imale su snažan uticaj na smanjenje stope viška smrtnosti. Suprotno, viši stepen kolektivizma uticao je na veću stopu viška smrtnosti, ali je taj uticaj statistički značajan samo ako se iz jednačine isključi glavna komponenta kojom se meri kvalitet institucija. Ovakav rezultat je očekivan jer je u prethodnim istraživanjima [4] dobijeno da su formalne institucije slabije u društvima u kojima je kolektivizam snažniji (u našem uzorku korelacija *cm* i *pc1* je -0,69). Osim kolektivizma na stopu viška smrtnosti značajno je uticala i prosečna starost stanovništva – u zemljama koje su imale starije stanovništvo povećanje smrtnosti tokom pandemije je bilo veće. Dobijeno je da je stopa viška smrtnosti u razvijenim zemljama niža, što znači da su preovladale prednosti ovih zemalja (kvalitetnije institucije, bolji zdravstveni sistem i dr.) u odnosu na njihove slabosti (starije stanovništvo, veća urbanizacija i dr.). Prema Modelu 1 stopa vakcinacije u prethodnom mesecu (*pfv<sub>it-1</sub>*) nije statistički značajno uticala na stopu

<sup>2</sup> Baza podataka "Worldwide Governance Indicators", Svetska banke.

<sup>3</sup> Razlog za to bi mogla da bude visoka linearna zavisnost između *avgosi* i *stdosi* (koeficijent korelacije je 0,52).

smrtnosti u narednom mesecu ( $edrate_{it}$ ), a nije ni statistički značajan uticaj prosečne temperatura u određenom mesecu ( $amt_{it}$ ).

U jednačinama 4-7 prikazani su rezultati ocenjivanja determinanti kumulativne stope viška smrtnosti (Model 2). Rezultati dobijeni za stopu kumulativne smrtnosti su velikim delom slični rezultatima koji su dobijeni za stopu smrtnosti po mesecima (jednačine 1-3). Starija populacija, slabiji zdravstveni sistem (veća smrtnost odojčadi) i snažniji kolektivizam utiču na rast stope smrtnosti, dok veća ulaganja u zdravstvo u periodu pandemije, viši nivo razvijenosti i bolje formalne institucije (bolji kvalitet upravljanja od strane države) utiču na smanjenje kumulativne stope smrtnosti. Za razliku od Modela 1 u Modelu 2 kumulativna stopa smrtnosti statistički značajno opada sa rastom kumulativne stope vakcinacije, kao i sa rastom prosečne temperature.

**Tabela 1:** Ekonometrijska ocena determinanti viška smrtnosti

Promenljive	$edrate_{it}$	$edrate_{it}$	$edrate_{it}$	$ed\ cumrate_{it}$	$ed\ cumrate_{it}$	$ed\ cumrate_{it}$	$ed\ cumrate_{it}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$avgosi_{it}$	0.218*** (0.038)	0.213*** (0.037)	0.223*** (0.0377)	-0.202 (0.219)		0.104 (0.2590)	-0.0238 (0.2570)
$stdosi_{it}$	-0.409*** (0.112)	-0.408*** (0.109)	-0.375*** (0.1130)	-7.897*** (0.701)	-8.220*** (0.638)	-9.985*** (0.7440)	-8.188*** (0.8350)
$pfv_{it-1}$	-0.0111 (0.019)	-0.0117 (0.018)	-0.0191 (0.0192)	-1.278*** (0.130)	-1.270*** (0.129)		-1.508*** (0.1460)
$che_i$	-0.334 (0.351)	-0.841** (0.332)	-0.13 (0.3440)			4.607* (2.3810)	5.900** (2.4860)
$ash_i$	-0.377*** (0.103)		-0.478*** (0.1110)	-3.229*** (0.682)	-3.305*** (0.684)	-4.562*** (0.8570)	-4.482*** (0.9040)
$inm_i$	0.328 (0.2320)	0.436** (0.2080)	1.007*** (0.2150)	2.003 (1.2180)	2.042* (1.2250)	9.446*** (1.3750)	8.275*** (1.5030)
$glsi_i$	0.0263 (0.0855)	-0.0314 (0.0810)	-0.0695 (0.0849)	3.062*** (0.5240)	2.999*** (0.5180)	0.327 (0.5880)	0.182 (0.6200)
$mage_i$	1.006*** (0.175)	1.032*** (0.172)	1.093*** (0.1810)	6.429*** (0.976)	6.502*** (0.990)	5.150*** (1.2170)	5.376*** (1.3020)
$cm_i$	-0.0342 (0.0518)	-0.0746 (0.0503)	0.183*** (0.0472)			1.968*** (0.3050)	2.121*** (0.3230)
$amt_{it}$	0.00744 (0.0061)	0.00596 (0.0060)	0.00816 (0.0063)	-0.132*** (0.0372)	-0.133*** (0.0372)	-0.0914** (0.0435)	-0.0728 (0.0443)
$yppp2019_i$	- (0.0000)	-0.000137*** (0.0000)	-0.000158*** (0.0000)				
$pcl_i$	-2.497*** (0.6220)	-2.314*** (0.5950)		-0.000350* (0.0002)	-0.000375* (0.0002)		
$pcl_i$				-36.56*** (3.4340)	-36.21*** (3.3630)		
<i>Constant</i>	-18.79* (9.7430)	-13.61 (9.1490)	-39.51*** (8.9540)	-6.388 (52.8100)	-7.552 (53.4500)	24.03 (58.7500)	-9.712 (63.7900)
Observations	1.251	1.293	1.251	1.503	1.503	1.320	1.251
R-squared	0.141	0.132	0.105	0.343	0.344	0.232	0.292
Number of nwcountrie	60	62	60	72	72	60	60

Napomena: U zagradama su navedene standardne greške; \*\*\* p<0.01. \*\* p<0.05. \* p<0.1

#### 4. ZAKLJUČAK

U cilju ublažavanja negativnih zdravstvenih posledica pandemije skoro sve zemlje su uvele snažna epidemiološka ograničavanja, povećale izdatke za javno zdravstvo i primenile politiku masovne vakcinacije. Međutim, uspeh u ublažavanju zdravstvenih posledica pandemije zavisio je i od performansi zdravstvenog sistema pre pandemije, nivoa razvijenosti zemalja, kvaliteta institucija, karakteristika stanovništva i prosečne temperature. Nakon dve godine pandemije moguće je na osnovu statističkih podataka oceniti da li su i u kojoj meri preduzete mere i karakteristike zdravstva, stanovništva i države uticale na upešnost zemalja u suzbijanju pandemije. Kao sintetički pokazatelj uspešnosti zemlja u suzbijanju zdravstvenih posledica pandemije korišćena je stopa viška smrtnosti koja pokazuje povećanje broja umrlih na sto hiljada stanovnika u odnosu na uobičajeni broj umrlih u predpandemijskom periodu. Višak smrtnosti predstavlja sintetički pokazatelj uspešnosti društva u suočavanju sa pandemijom, jer obuhvata umrle od COVID-19, ali u povećanje smrtnosti od drugih bolesti u periodu pandemije. Na osnovu podataka za preko 60 zemalja, primenom ekonometrijskog panel modela, dobijeno je da je su manju stopu viška smrtnosti imale zemlje koje su primenile strožije epidemiološke mere, a koje su pri tome češće varirale strogost mera, kako bi ih prilagodile epidemiološkoj situaciji. Viša stopa vakcinacije i povećanje izdataka za zdravstvo u periodu pandemije takođe su uticali na nižu stopu viška smrtnosti. Osim mera koje su preduzete u toku pandemije stopa viška smrtnosti je zavisila od kvaliteta zdravstva i institucija i karakteristika stanovništva koje su postojale pre pandemije. U zemljama koje su pre pandemije imale kvalitetnije zdravstvo i institucije, mlađe stanovništvo, manji stepen kolektivizma (veći individualizam) stopa viška smrtnosti je manje povećana. Takođe, viša temperatura je povezana sa nižom kumulativnom stopom viška smrtnosti.

#### LITERATURA

- [1] Agrawal. V., J. H. Cantor N. Sood and C. M. Whaley. 2021. "The Impact of the COVID-19 Pandemic and Policy Responses on Excess Mortality." NBER Working Papers 28930. <http://www.nber.org/papers/w28930>
- [2] Aizenman J., A. Cukierman, Y. Jinjarak and W.Xin. 2021. "International Evidence on Vaccines and the Mortality to Infections Ratio in the Pre-Omicron Era." NBER Working Papers 29498. <http://www.nber.org/papers/w29498>
- [3] Aizenman J., A. Cukierman, Y. Jinjarak, S. Nair-Desai and W.Xin. 2022. "A Two Covid-19 Years Quartile Comparison of Official with Excess Mortality: Voice and Accountability and the Impact of Vaccines" NBER Working Paper No. 29778
- [4] Alesina. A.. & Giuliano. P. (2015). Culture and institutions. *Journal of Economic Literature*. 53(4).898–944
- [5] Brodeur A. Gray D. Islam A. Bhuiyan S. A literature review of the economics of COVID-19. *Journal of Economic Surveys*. 2021;35:1007–1044. <https://doi.org/10.1111/joes.12423>
- [6] Chang. D.. Chang. X.. He. Y. et al. The determinants of COVID-19 morbidity and mortality across countries. *Sci Rep* 12. 5888 (2022). <https://doi.org/10.1038/s41598-022-09783-9>
- [7] Mulligan. C. B.. Deaths of Despair and the Incidence of Excess Mortality in 2020 (December 23. 2020). NBER Working Paper 28303. <http://www.nber.org/papers/w28303>
- [8] Webster, G. D., Howell, J. L., Losee, J. E., Mahar, E. A., & Wongsomboon, V. (2021). Culture, COVID-19, and collectivism: A paradox of American exceptionalism?. *Personality and Individual Differences*, 178, 110853.





## FISKALNI I TEKUĆI BILANS U ZEMLJAMA CENTRALNE I ISTOČNE EVROPE: PRIMENA DINAMIČKOG PANEL MODELA

## FISCAL AND CURRENT ACCOUNT BALANCE IN CENTRAL AND EASTERN EUROPEAN COUNTRIES: DYNAMIC PANEL MODEL APPROACH

JELENA RAŠKOVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, jelena.raskovic@ekof.bg.ac.rs

**Rezime:** Cilj ovog rada je da empirijski istraži uticaj fiskalnog bilansa na tekući bilans u zemljama Centralne i Istočne Evrope (CIE) koje su članice EU. Zbog visoke perzistentnosti salda tekućeg računa, empirijska analiza se zasniva na korišćenju dinamičkog panel modela i Hsiao & Anderson pristupu instrumentalnih promenljivih za ocenu parametara modela. Rezultati ukazuju da postoji prilično jak pozitivan uticaj fiskalnog bilansa na tekući bilans ( $\beta=0,33$ ), potvrđujući hipotezu blizanačkog deficita.

**Ključne reči:** blizanački deficit, Rikardijanska jednakost, dinamički panel model, Hsiao&Anderson pristup instrumentalnih promenljivih

**Abstract:** The aim of this paper is to empirically investigate the impact of the fiscal balance on the current account in the Central and Eastern European (CEE) countries that are members of the EU. Due to the high persistence of the current account balance, the empirical analysis is based on the use of a dynamic panel model and the Hsiao&Anderson estimation approach of instrumental variables. The results indicate that there is fairly strong and positive relationship between these two balances ( $\beta=0,33$ ) confirming the twin deficit hypothesis.

**Keywords:** twin deficit, Ricardian equivalence, dynamic panel model, Hsiao&Anderson instrumental variable approach

### 1. UVOD

Od sredine 2000-ih, naročito tokom finansijske i ekonomske krize od 2007. do 2009. godine, unutrašnje i spoljnoekonomske neravnoteže su se povećale u mnogim zemljama Evropske Unije. Ovo je ponovo aktuelizovalo debatu o tome da li fiskalni deficiti utiču na spoljnoekonomsku ravnotežu. Stoga, fiskalna politika koju sprovode vlade mogu imati značajne implikacije na tekući račun. Predmet ovog istraživanja je testiranje uticaja fiskalnog bilansa na tekući bilans u zemljama CIE koje su članice EU u periodu od 2000. do 2019. godine. Reč je o malim otvorenim ekonomijama, koje su visoko zavisne od inostranog kapitala, kao i uvoza tehnologije i često sa visokim učešćem javnog duga u ukupnom BDP. Fiskalna politika kroz povećanje trošenja tj. kroz budžetski deficit, ima značajne implikacije na spoljnoekonomsku ravnotežu u otvorenijim privredama [8]. Zbog toga je ispitivanje efekata fiskalne politike na spoljnoekonomsku ravnotežu u ovim zemljama vrlo značajna za kreatore makroekonomske politike. Sa istočnim proširenjem EU, činilo se da centralno-istočni region sustiže Zapadnu Evropu u pogledu ekonomskih performansi. Međutim, finansijska i ekonomska kriza iz 2008. godine, donekle je poremetila dotadašnji trend konvergencije i dovela do povećanja razlika između ekonomskih performansi pojedinih zemalja.

U ekonomskoj teoriji postoje dva različita pristupa ovoj relaciji. Prvi se odnosi na hipotezu o blizanačkom deficitu, koja predviđa da su veliki budžetski deficiti uglavnom praćeni visokim deficitima tekućeg računa. Jedno od mogućih teorijskih objašnjenja uzročne povezanosti fiskalnog i tekućeg bilansa izraženo je kroz model preklapajućih generacija sa konačnim vremenskim horizontom [9]. Po ovom pristupu, pad javne štednje usled smanjenja poreza, dovešće do povećanja privatne štednje, ali i do povećanja potrošnje, što će izazvati pad nacionalne štednje. To dalje uslovljava rast kamatne stope, istiskujući domaće investicije i posebno smanjuje dugoročne investicije. U malim otvorenim ekonomijama, to utiče na priliv kapitala u obliku stranih direktnih investicija i kredita, zatim apresijacije deviznog kursa, i konačno pogoršanje tekućeg

bilansa. Ovo odgovara kenzijanskom pristupu. Sa druge strane, ukoliko se tržišni učesnici ponašaju racionalno, pad javne štednje, kao rezultat smanjenja poreza, uz konstantan nivo javne potrošnje, dovelo bi samo do povećanja privatne štednje, jer tržišni učesnici očekuju rast poreza u budućnosti. Bilans tekućeg računa se neće promeniti, jer će rast privatne štednje, pri konstantnoj kamatnoj stopi biti dovoljan da se izbegne zaduživanje u inostranstvu. Ovo poslednje je u skladu sa rikardijanskom teorijom.

U dosadašnjoj empirijskoj literaturi, preovladava kenzijanski pristup, tj. da fiskalni deficit značajno utiče na deficit tekućeg bilansa. Literatura je prilično obimna, i uključuje i istraživanja koja istovremeno obuhvataju i zemlje u razvoju i razvijene zemlje ([1], [6], [12]). Hipoteza blizanačkog deficita za zemlje CIE je, do sada, empirijski bila obuhvaćena u nekoliko radova [10], [11], [13], [15], [16].

Ovaj rad ima za cilj da istraži vezu između fiskalnog i tekućeg bilansa i utvrdi da li se potrošači u zemljama CIE koje su članice EU ponašaju kenzijanski ili rikardijanski. Panel analiza je sprovedena na uzorku 11 zemalja CIE, i to, Bugarskoj, Estoniji, Letoniji, Litvaniji, Hrvatskoj, Mađarskoj, Poljskoj, Rumuniji, Sloveniji, Slovačkoj i Češkoj. U ovom istraživanju su korišćeni godišnji podaci za period 2000-2019. godine. Hipoteza blizanačkog deficita je vrlo važna za zemlje CIE, jer su u procesu tranzicije ka tržišnoj ekonomiji prošle kroz visoke budžetske i tekuće neravnoteže.

Doprinos ovog istraživanja postojećoj literaturi ogleda se u dva pravca. Panel analiza podataka za 11 zemalja CIE članica EU, pruža specifičnu prednost u pogledu dubine i obuhvatnosti podataka u odnosu na prethodne analize za ovu grupu zemalja. U dosadašnjoj literaturi, ove zemlje su uglavnom, samo parcijalno testirane ili kao deo neke veće grupe zemalja. Drugo, ova analiza ima i praktičnu primenu, jer ovi rezultati pokazuju na koji način i u kojoj meri fiskalna politika ima uticaj u korigovanju spoljnoekonomskog bilansa.

Struktura ovog rada je organizovana na sledeći način. Detaljna metodologija koja je korišćena u ovom radu je predstavljena u poglavlju 2, empirijski podaci u poglavlju 3 i zaključna razmatranja u poglavlju 4.

## 2. METODOLOGIJA

Kao što je u uvodnom poglavlju već napomenuto, ovaj rad ima za cilj da testira uticaj fiskalnog bilansa na saldo tekućeg bilansa, kroz testiranje hipoteze o blizanačkom deficitu. Empirijski model za testiranje ove hipoteze predstavljen je u jednačini (1):

$$TB_{it} = \alpha + \beta FB_{it} + \theta Z_{it} + \varepsilon_{it} \quad (1)$$

gde je  $TB_{it}$  tekući bilans (%BDP) za zemlju  $i$  ( $i = 1, \dots, N$ ) u vremenu  $t$  ( $t = 1, \dots, T$ ),  $FB_{it}$  je konsolidovani fiskalni bilans ili javna štednja (T-G) predstavljen kao učešće u BDP,  $Z_{it}$  je skup kontrolnih promenljivih,  $\alpha$  je konstanta, a  $\varepsilon_{it}$  slučajna greška modela. Glavna razlika između teorije rikardijanske jednakosti i hipoteze blizanačkih deficita odnosi se na znak i statističku značajnost koeficijenta  $\beta$  koji predstavlja odgovor salda tekućeg bilansa na varijacije u fiskalnom bilansu. Hipoteza blizanačkih deficita predviđa da pogoršanje budžetskog bilansa dovodi do pogoršanja stanja na tekućem računu ( $\beta > 0$ ), dok rikardijanska teorija predviđa da je  $\beta = 0$ .

Kontrolne promenljive, koje pored fiskalnog bilansa determinišu tekući bilans su: rast BDP, bruto nacionalni dohodak po glavi stanovnika (relativan odnos prema BND pc u SAD), promena efektivnog realnog deviznog kursa, odnosi razmene, otvorenost zemlje za trgovinu, energetska zavisnost zemlje, stopa izdržavanog stanovništva i plasirani krediti stanovništvu kao učešće u BDP. Izbor kontrolnih promenljivih je zasnovan na brojnim istraživanjima u literaturi koji analiziraju determinante tekućeg bilansa, a jedan od novijih radova koji je imao najveći uticaj na izbor kontrolnih promenljivih u ovom istraživanju je [2].

Ekonometrijski metod koji je korišćen u ovom radu je panel analiza podataka. Pre svega, ubacivanjem veštačke promenljive ( $V$ ) među regresore, obuhvaćen je i uticaj finansijske krize u periodu od 2006. do 2009. godine. U prvom koraku, u osnovnom modelu, ispituje se značajnost individualnih karakteristika zemalja (fiksni efekata), kao i vremenskih efekata. Hausmanov test je pokazao da su fiksni efekti prisutni u modelu, dok vremenski efekti nisu značajni. S obzirom da je utvrđena i visoka perzistentnost zavisne promenljive odnosno tekućeg bilansa, relevantan model za testiranje postaje dinamički panel model:

$$TB_{it} = \alpha + \gamma TB_{i,t-1} + \beta FB_{it} + \theta Z_{it} + c_i + \varepsilon_{it} \quad (2)$$

gde  $c_i$  predstavlja individualne karakteristike zemlje.

S obzirom na to da se pomaknuta vrednost zavisne promenljive pojavljuje u skupu objašnjavajućih promenljivih, stroga egzogenost nezavisnih promenljivih više ne važi, a ONK metod ocenjivanja sa veštačkim promenljivim, postaje nekonzistentan kada  $N \rightarrow \infty$ , a  $T$  je fiksno. Takođe, rast bruto domaćeg proizvoda, kao i realni devizni kurs su potencijalno endogene promenljive, tako da endogenost može proizići iz drugih objašnjavajućih promenljivih tj. postojanja obrnutih uzročnosti (simultanost), pa mogu biti korelisani sa slučajnom greškom. Metod instrumentalnih promenljivih i uopšteni metod momenata (engl.

general method of moments ili GMM), istovremeno rešavaju i problem perzistentnosti zavisne promenljive i problem endogenosti.

U ovom istraživanju, korišćen je Hsiao&Anderson pristup instrumentalnih promenljivih (HA pristup), predstavljen u [3], koji se svodi se na ocenu parametara korišćenjem dvostepenih najmanjih kvadrata. U prvom koraku, uzimaju se prve difference originalnog modela, čijom transformacijom se eliminišu i individualni efekti i konstanta (jednačina 3), ali i dalje ostaje korelacija između prve difference pomaknute zavisne promenljive i slučajne greške (koja je sada proces pokretnih sredina prvog reda). Sa uklonjenim individualnim fiksnim efektima, možemo oceniti parametre modela pomoću instrumentalnih promenljivih, konstruisanjem instrumenata za pomaknutu zavisnu promenljivu i ostale potencijalno endogene objašnjavajuće promenljive. Adekvatne instrumentalne promenljive i za pomaknutu vrednost zavisne promenljive i druge objašnjavajuće promenljive, mogu biti njihove docnje. Ukoliko je  $\varepsilon$  slučajan proces, druga i treća docnja mogu biti visoko korelisane sa pomaknutom vrednošću zavisne promenljive tj. prvom diferencom, ali nekorelisane sa greškom modela.

$$\Delta TB_{it} = \gamma \Delta TB_{i,t-1} + \beta FB_{it} + \theta \Delta Z_{it} + \Delta \varepsilon_{it} \quad (3)$$

Pomaknutu vrednost tekućeg bilansa smo instrumentalizovali sa njenom drugom i trećom docnjom, a rast BDPa sa njegovom prvom docnjom i privatnim investicijama. Realni devizni kurs je takođe, potencijalno endogen, ali budući da HA metod nije pouzdan za veliki broj instrumenata, kao i da realni kurs ima odložen pun efekat na tekući bilans, u model je uključen sa docnjom, čime smo izbegli njegovu eventualnu korelisanost sa greškom modela.

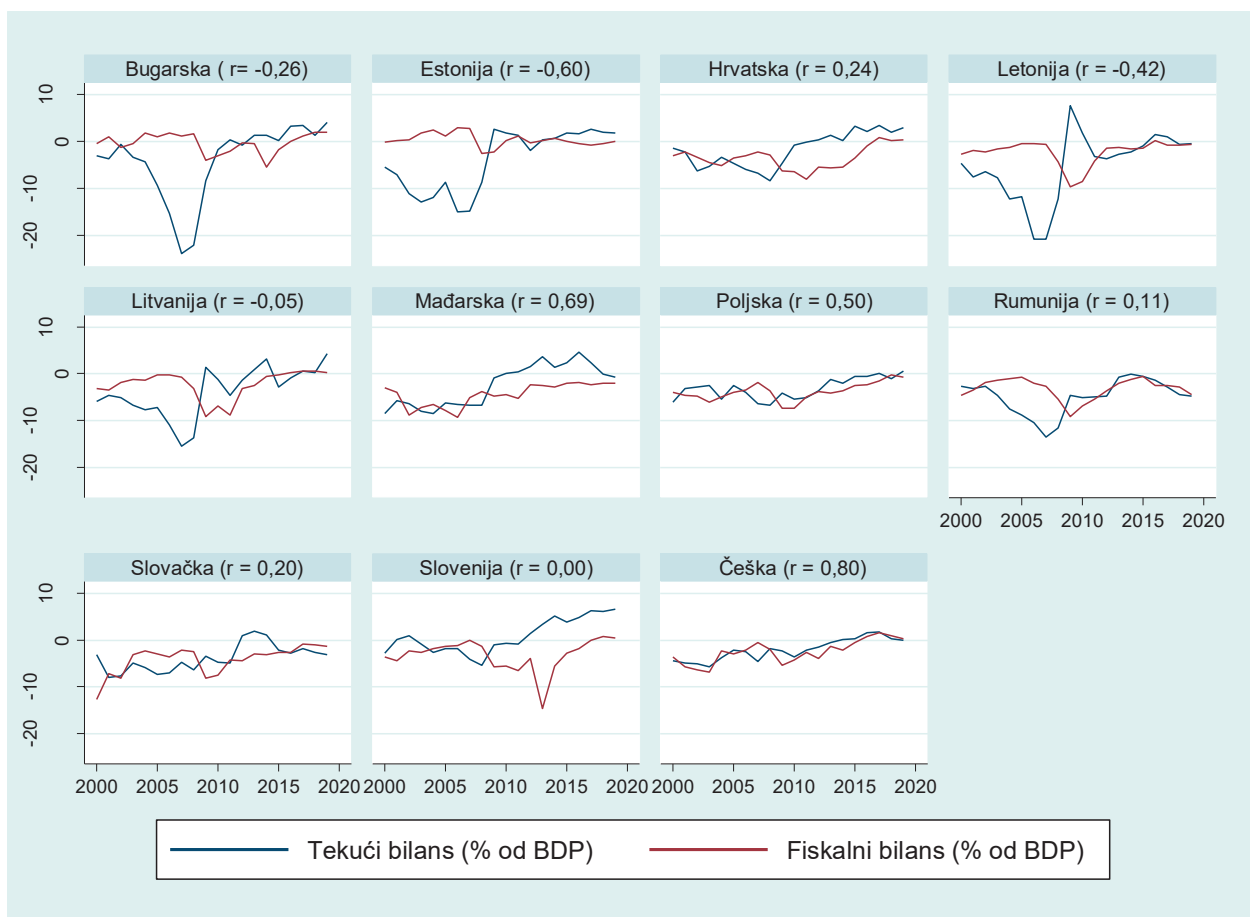
Za razliku od HA pristupa instrumentalnih promenljivih, GMM pristup koji su razvili Arellano&Bond [4], kao i unapređena verzija sistemski GMM ([5], [7]), omogućava da sve objašnjavajuće promenljive mogu biti potencijalno endogene. Iako ovaj metod daje efikasnije ocene od HA pristupa, problem je upotreba velikog broja instrumenata pri malom  $N$ . Kada broj instrumentalnih promenljivih premaši broj zemalja, GMM ocene "pate" od pristrasnosti prevelikog broja informacija [14]. Ovaj pristup je najpogodniji za uzorke sa velikim  $N$ , a malim  $T$ . S obzirom da ovo istraživanje analizira samo 11 zemalja, za period od 20 godina, nismo uspeli da dobijemo pouzdane ocene parametara modela.

### 3. EMPIRIJSKI REZULTATI

Prvo su predstavljeni prosti koeficijenti korelacije po zemljama na grafikonu 1, koji pomažu da se deskriptivnom analizom podataka dobiju neke zakonomernosti u kretanju fiskalnog i tekućeg bilansa. Ukupan koeficijent korelacija za sve zemlje je blizu 0 i iznosi -0,05. Ovo je i očekivan nivo korelacije, jer se na grafikonu 1 jasno vidi da u većini zemalja u predkriznom periodu, a pogotovu za vreme svetske finansijske krize iz 2008, dolazi do potpunog razmimoilaženja ova dva bilansa, dok nakon krize se vraćaju na približno istu liniju trenda. Uočava se jaka pozitivna veza između ova dva bilansa u Češkoj, Poljskoj i Mađarskoj (0,80, 0,50 i 0,69 respektivno), dok se u baltičkim zemljama Estoniji i Letoniji, a i u Bugarskoj uočava jaka negativna veza (-0,60, -0,42 i -0,24). S obzirom da se iz podataka jasno vidi da tokom krize dolazi do prekida veze fiskalnog i tekućeg bilansa, ovaj rezultat je i očekivan, jer su Estonija, Letonija i Bugarska među zemljama CIE koje su tokom svetske finansijske krize 2007- 2009. godine, zabeležile najveće deficite na tekućem računu.

Period finansijske krize, obuhvaćen je veštačkom promenljivom  $V$ , čime je isključen njen uticaj na konačne rezultate prilikom ekonometrijskog ocenjivanja panela.

Što se tiče rezultata dinamičkog panela, HA pristup instrumentalnih promenljivih nam sugerise da je koeficijent fiskalnog bilansa pozitivan i visoko statistički značajan na nivou od 1%, podržavajući hipotezu blizanačkog deficita tj. kenzijanski pristup. Naime,  $\beta = 0,33$ , što ukazuje na to da rast konsolidovanog fiskalnog deficita (kao %BDP), ima tendenciju da pogorša deficit tekućeg bilansa (kao % BDP). Ovo implicira odbacivanje teorije rikardijanske jednakosti, što znači da je javni dug zapravo teret za buduće poreske obveznike. Shodno tome, to predstavlja problem za finansiranje budžetskih rashoda i može dovesti čak i do fiskalne krize.



**Grafikon 1:** Tekući i fiskalni bilans po zemljama

Osim konsolidovanog fiskalnog bilansa, uključene su i druge objašnjavajuće promenljive koje determinišu tekući bilans. Realni devizni kurs i stopa izdržavanog stanovništva su prisutne u modelu sa jednom docnjom, zbog vremenskog razmaka uticaja na zavisnu promenljivu i potencijalne endogenosti. Takođe, analizom prirode vremenskih serija, utvrđeno je da vrednost kredita plasiran privatnom sektoru prikazan kao učešće u BDP, kao i stopa izdržavanog stanovništva, imaju po jedan jedinični koren, tako da se javljaju u formi prve diference. Koeficijent uz docnju tekućeg bilansa je takođe visok i visoko statistički značajan, ukazujući na opravdanost primene dinamičkog panela.

Rast BDP ima najveći uticaj na tekući bilans i to negativan, jer sa ubrzanim rastom, raste i uvoz repromaterijala i investicija u tehnologiju, što dovodi do većeg deficita tekućeg bilansa. Dalje, rezultati regresije ukazuju da poboljšanje odnosa razmene ima pozitivan efekat na tekući bilans. Ovo je u skladu sa Harberger-Laursen-Metzler efektom, koji predviđa da bi pad tekućeg dohotka kao posledica pogoršanja odnosa razmene, uticao na smanjenje privatne štednje, a samim tim i na pad tekućeg bilansa.

Rezultati takođe pokazuju da promena koeficijenta izdržavanog stanovništva ima očekivani negativni predznak, što ukazuje na važnost demografskih faktora za stanje tekućeg bilansa. Pozitivan predznak uz relativni bruto nacionalni dohodak, implicira da ispod prosečni dohodak po glavni stanovnika pogoršava saldo tekućeg bilansa. Pretpostavlja se da će siromašnije zemlje rasti brže od proseka i da se stoga zadužuju na ime budućih prihoda.

Uticaj realnog efektivnog deviznog kursa na tekući bilans je dvosmislen. Naime, u procesu ubrzanog rasta radi sustizanja razvijenih ekonomija, dolazi do apresijacije realnog deviznog kursa (Balaša-Samjuelsonov efekat), a samim tim i rasta očekivanog dohotka. Zbog toga, shodno hipotezi o izravnjanju potrošnje, apresijacija deviznog kursa može dovesti do poboljšanja tekućeg bilansa. Sa druge strane, realna apresijacija deviznog kursa, poskupljuje izvoz, a pojeftinjuje uvoz, pogoršavajući spoljni bilans. U našem primeru, vidimo da je koeficijent negativan, ali statistički neznačajan, što je potencijalno rezultat poravnjanja ova dva efekta. Takođe, razvoj finansijskih tržišta meren kao učešće plasiranih kredita privatnog sektora u BDP, ima negativan, ali statistički neznačajan uticaj na saldo tekućeg bilansa.

**Tabela 1:** Dinamički panel model sa fiksnim efektima (HA pristup instrumentalnih promenljivih)

Promenljive	Koeficijent	Standardna greška	P - vrednost
TB <sub>t-1</sub>	0,5304***	0,0774	0,000
<b>FB<sub>t</sub></b>	<b>0,3332***</b>	<b>0,1206</b>	<b>0,006</b>
BDP_rast <sub>t</sub>	-0,7815***	0,1150	0,000
BND <sub>t</sub>	0.0934*	0,0496	0,061
RDK <sub>t-1</sub>	-0,0367	0,0448	0,414
OR <sub>t</sub>	0,2854*	0,2020	0,099
OTV	0,0557*	0,0317	0,081
ENERG <sub>t</sub>	-0.0571**	0,0249	0,023
KS <sub>t</sub>	-0.0583	0,0418	0,165
IS <sub>t-1</sub>	-1.2492*	0,6960	0,075
V	-1.1614***	0,4025	0,004
<b>Testovi specifikacije</b>			
Sargan-Hansen test	3,754		0,1531
Stock-Yogo test	10,286	Max pristrasnost	10%
Test endogenosti	10,767		0,005

\*\*\* $p < 0,01$ , \*\*  $p < 0,05$ , \*  $p < 0,10$

V – veštačka promenljiva koja uzima nenulte vrednosti 1 za 2006. i 2009. godinu i 2,5 za 2007. i 2008. godinu.

Konačno, zemlje koje su više zavisne od uvoza nafte, imaju veći deficit tekućeg bilansa, tako da je predznak negativan. Otvorenost zemlje za trgovinu, značajno pozitivno utiče na tekući bilans, što znači da sa rastom otvorenosti zemlje, raste i suficit ili se smanjuje deficit na tekućem računu.

Što se tiče dijagnostičkih provera, sproveden je Hansen-Sarganov test o prekomernoj identifikovanosti. Kao što se vidi iz tabele 1, ne možemo odbaciti nultu hipotezu o nekorelisanosti instrumenata sa greškom modela (na nivou značajnosti od 1%). Takođe, Cragg-Donaldova statistika, koristeći Stock Yogo kritične vrednosti, odbacuje nultu hipotezu da su instrumenti slabo korelisanani sa endogenim regresorima (nivo pristrasnosti je 10%). Konačno, test endogenosti potvrđuje da su izabrane potencijalno endogene promenljive zaista endogene.

#### 4. ZAKLJUČAK

U ovom radu se ispituje odnos između konsolidovanog fiskalnog bilansa i tekućeg bilansa u zemljama CIE koje su članice EU u periodu 2000 - 2019. godine, koristeći dinamički panel model sa fiksnim efektima. Naši rezultati dinamičkog panela sugerišu da u ovim zemljama postoji prilično jak pozitivan odnos između ova dva bilansa ( $\beta=0,33$ ), potvrđujući hipotezu blizanačkog deficita. Povećanje fiskalnog deficita neće biti neutralisano povećanjem privatne štednje, već dovodi i do rasta potrošnje, tako da budžetski račun ima uticaj na saldo tekućeg računa. Ovo implicira da se potrošači u zemljama CIE koje su se pridružile EU, u svojim odlukama o potrošnji i štednji, okrenuti više sadašnjosti nego budućnosti, što je u skladu sa kenzijanskim pristupom.

Ovo ima direktan uticaj na način na koji kreatori makroekonomske politike sprovode ekonomske mere. Suficit državnog budžeta ima pozitivan i značajan efekat na saldo tekućeg računa. Rezultati pokazuju da je smanjenje fiskalnog deficita (kao %BDP) povezano sa poboljšanjem tekućeg bilans (kao %BDP). Nalazi za ovu grupu zemalja su u velikoj meri konzistentni sa konvencionalnim pristupom, koji dominira u postojećoj empirijskoj literaturi.

**LITERATURA**

- [1] Afonso, A., Huart, F., Tovar Jalles, J., & Stanek, P. (2022). Twin deficits revisited: A role for fiscal institutions? *Journal of International Money and Finance*, 121, 102506.
- [2] Altayligil, Y. B., & Çetrez, M. (2020). Macroeconomic, institutional and financial determinants of current account balances: A panel data assessment. *Journal of Economic Structures*, 9(1), 49.
- [3] Anderson, T. W., & Hsiao, C. (1981). Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association*, 76(375), 598–606.
- [4] Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2), 277–297.
- [5] Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), 29–51.
- [6] Blanchard, O., & Giavazzi, F. (2002). Current Account Deficits in the Euro Area: The End of the Feldstein-Horioka Puzzle? *Brookings Papers on Economic Activity*, 2002(2), 147–186.
- [7] Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115–143.
- [8] Corsetti, G., & Müller, G. J. (2006). Twin deficits: Squaring theory, evidence and common sense. *Economic Policy*, 21(48), 598–638.
- [9] Jacob A. Frenkel & Assaf Razin. (1992). *Fiscal Policies and the World Economy*. Cambridge, MA: MIT Press.
- [10] Josifidis, K., Dragutinović-Mitrović, R., & Bodor, S. (2021). The effect of fiscal deficits on the external imbalances in the European Union. *Panoeconomicus*, 68(5), 625–652.
- [11] Matej, A. (2019). Interactions between the budget and the current account balance: Twin deficits in selected Central and East European countries. *Journal of Management and Financial Sciences*, 38, 47-60-47–60.
- [12] Mohammadi, H. (2004). Budget deficits and the current account balance: New evidence from panel data. *Journal of Economics and Finance*, 28(1), 39–45.
- [13] Obadic, A., Globan, T., & Nadoveza, O. (2014). Contradicting the twin deficits hypothesis: The role of tax revenues composition. *Panoeconomicus*, 61(6), 653–667.
- [14] Roodman, D. (2009). How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *The Stata Journal*, 9(1), 86–136.
- [15] Tosun, M. U., Iyidogan, P. V., & Telatar, E. (2014). The Twin Deficits in Selected Central and Eastern European Economies: Bounds Testing Approach with Causality Analysis. *Journal for Economic Forecasting*, 17(2), 141–160.
- [16] Turan, T., & Karakas, M. (2018). Asymmetries in twin deficit hypothesis: evidence from CEE countries. *Ekonomický Časopis (Journal of Economics)*, 66(6), 580–597.

## GDP GROWTH DETERMINANTS IN SELECTED POST-TRANSITION EUROPEAN COUNTRIES

MILUTIN JEŠIĆ<sup>1</sup>

<sup>1</sup> University of Belgrade – Faculty of Economics and Business milutin.jesic@ekof.bg.ac.rs

**Abstract:** *In this paper we empirically investigated the determinants of GDP growth in selected post-transition European countries (Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Serbia, Slovakia, Slovenia, Turkey, and Ukraine). Dynamic panel data modeling has been employed, in particular System GMM method. The time span includes 2009 – 2018 period. Findings indicate that GDP growth in these countries is under the influence of factors such as: inflation, public debt, initial level of development, unemployment, structural budget balance, current account balance and Eurozone membership. Policy implications that can be derived from the results are consistent with the policy prescriptions founded in the economic theory, and relate to significance of fiscal responsibility and price stability for economic growth. Therefore, macroeconomic stability is conditio sine qua non of GDP growth.*

**Keywords:** *GDP growth, GDP growth determinants, post-transition European countries, dynamic panel data, System GMM*

### 1. INTRODUCTION

Investigation of GDP growth determinants is one of the most interesting topics for the research in macroeconomics. Different theoretical assumptions triggered conduction of variety of empirical studies which concluded that initial conditions, macroeconomic stability and structural policies are key determinants of GDP growth.

Determinants of GDP growth are specific to countries of different level of development. In addition, institutional framework shapes the boundaries of GDP growth possibilities. Membership in the economic unions is of importance too, since in that case free riding effects can be exploited, but also negative spillover effects are real threat to countries that belong to the specific union.

Motivation for this research can be found in willingness to investigate what are the key determinants of GDP growth in the selected sample of post-transition countries. Few of these countries are already EU member, but some of them are only candidates for the EU membership, or not even the candidate for the EU. Special interest in the research is dedicated to macroeconomic policy influence on GDP growth. Although some policy instruments do not have direct effect on growth, indirect effects can be significant.

The paper is structured in the following way. First, we briefly present literature review. Data sources and variable descriptions are presented in the third part of the paper. The fourth part of the paper is dedicated to the methodology used in the research. Next part of the paper deals with the results. Final section gives brief conclusions and policy implications.

### 2. LITERATURE REVIEW

The topic of GDP growth determinants is one of the most interesting for the research in macroeconomics. In addition, this is important and relevant topic for developing countries and countries that are not fully integrated in specific broader economic or monetary unions. Besides the theoretical papers that deals with these determinants, our focus now is on empirical papers that analyze the GDP growth patterns.

Barro (1999) analyzed factors of economic growth for panel of 100 countries for 1960-1995 period by extended neoclassical growth model and showed that investment share, terms of trade, years of schooling, rule of law index, democracy index and international openness positively influenced the economic growth, while government consumption, total fertility rate, and inflation negatively impact it. Bleaney, Gemmell, and Kneller (2001) investigated the relation between fiscal policies and economic growth on the data for 22 developed countries for 1970-1995 period, and found that productive government expenditure had a significant positive influence on the long-run economic growth rate, while distortionary fiscal policies had a negative influence on the long-run economic growth rate. Bayraktar (2006) analyzed the relationship between growth rates and various macroeconomic indicators in Turkey for 1968-1998 period, and conclude that investments, human capital development and inflation are determinants of economic growth in Turkey. Checherita-Westphal and

Rother (2012) examined the relationship between government debt and economic growth in twelve Euro countries for 1970-2008 period and found that government balance, private savings, and trade openness are positively related to economic growth, while population growth and real interest rates are negatively related. Government debt positively influences the economic growth. However, the squared government debt as a variable negatively impacts the economic growth, meaning that there is threshold effects immanence. After some certain level of government debt, negative effects become dominant. Prochniak (2011) analyzed factors of growth in 10 CEE economies during the 1993-2009 period, by OLS estimation. He found that investment rate, human capital development, financial sector development, high services share in GDP, high share of working age population, development of information, communication and technology (ICT), high private sector share in GDP, economic freedom, and progress in market and structural reforms positively affect GDP growth, while budget deficits, public debt, interest rates, and inflation negatively affect economic growth. Josifidis, Dragutinović-Mitrović and Ivančev (2012) analyzed the heterogeneity of growth in the Western Balkan and emerging European economies for 1997-2009 period, by dynamic panel data models. They found that macroeconomic stabilization and reforms are significant determinants of GDP growth, but that foreign direct investments and economic integrations are key determinants. Fetahi-Vehapi, Sadiku, and Petkovski (2015) analyzed determinants of economic growth in 10 South-Eastern European countries during the period 1996-2012 by fixed effects panel regression estimation method. They found that trade openness, GDP per capita, human capital development, gross fixed capital formation, and foreign direct investment positively impact economic growth, while population negatively impacts it. Arsić, Mladenović, and Nojković (2021) examined the response of economic growth to the public debt uncertainty in 10 emerging European economies between 2000 and 2015. They found the negative effect that public debt uncertainty has on GDP growth in emerging European economies, and that impact has been amplified over the crisis episode that started in 2008.

### 3. DATA

This research is dealing with the determinants of GDP growth in sample of countries consists of the following ones: Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Serbia, Slovakia, Slovenia, Turkey, and Ukraine. The time span ranges from 2009 to 2018.

Data sources are IMF's World Economic Outlook database and World Bank's WDI database. Sources and description of the data can be found in the table 1.

**Table 1:** Data description

Variables	Description	Source
<i>Y</i>	GDP growth	World Bank, WDI database
<i>CPI</i>	CPI measured inflation at the end of the period	IMF, World Economic Outlook database
<i>DEBT</i>	Public debt in % of GDP.	World Bank, WDI database
<i>GDPPC</i>	Logarithm of GDP per capita	World Bank, WDI database
<i>U</i>	Unemployment, total (% of total labor force)	World Bank, WDI database
<i>S.BALANCE</i>	Structural budget balance in percentage of potential GDP.	IMF, World Economic Outlook database
<i>CA</i>	Current account balance in % of GDP.	IMF, World Economic Outlook database
<i>EUROZONE</i>	Takes value 1 if the country is member of Eurozone, and 0 otherwise.	

Dependent variable in our analysis is GDP growth (*Y*). The path of that variable is estimated dependent on its lagged value and other covariates.

Inflation rate (*CPI*) is one natural candidate to be included as an explanatory variable. We think that it can be significant driver of GDP growth tendencies. The expected sign of the influence is ambiguous. However, because of yearly data, and panel of observed countries, we have good reasons to believe in dominance of negative consequences of inflation on economic growth.

Fiscal policy empirical studies are very often occupied with influence of public debt (*DEBT*) on GDP growth. The key question in this kind of research is the sensitivity of interest rate in conditions of debt rising and transmission mechanism channels on GDP growth. Dominant view in the literature is negative influence from the debt rising on the economic growth. The variable is extracted from the WDI database.

Different development of the observed countries can be of importance in the analysis. Therefore, we include logarithm of GDP per capita (*IGDPPC*), as a variable that can potentially explain economic growth patterns. The expected sign of the influence is ambiguous. One strand of literature argues that in developing countries



there is more maneuver space for growth, while the second strand argues contrary, relating that to the level of development which can be stimulating to growth. Having in mind our group of the observed countries, we are more willing to expect positive sign of the respective coefficient.

Beside these variables that play significant role in explanation the GDP growth tendencies, we include other covariates as a control variables in different specifications. They are presented in the following paragraphs.

Unemployment rate ( $U$ ) can potentially explain GDP growth differences. This is also one of the most interesting topics in macroeconomics, due to the popularity of Okun's curve. However, nowadays Okun's curve has a different structural form than the curve developed in early stages. This is why we classified this variable in control one. Economic theory suggests different slopes of this kind of curve, and therefore there is absence of agreement on the problem. However, we will follow traditional stance and assume negative impact of unemployment on GDP growth.

One additional variable that we want to include in our model specifications is structural budget balance ( $S.BALANCE$ ), because it can be useful in measuring the effects of fiscal (ir)responsibility on GDP growth rate. We used IMF's World Economic Outlook database to extract this variable. Again, we assume that fiscal responsibility supports GDP growth, and therefore expect positive sign of the estimated coefficient.

Current account balance ( $CA$ ) can be potential covariate. There are direct and indirect channel through which current account balance can influence GDP growth. Direct channel is related to main GDP component structure. Increase in current account balance positively affects GDP growth. Contrary, due to the effect of aggregate demand, the effect is of opposite direction. The overall effect depends on the relative strength of these two channels.

Finally, we wanted to control for Eurozone ( $EUROZONE$ ) membership, as possible determinant of growth. In our sample, Estonia, Latvia, Lithuania, Slovakia, and Slovenia are members of the Eurozone. All these countries are relatively small in comparison to other countries in the panel. In addition, few countries in the panel that have their own currency are one of the most opened countries measured by sum of exports and imports relative to GDP. All these factors can determine the expected sign of the coefficient.

The following table 2 gives summary descriptive statistics regarding used variables.

**Table 2:** Summary descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
$Y$	1.62	4.20	-14.81	11.11
$CPI$	3.14	5.09	-1.95	43.31
$DEBT$	45.16	19.59	6.11	83.99
$GDPPC$	1.40	0.06	1.20	1.49
$U$	11.05	5.63	2.24	28.01
$S.BALANCE$	-2.10	2.16	-8.54	2.11
$CA$	-1.55	3.49	-10.80	7.71

Source: Author's calculation

#### 4. METHODOLOGY

The model specification can be noted in the following way:

$$Y_{it} = \rho Y_{it-1} + \alpha CPI_{it} + \beta DEBT_{it} + \gamma GDPPC_{it} + \delta X'_{it} + \mu_i + \varepsilon_{it} \quad (1)$$

Where  $Y$  is GDP growth,  $CPI$  is inflation rate,  $DEBT$  is public debt,  $GDPPC$  is logarithm of GDP per capita,  $X$  is a vector of control variables such as  $U$  (unemployment),  $S.BALANCE$  (structural budget balance),  $CA$  (current account balance),  $EUROZONE$  (dummy variable that takes value 1 if the country is member of Eurozone, and 0 otherwise),  $\mu_i$  captures unobserved country-specific effects and  $\varepsilon_{it}$  is the error term.

Although modeling in panel data framework can be done by employing various techniques such as pooled OLS method, fixed and random effects estimation, instrumental variables methods, dynamic panel data estimation, we opt for dynamic panel data estimation. This technique is especially relevant to our case, since there is high possibility of latency in dependent variable path. If the dynamic panel data model is estimated by some of the standard panel data techniques, then the results would be biased and estimations would be inconsistent. Many empirical studies perform Fixed Effects (FE), but it can cause bias especially if the number of periods is low (Nickell, 1981; Kiviet, 1995). Therefore, we opted for standard dynamic panel data modeling. However, there are different estimators that can be used, like Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998) method. We used one-step procedure by Blundell and Bond (System GMM method). This method accounts for endogeneity of lagged dependent variable. It reduces finite sample bias and is dominant according to that criterium in comparison with other dynamic panel data estimators (Baltagi, 2008). It is asymptotically more efficient. Soto (2009) finds that this method gives the best results in

the case of small N, as it is the case in our study, and that its application on small samples does not have significant repercussions on the properties of the estimator.

Regarding the underline methodology of the method, System GMM encompasses a system of equations where lagged first differences of dependent variable are instruments for the equations in level, whereas lagged levels of dependent variable are used as instruments for equations in first differences (Blundell and Bond, 1998). Instruments for other endogenous variables can be used. Roodman (2009) investigated a potential problem with too many instruments. It can be solved by limiting number of lags and collapsing the instrument matrix. We use the both methods in order to limit the number of instruments to desirable level.

Regarding standard statistical tests in this kind of research, first we use Hansen test for over-identifying restrictions, which investigates the instruments validity, i.e. it confirms the exogeneity of instruments. Second test (Arellano - Bond test) is used for checking the presence of second order serial correlation of differenced residuals. In addition, Blundell and Bond estimator assumes the stationarity of all variables. This property is validated by usage of panel unit root tests (Levin, Lin and Chu (LLC), Im, Pesaran, Shin (IPS) and ADF and PP Fisher - type tests). These are first generation tests which have advantages in small samples (Breitung, 2000; Im *et al.*, 2003). The results of the unit root testing are presented in the following table 3.

**Table 3:** Unit root tests results

	LLC		IPS		ADF Fisher		PP Fisher	
	Statistic	p value	Statistic	p value	Statistic	p value	Statistic	p value
<i>Y</i>	-9.1076	0.0000	-6.1162	0.0000	25.1378	0.0000	46.4950	0.0000
<i>CPI</i>	-6.0702	0.0000	-1.6645	0.0480	-0.2002	0.5793	-0.2002	0.5793
<i>DEBT</i>	-5.3560	0.0000	-3.0700	0.0011	24.4775	0.0000	7.9264	0.0000
<i>GDPPC</i>	-1.6128	0.0534	-9.4286	0.0000	28.6755	0.0000	-2.2293	0.9871
<i>U</i>	-1.0838	0.1392	0.8218	0.7944	20.7200	0.0000	-2.6754	0.9963
<i>S.BALANCE</i>	-5.1422	0.0000	-1.4017	0.0805	26.9316	0.0000	4.6023	0.0000
<i>CA</i>	-5.3619	0.0000	-1.9213	0.0273	18.5442	0.0000	4.7839	0.0000

**Note:** Lags are chosen by AIC.

**Source:** Author's calculation

Although some of the tests find non-stationarity of some variables, we proceed to estimation due to the results of other tests.

## 5. RESULTS

The first specification starts with the simple framework consisting of only few most important determinants of GDP growth, while other two specifications include the additional variables that are potential determinants of GDP growth. The following table 4 represents the results for these three specifications of the model.

**Table 4:** Estimated results

Dependent variable: <i>Y</i>	Specification 1	Specification 2	Specification 3
<i>Y (-1)</i>	0.1349* (0.0772)	0.0995 (0.0682)	0.2106*** (0.0840)
<i>CPI</i>	-0.1471* (0.0895)	-0.1902** (0.0972)	-0.1488* (0.0878)
<i>DEBT</i>	-0.0744*** (0.0299)	-0.0511*** (0.0200)	-0.1107*** (0.0250)
<i>GDPPC</i>	6.0527*** (1.1578)	4.8554*** (0.9725)	6.6276*** (1.2450)
<i>U</i>	-0.1828* (0.1087)		
<i>S.BALANCE</i>		0.6711*** (0.2572)	0.2256* (0.1268)
<i>CA</i>		-0.4139*** (0.1479)	
<i>EUROZONE</i>			-6.0079*** (1.6443)

<i>EUROZONE*DEBT</i>			0.0941*** (0.0241)
Arrelano-Bond AR(2) test	-0.00 [0.9970]	-0.49 [0.6250]	-0.47 [0.6400]
Hansen test	9.09 [0.1050]	6.97 [0.2230]	8.76 [0.1190]
No. of observations	135	135	135
No. of countries	15	15	15

**Notes:** \*\*\*, \*\*, \* Denote statistical significance at the 1%, 5%, 10% level, respectively. In parentheses are robust standard errors. System GMM with robust standard errors is applied. Instruments used for level equation are lagged first differences of *Y*, *CPI*, *DEBT*, *GDPPC*, *U*, *CA* (potentially endogenous variables). Instruments used for the first-differenced equations are lagged levels (one period) of dependent variable, and of the potentially endogenous variables. The exogenous covariates (*S\_BALANCE*, *EUROZONE* and interaction term) are instrumented by themselves in the level equations. In square brackets for Arrelano-Bond and Hansen tests are p values.

**Source:** Author's calculation

The estimated specifications perform good statistical properties and can be used for further statistical inference. Generally speaking, the presented results are expected.

The GDP growth variable has the latency in its dynamics. One of the important conclusions can be derived from coefficient of *CPI*, which has negative sign and it is in accordance with many studies (Barro, 1999; Bayraktar, 2006; Prochniak, 2011). That means that inflation has harmful influence on GDP growth. *DEBT* also negatively affects GDP growth, which is expected based on the empirical papers in this area (Barro, 1999; Bleaney, Gemmell, and Kneller, 2001; Prochniak, 2011). Initial level of development proxied by GDP per capita according to our results has positive effect on GDP growth, which is consistent with the findings of Fetahi-Vehapi, Sadiku, and Petkovski (2015). The possible explanation can be found in the characteristics of the observed countries, where all countries in our sample that have higher level of GDP per capita are still converging to more developed countries of Western Europe, and they have greater maneuver space to boost its GDP than already developed economies. Contrary, the rest of countries in our sample, among economic one, have many political problems and institutional issues that have to be fixed. According to theoretical foundation, unemployment rate is negatively correlated to GDP growth.

In addition to these variables, we wanted to control for other ones, that can potentially be significant determinants of GDP growth. The first one is structural budget balance. As expected, fiscal responsibility in the discretionary fiscal policy sphere contributes significantly to the GDP growth. It is in accordance with the findings on public debt influence on growth, and therefore we can be confident that our results are robust in that domain. Current account balance significantly negatively affects GDP growth. Finally, we wanted to see do membership in the Eurozone affects GDP growth, and whether there is significance in interaction between this dummy variable and other covariates. Our findings indicate that countries that are Eurozone members have smaller growth rate, possible because of monetary policy that is unified on level of Eurozone, and which is not always suitable to periphery countries. In addition, the effect of *DEBT* on GDP growth is smaller if the country is member of Eurozone, which can be explained by many above-state agreements in the fiscal area which is found desirable by many investors.

## 6. CONCLUSIONS

This research focuses on the determinants of GDP growth for 2009 – 2018 period in sample of countries that consists of the following countries: Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Serbia, Slovakia, Slovenia, Turkey, and Ukraine. We opted for standard dynamic panel data modeling, specifically System GMM method, which encompasses a system of equations where lagged first differences of dependent variable are instruments for the equations in level, whereas lagged levels of dependent variable are used as instruments for equations in first differences.

We argue that GDP growth in these countries is under influence of many factors from macroeconomic policy side. The GDP growth variable shows the latency in its dynamics. Inflation has negative influence on GDP growth, as well as public debt. Initial level of development proxied by GDP per capita has positive effect on GDP growth. Structural budget balance positively affects economic growth, which is in accordance with the theory that fiscal responsibility contributes to GDP growth. In addition to these variables we control for many other, in order to make our estimations robust.

Based on the findings, we can conclude that policymakers have to be aware of the influence that their policy actions can have on GDP growth. Responsible policies are recognized as pro-growth oriented and due

to the fact that in these countries institutional framework is not developed as it is in the developed countries, the future GDP growth tendencies will be under influence of this structural features of the observed economies.

## Acknowledgement

The paper is the result of research funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

## REFERENCES

- [1] Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277.
- [2] Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), 29-51.
- [3] Arsić, M., Mladenović, Z., & Nojković, A. (2021). Debt uncertainty and economic growth in emerging European economies: Some empirical evidence. *Emerging Markets Finance and Trade*, 57(12), 3565-3585.
- [4] Baltagi, B. (2008). *Econometric Analysis of Panel Data*. John Wiley & Sons.
- [5] Barro, R.J. (1999). Determinants of Economic Growth: Implications of the Global Evidence for Chile. *Cuadernos de Economía*, 36(107), 443-478.
- [6] Bayraktar, B. (2006). Investigation on sources of growth for Turkey. *Canadian Journal of Development Studies / Revue canadienne d'études du développement*, 27(1), 25-38.
- [7] Bleaney, M., Gemmell, N., & Kneller, R. (2001). Testing the endogenous growth model: Public expenditure, taxation, and growth over the long run. *Canadian Journal of Economics/Revue canadienne d'économique*, 34(1), 36-57.
- [8] Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143.
- [9] Breitung, J. (2000). The local power of some unit root tests for panel data. In Baltagi, B.H., Fomby, T.B., & Hill, C.R. (eds): *Nonstationary Panels, Panel Cointegration and Dynamic Panels, Advances in Econometrics*, JAI, 161-177.
- [10] Checherita-Westphal, C., & Rother, P. (2012). The impact of high government debt on economic growth and its channels: An empirical investigation for the euro area. *European Economic Review*, 56(7), 1392-1405.
- [11] Fetahi-Vehapi, M., Sadiku, L., & Petkovski, M. (2015). Empirical analysis of the effects of trade openness on economic growth: An evidence for south East European countries. *Procedia Economics and Finance*, 19, 17-26.
- [12] Im, K. S., Pesaran, M., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1), 53-74.
- [13] IMF. (2021). World Economic Outlook database. <https://www.imf.org/en/Publications/SPROLLS/world-economic-outlook-databases#sort=%40imfdate%20descending>
- [14] Josifidis, K., Dragutinović-Mitrović, R., & Ivančev, O. (2012). Heterogeneity of growth in the west Balkans and emerging Europe: A dynamic panel data model approach. *Panoeconomicus*, 59(2), 157-183.
- [15] Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics*, 68(1), 53-78.
- [16] Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6), 1417.
- [17] Prochniak, M. (2011). Determinants of economic growth in central and Eastern Europe: The global crisis perspective. *Post-Communist Economies*, 23(4), 449-468.
- [18] Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), 135-158.
- [19] Soto, M. (2009). *System GMM Estimation with A Small Sample*. UFAE and IAE Working Paper 78009.
- [20] World Bank, (2021). WDI database. <https://databank.worldbank.org/source/world-development-indicators>

## OCENA PERZISTENTNOSTI INDEKSA NEIZVESNOSTI EKONOMSKE POLITIKE: NEKI EMPIRIJSKI REZULTATI

### ESTIMATING PERSISTENCE OF ECONOMIC POLICY UNCERTAINTY INDEX: SOME EMPIRICAL RESULTS

ZORICA MLADENović<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, zorica.mladenovic@ekof.bg.ac.rs

**Rezime:** Rad sadrži rezultate ocenjivanja nivoa perzistentnosti vremenskih serija indeksa neizvesnosti ekonomske politike Evrope, SAD-a i Kine u periodu: januar 2001 – april 2022. Analiza se zasniva na oceni parametra razlomljene integrisanosti ARFIMA modela. Modeliranje podrazumeva vremenski promenljivu uslovnu varijansu slučajne greške modela, koja je ocenjena prema GARCH specifikaciji uz mogućnost prisustva duge memorije u volatilnosti. Rezultati sugerišu relativno visoku perzistentnost sva tri indeksa neizvesnosti, kao i ekstremnu perzistentnost volatilnosti u Kini. Identifikovana svojstva neizvesnosti mogu negativno uticati na stabilnost makroekonomskih i finansijskih performansi.

**Cljučne reči:** ARFIMA modeli, duga memorija, GARCH modeli, neizvesnost, perzistentnost

**Abstract:** The paper contains results of estimating level of persistence of economic policy uncertainty index time series for Europe, U.S, and China based on the sample: January 2001 – April 2022. Analysis relies on the estimation of the fractional integration parameter of ARFIMA model. Econometric approach assumes time-varying conditional variance of an error term captured by GARCH specification, which allows for the long memory in volatility. Results suggest persistence of relatively high magnitude for all three indexes, and of extreme level in volatility for China. Identified features of uncertainty may have negative impact on macroeconomic and financial performance.

**Keywords:** ARFIMA models, GARCH models, long memory, persistence, uncertainty

## 1. UVOD

Neizvesnost ekonomske politike označava nemogućnost ekonomskih subjekata da predvide, kako buduće ekonomske politike, tako i efekte već donetih ekonomskih mera [18]. Uticaj ekonomske neizvesnosti na makroekonomsku stabilnost predstavlja jednu od najzastupljenijih tema naučnih i stručnih ekonomskih rasprava poslednjih godina.

Sama veličina neizvesnosti je izvedena kategorija. Njenoj kvantifikaciji posvećen je veći broj različitih metoda. Kao alternativa pristupima koji na posredan način mere ekonomsku neizvesnost, razvijen je koncept indeksa neizvesnosti ekonomske politike (engl. skraćenica EPU index – economic policy uncertainty index) [5], [7]. EPU indeks određuje se na mesečnom nivou dominantno prema učestalosti pojavljivanja u štampanim medijima ključnih reči koje asociraju na ekonomsku i političku neizvesnost. Time se aproksimira sumarna neizvesnost ekonomske politike koja je dostupna u javnim izvorima [5].

EPU indeks obrazovan je prvo za SAD. Već početne analize pokazale su njegovu visoku korelisanost sa rastom volatilnosti prinosa finansijskih instrumenata i padom nivoa investicija i zaposlenosti u pojedinim ekonomskim sektorima [5],[8]. Posmatrano na makro nivou, neanticipirani pozitivni slučajni šokovi u neizvesnosti utiču na pad investicija i proizvodnje, kako u SAD-u, tako i drugim ekonomijama. Indeks je sada dostupan za veći broj zemalja ([www.policyuncertainty.com](http://www.policyuncertainty.com)), dok su za pojedine ekonomije definisane i uže mere neizvesnosti (monetarna, fiskalna, poreska, itd.).

Koncept EPU indeksa intenzivno se koristi u empirijskom modeliranju. Može se izdvojiti nekoliko pravaca analize. Jedna grupa istraživanja odnosi se na modeliranje dinamičke strukture ukupne i pojedinačnih kategorija neizvesnosti na nivou date ekonomije. Druga grupa istraživanja ispituje prirodu uticaja neizvesnosti

na ključne ekonomske veličine. Treća grupa analiza meri efekat preliivanja uticaja neizvesnosti između različitih zemalja i regiona.

Jednodimenziono modeliranje EPU indeksa prirodno otvara pitanje stepena njegove perzistentnosti. Visoka perzistentnost sugerise prisustvo duge memorije, odnosno trajno dejstvo neočekivanih slučajnih šokova.

Formalno posmatrano, perzistentnost se meri prema zbirnoj vrednosti parametara odgovarajuće reprezentacije linearnog procesa [15]. Ta beskonačna vrednost može se aproksimirati na osnovu različitih pristupa [15], [16]. Jedan od njih zasniva se na oceni parametra razlomljene (frakcijalne) integrisanosti u okviru autoregresionih modela pokretnih proseka za razlomljeno (frakcijalno) integrisane vremenske serije (engl. skraćena ARFIMA) [12], [13], [14]. Ideja o prisustvu duge memorije zastupljena je i kod modeliranja vremenski promenljive heteroskedastičnosti uopštenog autoregresionog tipa (engl. skraćena GARCH). Razvijena je posebna grupa modela kojom se dozvoljava razlomljena integrisanost volatilnosti – FIGARCH modeli [2], [3], [9].

Cilj ovog rada je ocena perzistentnosti indeksa neizvesnosti ekonomske politike za tri relevantne ekonomske oblasti na globalnom nivou – Evropu, SAD i Kinu, prema konceptu ARFIMA modela. Pri tome, za modeliranje vremenski promenljive uslovne varijanse koristi se GARCH specifikacija. Ispituje se prisustvo duge memorije i u volatilnosti.

Analizi duge memorije ekonomske i užih kategorija neizvesnosti posvećeno je nekoliko istraživanja [1], [11], [18], [19]. Ona se odnose na podatke do kraja 2018. ili sredine 2019. godine. Za najveći broj ekonomija ustanovljen je relativno visok nivo perzistentnosti, koji ipak ne sugerise nestacionarnost jediničnog korena.

Doprinos ovog rada sastoji se u sledećem. Modelira se dinamika vremenske serije tako što se istovremeno ocenjuje perzistentnost nivoa i volatilnosti. Period analize (do aprila 2022) obuhvata novije egzogene efekte, koji su značajno uticali na rast ekonomske neizvesnosti (pandemija COVID-19 i početak rata u Ukrajini). Naše modeliranje uključuje i evropski indeks neizvesnosti, koji u dostupnim radovima nije bio predmet razmatranja.

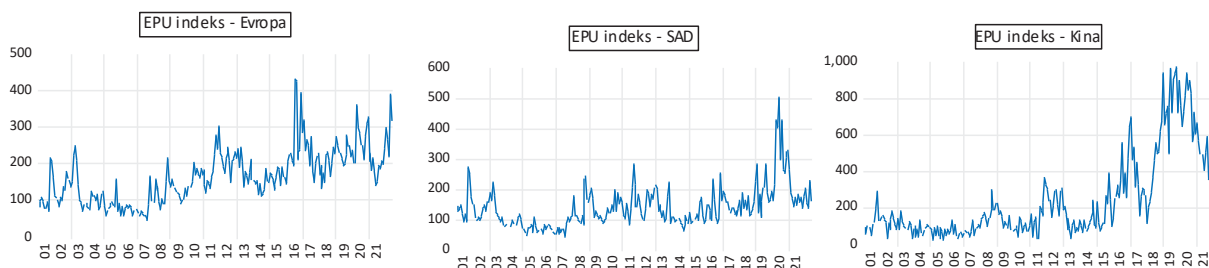
Struktura rada je sledeća. U delu 2. dati su rezultati preliminarne statističke analize. Rezultati ekonometrijskog modeliranja izloženi su u delu 3. Zaključni komentari obuhvaćeni su delom 4.

## 2. PRELIMINARNI EMPIRIJSKI REZULTATI

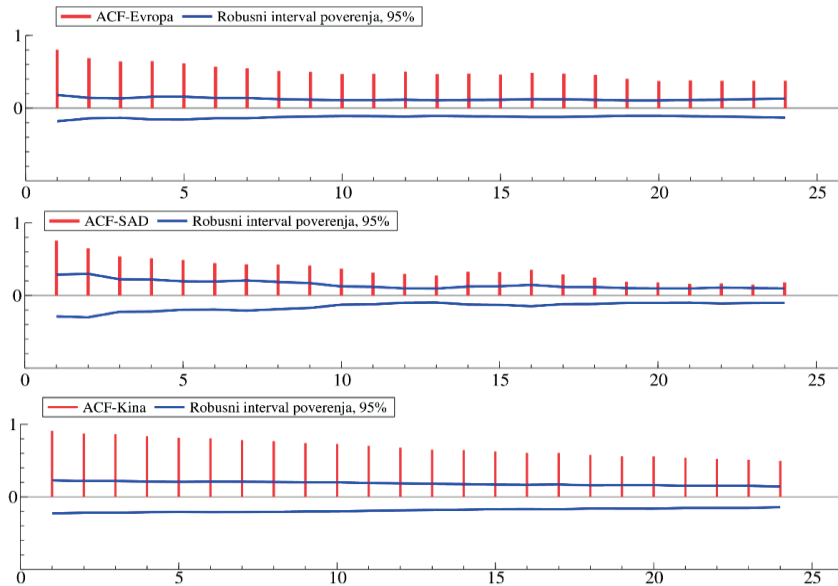
Vremenske serije EPU indeksa neizvesnosti za Evropu, SAD i Kinu razmatrane su na mesečnom nivou u periodu: januar 2001 – april 2022. (256 podataka). Podaci su preuzeti sa sajta [www.policyuncertainty.com](http://www.policyuncertainty.com) [4], [5]. Karakterise ih rastući trend (Slika 1) sa izraženim kratkoročnim varijacijama i većim brojem jednokratnih nestandardnih opservacija. Primenom konvencionalnih testova jediničnog korena (ADF, KPSS i ERS) trend je identifikovan kao deterministički za EPU indekse Evrope i SAD-a. Rezultati ukazuju na prisustvo jediničnog korena EPU indeksa Kine. Ukoliko se prilikom testiranja omogući provera postojanja trajnog strukturnog loma endogenog tipa, tada se dobija nalaz da je EPU indeks Kine vremenska serija koja stacionarno oscilira oko determinističkog trenda sa strukturnim lomom u maju 2018. godine (kao posledica američko-kineskog trgovinskog rata). Otuda se u daljem radu razmatraju nivoi sva tri indeksa.

Jedan od uobičajenih instrumenata analize vremenskih serija jeste ocena obične autokorelacione funkcije (engl. ACF). Za svaku od tri posmatrane vremenske serije ocene ove funkcije sporo opadaju sa rastom doznji, što je znak osetljivosti na uticaj neočekivanih slučajnih šokova (Slika 2). Ocene na prvoj doznji su redom 0.80, 0.75 i 0.91 za Evropu, SAD i Kinu. Identičan zaključak sugerisu vrednosti R/S testa Hurst-Mandelbrota i Loa [16] u Tabeli 1. Ova tabela obuhvata i rezultate testiranja koji ukazuju na postojanje vremenski promenljive uslovne varijanse.

Nivo duge memorije preliminarno je ocenjen na osnovu GPH [10] i Robinsonovog [17] algoritma (Tabela 2). Dobijene ocene upućuju na relativno visoku perzistentnost sve tri vremenske serije, koja je posebno izražena za Kinu.



Slika 1: EPU indeksi neizvesnosti



**Slika 2:** Ocene obične autokorelacione funkcije za prve 24 docnje

**Tabela 1:** Osnovni pokazatelji postojanja duge memorije i vremenski promenljive uslovne varijanse

EPU indeks	Q(24)	Q <sup>2</sup> (24)	ARCH(10)	Hurst-Mandelbrot R/S	Lo R/S	
					10	24
Evropa	113.5(0.0)	937(0.0)	31.5(0.0)	5.55	2.03	1.46
SAD	59.3(0.0)	575.9 (0.0)	33.6(0.0)	3.69	1.44	1.09
Kina	76.7(0.0)	2394(0.0)	106.2(0.0)	5.70	1.95	1.30

Napomene: U zagradi je data p-vrednost. Q(24) i Q<sup>2</sup>(24) su Boks-Ljungove statistike za proveru zbirne autokorelacije reda 24 respektivno u nivou i kvadratu vremenske serije. ARCH(10) je test autoregresione strukture reda 10 kvadrata vremenske serije. Sa R/S označen je test raspona normiran standardnom devijacijom (engl. range over standard deviation), koji je dat u dve različite varijante. Interval poverenja sa verovatnoćom 95% R/S testa je (0.81,1.86). Prilikom određivanja vrednosti R/S statistike Loa korišćena je ocena dugoročne varijanse bazirana na autokorelaciji reda 10 i 24.

**Tabela 2:** Ocena nivoa perzistentnosti polaznih podataka

EPU indeks	T <sup>0.5</sup> - 1=15		T <sup>0.5</sup> =16		T <sup>0.5</sup> + 1=17	
	GPH	Robinsonov	GPH	Robinsonov	GPH	Robinsonov
Evropa	0.74	0.64	0.70	0.65	0.62	0.58
SAD	0.68	0.62	0.64	0.62	0.57	0.53
Kina	0.98	0.96	0.98	0.96	0.94	0.94
<i>Standardna greška ocene</i>	<i>0.22</i>	<i>0.13</i>	<i>0.21</i>	<i>0.13</i>	<i>0.20</i>	<i>0.12</i>

Napomene: T je dužina vremenske serije. Kao u [1], ocene su dobijene za različit broj spektralnih ordinata.

### 3. REZULTATI EKONOMETRIJSKOG MODELIRANJA

Osnovni okvir ekonometrijske analize jeste ARFIMA-FIGARCH model koji se sastoji od jednačine srednje vrednosti i jednačine volatilnosti.

Jednačina srednje vrednosti opisuje se ARFIMA( $p,d,q$ ) modelom sledećeg oblika [13], [14]:

$$(1 - \phi_1 L - \dots - \phi_p L^p)(1 - L)^d X_t = (1 - \theta_1 L - \dots - \theta_q L^q) e_t \quad (1)$$

gde je:  $d$  parametar razlomljene integrisanosti,  $L$  operator docnje prvog reda i  $e_t$  slučajna greška modela. Sa  $\phi_1, \dots, \phi_p$  označeni su autoregresioni parametri, a sa  $\theta_1, \dots, \theta_q$  parametri pokretnih proseka. Ključni predmet naše analize je parametar razlomljene integrisanosti  $d$ .

Za  $0 < d < 0.5$ ,  $X_t$  je stacionarna vremenska serija sa dugom memorijom, čija obična autokorelaciona funkcija uzima samo pozitivne vrednosti koje opadaju po hiperboličnoj putanji ka nuli sa rastom docnji [2], [12], [16]. Ukoliko je:  $0.5 \leq d < 1$ , tada varijansa  $X_t$  više nije stabilna, tako da je vremenska serija formalno nestacionarna sa dugom memorijom [16]. Međutim, dugoročni uticaj slučajnih šokova nije intenzivan u meri u kojoj je to slučaj kod prisustva jediničnog korena ( $d = 1$ ) [16].

Slučajna greška modela (1) data je sa  $e_t = \sigma_t u_t$ , gde je  $u_t$  niz nezavisnih i jednako raspodeljenih slučajnih promenljivih nulte srednje vrednosti i jedinične varijanse. Sa  $\sigma_t^2$  označena je uslovna varijansa.

Ukoliko pretpostavimo prisustvo duge memorije za  $\sigma_t^2$ , kao i relevantnost dinamike samo na prvoj docnji, tada za modeliranje volatilnosti koristimo FIGARCH(1,  $d^*$ , 1) specifikaciju sledećeg tipa [9]:

$$(1 - \alpha_1 L)(1 - L)^{d^*} (e_t^2 - \sigma_t^2) = (1 - \beta_1 L) (\sigma_t^2 - \sigma_t^2) \quad (2)$$

U jednačini (2)  $d^*$  je parametar razlomljene integrisanosti volatilnosti,  $\sigma^2$  bezuslovna varijansa za  $e_t$ , dok su  $\alpha_1$  i  $\beta_1$  parametri. Uslovna varijansa je pozitivna ako važi:  $\sigma^2 > 0, 0 \leq \alpha_1 \leq \beta_1 \leq d^* \leq 1$  [9]. Za  $d^* = 0$ , FIGARCH model se svodi na GARCH. Izraz (2) predstavlja modifikaciju FIGARCH modela koji je prvo definisan u [2] i [3].

Parametri modela (1) i (2) ocenjeni su za svaku od tri serije EPU indeksa primenom BFGS algoritma numeričke optimizacije. Pretpostavljeno je da slučajna greška modela sledi  $t$  ili uopštenu eksponencijalnu raspodelu (engl. skraćenica: GED), da bi se bolje obuhvatilo svojstvo teških repova empirijske raspodele. Korišćen je softver Oxmetrics 8.1. Rezultati su dati u Tabeli 3.

EPU indeksi Evrope i SAD modelirani su na osnovu ARFIMA(0,  $d$ , 0) – GARCH(1,1) specifikacije sa ocenama parametra razlomljene integrisanosti redom 0.70 i 0.63. Ove vrednosti su slične prethodno dobijenim u preliminarnoj analizi. Rezultat sugerise nestabilnu varijansu nivoa indeksa neizvesnosti tokom vremena i njegovu relativno visoku osetljivost na uticaj neočekivanih slučajnih šokova. Trajnost efekata slučajnih šokova na volatilnost indeksa neizvesnosti nije uočena, jer parametar duge memorije nije statistički značajan u GARCH modelima.

Dobijena ocena parametra  $d$  za SAD je nešto veća od ocena datih u radovima [11] i [18]. Budući da naša analiza uključuje podatke zaključno sa aprilom 2022, može se zaključiti da su novi faktori rasta neizvesnosti (pandemija COVID-19 i početak rata u Ukrajini) uticali i na povećanje nivoa perzistentnosti.

Za EPU indeks Kine odabran je model ARFIMA(0,0.62,3) – FIGARCH(1,0.93,1). Ocena parametra duge memorije ARFIMA specifikacije, 0.62, slična je kao kod prethodna dva modela. Ona je znatno manja nego u preliminarnoj analizi, budući da je modeliranjem eksplicitno uključena informacija o postojanju trajnog strukturnog loma.

Parametar  $d$  je statistički značajno manji od 1 za sva tri razmatrana indeksa neizvesnosti, tako da ne postoji ekstremna perzistentnost njihovog nivoa.

U ocenjenom FIGARCH modelu EPU indeksa Kine identifikovana je jaka perzistentnost vremenski promenljive uslovne varijanse. Na osnovu dobijenih ocena zaključuje se da parametar perzistentnosti  $d^*$  nije statistički značajno različit od 1. Time se sugerise postojanost dugoročnog uticaja slučajnih šokova na dinamiku volatilnosti.

Prirodno se nameće pitanje međusobne usklađenosti, kako pojedinačno razmatranih EPU indeksa, tako i tipa njihove uzročnosti. Rezultati modeliranja vremenskih serija količnika EPU indeksa Kine i SAD-a, odnosno Evrope i SAD-a, ukazuju na takav stepen slaganja kojim se neutrališe jedan deo perzistentnosti. Preciznija analiza bi zahtevala primenu koncepta razlomljene kointegracije [12], [19]. U cilju sagledavanja pravca uzročnosti, od koristi su, ne samo standardni testovi, već i pristupi kojima se detektuju promene karaktera uzročnosti tokom posmatranog perioda.



**Tabela 3:** Ocene ARFIMA – FIGARCH modela za EPU indekse

Jednačina srednje vrednosti			
Ocena parametra	Evropa	SAD	Kina
$d$	0.70 (14.65)	0.63 (7.76)	0.62 (4.23)
$\theta_1$	-		-0.15 (-0.98)
$\theta_2$	-	-	-0.09 (-1.10)
$\theta_3$	-	-	0.13 (2.05)
Jednačina volatilnosti			
Ocena parametra	Evropa	SAD	Kina
$d^*$	-	-	0.93 (14.57)
$ARCH(1)$	0.29 (2.17)	0.41 (2.03)	0.22 (3.08)
$GARCH(1)$	0.11 (0.57)	0.29 (2.07)	0.86 (11.74)
Broj stepeni slobode $t$ -raspodele	6.74 (2.02)	-	-
Broj stepeni slobode GED raspodele	-	1.27 (7.28)	1.78 (7.83)
SC	9.959	10.072	11.456
Q(24)	21.95(0.58)	20.17(0.69)	29.12(0.11)
Q <sup>2</sup> (24)	19.16(0.64)	9.58(0.95)	25.21(0.29)
ARCH(10)	0.72(0.70)	4.41(0.35)	1.18(0.30)
Engle-Ng zbirni test asimetrije	0.56(0.90)	1.16(0.76)	5.03(0.17)

Napomene: U zagradi ispod ocene parametra dat je t-odnos. Obe ocenjene jednačine za sve tri serije sadrže slobodan član. U jednačine srednje vrednosti EPU indeksa Evrope i SAD-a uključene su impulsne veštačke promenljive sa nenultom vrednošću 1 za sledeće mesece: 2016M11, 2020M3 i 2022M3. Jednačina EPU indeksa Evrope dodatno sadrži impulsne veštačke promenljive za: 2011M11 i 2016M67, a jednačina EPU indeksa SAD-a za: 2001M9 i 2020M45. U jednačini srednje vrednosti EPU indeksa Kine prisutne su impulsne veštačke promenljive za 2018M12 i 2022M3, kao i stepenik veštačka promenljiva koja uzima nenultu vrednost 1 od 2018M5.

#### 4. ZAKLJUČAK

Visoka perzistentnost i nestabilnost fluktuacija ekonomske neizvesnosti mogu imati negativan uticaj na efikasnost monetarne i fiskalne politike, kao i na pojedinačne odluke ekonomskih subjekata. Otuda je ocena perzistentnosti značajna tema ekonometrijske analize, posebno imajući u vidu egzogene faktore koji su poslednjih godina uticali na rast ekonomske neizvesnosti [6].

Izložili smo rezultate ocenjivanja perzistentnosti indeksa neizvesnosti ekonomske politike Evrope, SAD-a i Kine, koji su dobijeni na osnovu ARFIMA-(FI)GARCH specifikacija. Istovremeno modeliranje većeg broja ključnih svojstava vremenske serije povećava preciznost statističkog zaključivanja. Ustanovili smo da indekse neizvesnosti karakteriše nestabilna varijansa i relativno visok stepen osetljivosti na dejstvo slučajnih šokova. Dodatno, za indeks neizvesnosti Kine ocenjena je ekstremno visoka perzistentnost volatilnosti.

Ova analiza odnosi se na period zaključno sa aprilom 2022, dok se postojeći rezultati za SAD i Kinu zasnivaju na podacima do kraja 2018, odnosno sredine 2019. godine. Iako primenjene metodologije nisu direktno uporedive, veće vrednosti naših ocena u odnosu na prethodno dobijene ukazuju na to da je sa rastom ekonomske neizvesnosti došlo do trajnijeg uticaja neočekivanih slučajnih faktora.

**LITERATURA**

- [1] Abakah, E.J.A., Caporale, G.M. & Gil-Alana L.A. (2021). Economic policy uncertainty: Persistence and cross-country linkages. *Research in International Business and Finance*, 58101442.
- [2] Baillie, R.T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73, 5–59.
- [3] Baillie, R.T., Bollerslev, T. & Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 74, 3–30.
- [4] Baker, S.R., Bloom, N., Davis S.J. & Wang, X. (2013). Economic policy uncertainty in China, University of Chicago.
- [5] Baker, S.R., Bloom, N. & Davis. S.J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131, 1593–1636.
- [6] Baker, S. R., Bloom, N., Davis S.J. & Terry S.J. (2020). COVID-induced Economic Uncertainty. National Bureau of Economic Research (NBER), Working Paper 26983.
- [7] Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77, 623–685.
- [8] Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28, 153–176.
- [9] Chung, C.F. (1999), Estimating the fractionally integrated GARCH model, National Taiwan University Working Papers.
- [10] Geweke, J. & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4, 221–238.
- [11] Gil-Alana, L.A. & Payne. J.E. (2020). Measuring the degree of persistence in the U.S. economic policy uncertainty index. *Applied Economics Letters*, 27, 831–835.
- [12] Gil-Alana, L.A. & Hualde, J. (2009). Fractional integration and cointegration: An overview and an empirical application, in Palgrave Handbook of Econometrics: Volume 2, Applied Econometrics (editors: Mills, T.C. & Patterson, K.), Palgrave Macmillan, 434–469.
- [13] Granger, C.W.J. & Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–39.
- [14] Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, 68, 165–176.
- [15] Markellos, R. and Mills, T.C. (2004). The econometric modelling of financial time series, 3<sup>rd</sup> ed., Cambridge University Press.
- [16] Mills, T.C. (2019). Applied Time Series Analysis, Academic Press.
- [17] Robinson, P.M. (1995). Gaussian semiparametric estimation of long run dependence, *Annals of Statistics*, 23, 1630–1661.
- [18] Solarin, S.A. & Gil-Alana, L.A. (2021). The persistence of economic policy uncertainty: Evidence of long range dependence, *Physica A*, 568, 125698.
- [19] Yaya, O.S., Abu, N. & Ogundunmade, T.P. (2021). Economic policy uncertainty in G7 countries: evidence of long-range dependence and cointegration, *Economic Change and Restructuring*, 54, 541 – 556.

## RANJIVOST MLADIH U SRBIJI<sup>1</sup>

GORANA KRSTIĆ<sup>1</sup>, ALEKSANDRA ANIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, [gorana.krstic@ekof.bg.ac.rs](mailto:gorana.krstic@ekof.bg.ac.rs)

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, [aleksandra.anic@ekof.bg.ac.rs](mailto:aleksandra.anic@ekof.bg.ac.rs)

**Rezime:** Korišćenjem mikro podataka Ankete o dohotku i uslovima života u periodu 2015–2018. analizirali smo rizik od siromaštva i socijalne isključenosti mladih u Srbiji. Pored pojedinačnih komponenata, stope rizika od siromaštva, veoma niskog intenziteta rada i izrazite materijalne uskraćenosti, ispitali smo i vezu obrazovnog nivoa roditelja i rizika od siromaštva mladih, kao i mladih koji žive u domaćinstvima niskog intenziteta rada. Mladi čiji su roditelji nižeg obrazovnog nivoa imaju mnogo veće stope siromaštva, i u većem procentu žive u domaćinstvima veoma niskog intenziteta rada. Mladi su u lošijem položaju u odnosu na ukupnu populaciju, pri čemu su više ranjivi mladi uzrasta 15–19 u odnosu na mlade uzrasta 20–24. Dodatno zabrinjava činjenica da čak četvrtina mladih u Srbiji je bila u riziku od siromaštva u najmanje tri od četiri uzastopne godine. Doprinos ovog istraživanja jeste izračunavanje dodatnih indikatora za mlade koji nisu dostupni u bazi podataka Eurostata niti RZS-a, a koji su značajni za praćenje ranjivosti mladih.

**Ključne reči:** Mladi, Ranjivost, SILC, Srbija

**Abstract:** We investigate youth at-risk-of-poverty or social exclusion in Serbia using micro data from the Survey of Income and Living Conditions over the period 2015–2018. At-risk-of-poverty, very low work intensity and severe material deprivation are analysed and also the relationship between two of these components with the education level of parents. Youth are more exposed to poverty risk and a higher percentage of them live in households with very low work intensity if their parents have low education attainment. Youth are in worse condition than the general population, particularly youth aged 15–19 are more vulnerable than youth aged 20–24. What is especially worrying is that one fourth of young people in Serbia are in risk-of-poverty in at least three out of four consecutive years. The contribution of this research is the provision of additional indicators for youth which are not available in neither Eurostat nor SORS database, which are important for monitoring youth vulnerability.

**Keywords:** Youth, vulnerabilities, SILC, Serbia

### 1. UVOD

Siromaštvo mladih nije do sada bilo u fokusu istraživanja za Srbiju, iako su stope siromaštva mladih veće od stopa siromaštva za ostale starosne grupe, kao i od stopa siromaštva za ukupnu populaciju. Istraživači su se uglavnom fokusirali na tranziciju mladih od škole do posla [8], kao i na položaj mladih na tržištu rada [2] budući da su to bitne determinante ranjivosti mladih. S obzirom na to da ne postoje istraživanja koja se bave ranjivosti mladih, u ovom radu ćemo detaljnije analizirati ranjive situacije u kojima se nalaze mladi na osnovu indikatora ranjivosti koji se koriste na nivou EU.

Predmet ovog istraživanja jeste analiza pokazatelja ranjivosti mladih u Srbiji korišćenjem podataka Ankete o dohotku i uslovima života za period od 2015–2018. godine. Mladi su lica uzrasta 15–24 godine, pri čemu ćemo se posebno osvrnuti na dve starosne grupe, 15–19 i 20–24. Analiziraćemo stopu rizika od siromaštva ili socijalne isključenosti, kao i njene komponente. To su stopa rizika od siromaštva, stopa izrazite materijalne uskraćenosti, i veoma nizak intenzitet rada. Ove indikatore ćemo posmatrati i prema obrazovanju roditelja. Prikazaćemo i stopu trajnog rizika od siromaštva i procenat mladih koji trajno žive u domaćinstvima veoma niskog intenziteta rada, kako bismo videli da li su mladi dugoročno izloženi siromaštvu.

Stopa rizika od siromaštva predstavlja udeo lica čiji je ekvivalentni raspoloživi dohodak manji od 60% medijane ekvivalentnog nacionalnog raspoloživog dohotka [4]. Stopa izrazite materijalne uskraćenosti izražava nesposobnost da se priušte najmanje četiri od devet stavki koje većina ljudi smatra poželjnim ili čak

<sup>1</sup> Rezultati u ovom radu su dobijeni u okviru opsežnog istraživanja mladih u Srbiji. ETF. (2021). “Youth situation in Serbia. Employment, skills and social inclusion”

neophodnim za odgovarajući životni standard.<sup>2</sup> Indikatori tržišta rada kao što su zaposlenost ili aktivnost odnose se samo na broj aktivnih ili zaposlenih lica, dok intenzitet rada članova domaćinstva ukazuje na količinu rada, odnosno na to koliko su odrasli članovi domaćinstva radili u odnosu na potencijalan broj meseci. Prednost ovog pokazatelja jeste ta što se intenzitet rada ne posmatra na nivou lica, već domaćinstva, pošto blagostanje lica ne zavisi samo od njegovog intenziteta rada već i od intenziteta rada ostalih članova u domaćinstvu [7]. Lica koja žive u domaćinstvu sa veoma niskim intenzitetom rada su sva lica starosti od 0 do 59 godina koja žive u domaćinstvu u kojima su radno sposobni članovi radili manje od 20% ukupnog broja meseci u kojima su mogli raditi u toku referentnog perioda. S obzirom na to da se lica često nalaze istovremeno i u riziku od siromaštva i žive u domaćinstvima veoma niskog intenziteta rada, posmatraćemo i udeo lica koja su ranjiva prema oba kriterijuma.

Stopa trajnog rizika od siromaštva pokazuje udeo lica koja su u riziku od siromaštva u tekućoj i najmanje dve od tri prethodne godine. Visoka stopa trajnog rizika od siromaštva ukazuje na to da su lica zarobljena u siromaštvu, odnosno da siromaštvo nije prolaznog karaktera. Procenat lica koja trajno žive u domaćinstvima veoma niskog intenziteta rada predstavlja udeo lica koja žive u ovim domaćinstvima u tekućoj godini i najmanje dve od prethodne tri godine.

Posle uvodnog dela, u drugom delu rada dajemo kratak prikaz literature o siromaštvu mladih u evropskim zemljama. U trećem delu predstavljeni su indikatori ranjivosti mladih koji su dostupni za Srbiju i druge evropske zemlje, kao i novi indikatori koji bi bili značajni za praćenje promene položaja mladih. U poslednjem četvrtom delu dati su zaključci.

## 2. LITERATURA

Postojeća istraživanja za zemlje Evrope pokazuju da su mladi uglavnom više izloženi siromaštvu od odraslih starijih od 25 godina, kao i u odnosu na ukupnu populaciju.

Aassve, Iacovou i Mencarini [1] su istraživali siromaštvo mladih u zemljama Evrope<sup>3</sup> korišćenjem Evropske panel studije domaćinstva (engl. *European Community Household Panel*) u periodu 1994–2000. godine. Bitne determinante rizika od siromaštva kod mladih su status na tržištu rada i stanovanje, odnosno zajednica u kojoj žive. Siromaštvo je manje kod mladih koji žive sa svojim roditeljima ili koji žive sa partnerom bez dece. Nasuprot tome, siromaštvo je veliko kod mladih koji žive sami ili koji su samohrani roditelji. Naravno, studenti (kao i ostali neaktivni) i nezaposleni mladi su u značajno većem riziku od siromaštva. Stopa rizika od siromaštva opada sa godinama, i niža je u tridesetim nego u dvadesetim. Stope siromaštva u zemljama Evrope su veće u Južnoj Evropi, kao i u zemljama „liberalnog režima“ na primer u Velikoj Britaniji i Irskoj. Razlike u stopama siromaštva postoje i između i unutar zemalja.

Fahmy [6] je ispitivao siromaštvo mladih u zemljama Evrope korišćenjem SILC podataka u 2009. godini, pri čemu je analiza bazirana na nekoliko indikatora: dohodno siromaštvo, materijalna deprivacija i subjektivno siromaštvo. Stope siromaštva mladih su veće nego stope siromaštva odraslih, pri čemu izrazito visoke stope siromaštva mladih su zabeležene u bivšim post-komunističkim zemljama i u tzv. južnim/mediteranskim državama blagostanja.

Siromaštvo mladih je ozbiljan problem sa kojima se suočava većina zemlja. Ono što se razlikuje između zemalja jeste da li je to siromaštvo privremenog karaktera, ili su mladi dugoročno zarobljeni u siromaštvu.

## 3. POKAZATELJI RANJIVOSTI MLADIH

Tabela 1 pokazuje stopu rizika od siromaštva ili socijalne isključenosti, kao i njene komponente za Srbiju u periodu 2015–2018. godine. Stopa rizika od siromaštva ili socijalne isključenosti je u posmatranom periodu veća za mlade nego za ukupnu populaciju, i razlika iznosi 4,5 procentna poena (p.p.) u 2018. godini. Ovaj kompozitni pokazatelj koji obuhvata sve tri dimenzije ranjivosti mladih (veoma nizak intenzitet rada, relativno siromaštvo i materijalna uskraćenost) pokazuje da je 38,8% mladih starosti od 15 do 24 godine u Srbiji iskusilo bar jednu od navedene tri dimenzije ranjivosti u 2018. godini. Vrednost ovog pokazatelja za Srbiju spada među najveće stope u poređenju sa 28 zemalja EU (Grčka je jedina zemlja sa višom stopom kod mladih: 45,9%)<sup>4</sup>.

<sup>2</sup> Tih devet stavki odnose se na nemogućnost domaćinstva da priušti: neočekivani trošak, odmor od nedelju dana van kuće; da izbegne kašnjenje u plaćanju troškova (hipoteku ili ratu za stan, komunalne usluge); obrok sa mesom ili ribom svakog drugog dana; odgovarajuće grejanje; veš mašinu; TV u boji; telefon; automobil.

<sup>3</sup> Zemlje uključene u analizu su Nemačka, Danska, Holandija, Belgija, Francuska, Velika Britanija, Irska, Italija, Grčka, Španija, Portugalija, Austrija i Finska.

<sup>4</sup> Podaci za EU na osnovu SILC za 2017.

Ranjivosti su više izloženi mladi uzrasta 15–19 nego 20–24. godine. Ako posmatramo pojedinačne komponente ovog indikatora, zapažamo slične trendove. Mladi su više izloženi ranjivosti nego ukupno stanovništvo, veća je ranjivost u uzrastu 15–19, nego 20–24, sa izuzetkom stope izrazite materijalne uskraćenosti. Stopa izrazite materijalne uskraćenosti je jednaka za ukupno stanovništvo i mlade u 2017. i 2018. godini. Skoro svaka četvrta mlada osoba živi u domaćinstvu veoma niskog intenziteta rada. U ukupnoj populaciji, svaka 8 osoba živi u domaćinstvu veoma niskog intenziteta rada. Vidimo da se učešće mladih koja žive u domaćinstvima veoma niskog intenziteta rada za uzrast 15–19 i 20–24 uglavnom ne razlikuje u posmatranom periodu. Primetan je trend smanjenja siromaštva i socijalne isključenosti u posmatranom periodu.

**Tabela 1:** Stopa rizika od siromaštva ili socijalne isključenosti i njene komponente, u %, 2015–2018.

	2015.	2016.	2017.	2018.
<b>Stopa rizika od siromaštva ili socijalne isključenosti</b>				
Ukupno stanovništvo	41,7	38,5	36,7	34,3
15–24 godine	46,2	45,0	40,1	38,8
15–19 godina	47,6	45,9	44,0	40,7
20–24 godine	44,9	44,3	37,1	37,2
Stopa rizika od siromaštva				
Ukupno stanovništvo	26,7	25,9	25,7	24,3
15–24 godine	32,1	33,9	30,2	30,2
15–19 godina	33,8	36,8	33,4	33,7
20–24 godine	30,7	31,3	27,7	27,3
Stopa izrazite materijalne uskraćenosti				
Ukupno stanovništvo	24,0	19,5	17,4	15,9
15–24 godine	24,9	20,6	17,1	15,5
15–19 godina	27,0	20,6	17,4	17,0
20–24 godine	23,1	20,5	16,9	14,3
Veoma nizak intenzitet rada				
Ukupno stanovništvo	15,6	15,7	14,5	13,0
15–24 godine	20,2	20,4	18,8	18,0
15–19 godina	21,0	22,4	19,3	17,8
20–24 godine	19,4	18,7	18,4	18,1

Izvor: SILC

U Tabeli 2 zapažamo da se rizik od siromaštva ili socijalne isključenosti za mlade 15–24 godine smanjuje sa povećanjem nivoa obrazovanja njihovih roditelja. Idući od nižeg do srednjeg obrazovanja, a zatim od srednjeg ka visokom nivou obrazovanja roditelja, stopa rizika od siromaštva ili socijalne isključenosti se gotovo prepolovila.

U periodu 2015–2018. godine, stopa rizika od siromaštva ili socijalne isključenosti porasla je za 14 p.p. među mladima koji žive sa roditeljima koji imaju niži nivo obrazovanja, ali se smanjila skoro u istoj meri među mladima čiji roditelji imaju visoko obrazovanje. Nešto manje smanjenje pokazatelja uočava se kod mladih čiji roditelji imaju srednje obrazovanje.

Takođe vidimo da 54,1% mladih čiji roditelji imaju niži stepen obrazovanja živi u domaćinstvima sa veoma niskim intenzitetom rada. Ovo je značajan porast sa 29,1% koliko je iznosilo u 2015. godini. Sa druge strane, u periodu 2015–2018. godine smanjio se udeo mladih čiji roditelji imaju srednje ili visoko obrazovanje u domaćinstvima veoma niskog intenziteta rada.

**Tabela 2:** Pokazatelji ranjivosti mladih prema nivou obrazovanja svojih roditelja, u %, 2015–2018.

	2015.	2016.	2017.	2018.
<b>Rizik od siromaštva ili socijalne isključenosti prema nivou obrazovanja roditelja (15-24 godine)</b>				
Niže obrazovanje roditelja	64,0	66,2	72,0	78,3
Srednje obrazovanje roditelja	48,3	47,1	39,9	36,4
Visoko obrazovanje roditelja	26,8	22,7	19,0	19,0

<b>Udeo lica koji žive u domaćinstvima sa veoma niskim intenzitetom rada prema obrazovanju roditelja (15-24 godine)</b>				
Niže obrazovanje roditelja	29,1	38,6	40,7	54,1
Srednje obrazovanje roditelja	20,3	18,3	15,7	13,4
Visoko obrazovanje roditelja	11,1	11,2	7,2	6,7

Izvor: Proračun autora na osnovu podataka iz SILC.

Sada pažnju usmeravamo na mlade koji žive u domaćinstvima veoma niskog intenziteta rada i istovremeno su izloženi riziku od siromaštva. Ova grupa mladih se nalazi u preseku ova dva skupa koja predstavljaju dve od tri komponente stope rizika od siromaštva ili socijalne isključenosti, i stoga su ranjiviji od onih pripadaju samo jednom ili drugom skupu. Ovaj indikator je dostupan u bazi podataka Eurostata za različite starosne intervale, ali nije dostupan za mlade uzrasta 15–24 i za dve grupe mladih 15–19 i 20–24.

U Srbiji 14,3% mladih starosti 15–24 godine živi u domaćinstvima sa veoma niskim intenzitetom rada i izloženi su riziku od siromaštva. Ovo su lica u domaćinstvima u kojima članovi ne rade ili rade veoma malo i koji takođe imaju relativno nizak dohodak. Udeo ove grupe mladih neznatno je smanjen, sa 14,9% u 2015. na 14,3% u 2018. godini ali je i dalje znatno viši nego u EU28 (4,5%), što je u skladu sa nižom prosečnom stopom domaćinstava sa veoma niskim intenzitetom rada i prosečnom stopom rizika od siromaštva u EU28.

**Tabela 3:** Stanovništvo u riziku od siromaštva koje živi u domaćinstvima veoma niskog intenziteta rada, u %, 2015–2018.

	2015.	2016.	2017.	2018.
Stanovništvo u riziku od siromaštva koje živi u domaćinstvima veoma niskog intenziteta rada				
Ukupno stanovništvo	10,9	10,9	10,4	9,6
15–24 godine	14,9	15,3	14,6	14,3
15–19 godina	15,7	17,6	14,7	15,7
20–24 godine	14,2	13,2	14,4	13,1

Napomena: Ukupno stanovništvo je za sve starosne grupe.

Izvor: SILC, RZS. Proračuni autora.

Od ukupnog broja mladih koji žive u domaćinstvima sa veoma niskim intenzitetom rada, skoro četiri od pet mladih osoba (79%) izloženo je riziku od siromaštva u 2018. godini (14.3%/18%). Ovo preklapanje dva skupa je razumljivo pošto je nizak intenzitet rada domaćinstva značajana determinanta niskog dohotka domaćinstva. Međutim, svaka druga mlada osoba koja se suočava sa rizikom od siromaštva (47,4%) živi u domaćinstvu veoma niskog intenziteta rada (14.3%/30.2%). To pokazuje da veoma nizak intenzitet rada domaćinstava nije jedini razlog zbog kog se mladi suočavaju sa rizikom od siromaštva, već su to i niske zarade mladih (čiji intenzitet rada nije toliko nizak) u kombinaciji sa socijalnim transferima, nedovoljnim da povećaju dohodak domaćinstva iznad praga rizika od siromaštva.

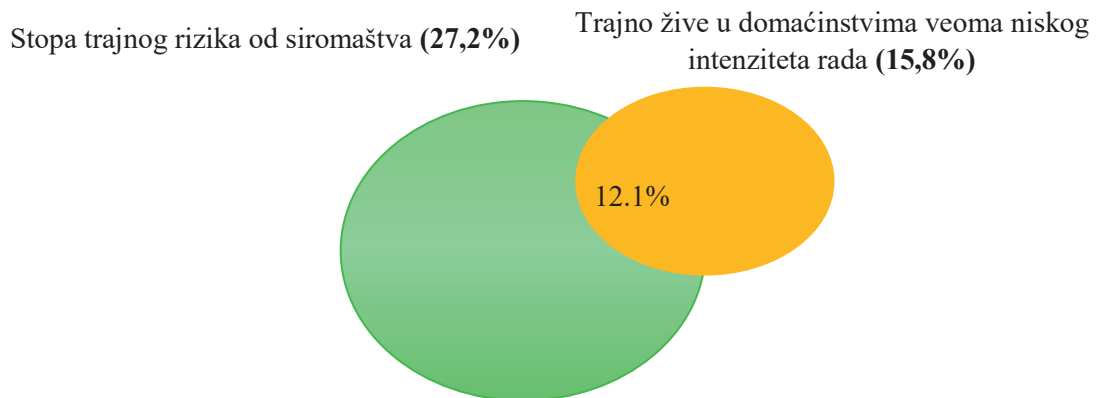
Stopa trajnog siromaštva za mlade starosti 18–24 godine je visoka jer je svaka četvrta mlada osoba u Srbiji bila u trajnom siromaštvu u 2018. godine (25,9%). Veća je verovatnoća da će se mladi naći u trajnom siromaštvu nego ukupno stanovništvo. Ova podgrupa mladih ljudi je izuzetno ranjiva pošto je njihov dohodak ispod granice relativnog siromaštva ne samo u tekućoj godini, već i u najmanje dve od prethodne tri godine. Praćenje promena ranjivosti mladih je posebno važna, jer verovatnoća da će osoba izaći iz siromaštva opada što duže ona ostaje u riziku od siromaštva [5]. Srbija ima najvišu stopu mladih u trajnom riziku od siromaštva među zemljama EU28, prosečna stopa u EU28 je 14,2%.

Siromaštvo mladih u Srbiji je dugoročna pojava, s obzirom na to da je 90% mladih ljudi starosti od 15 do 24 godine koji su bili siromašni takođe trajno siromašni. Ovo znači da su skoro svi mladi ljudi koji su se suočili sa rizikom od siromaštva u 2018, bili takođe siromašni tokom najmanje dve od prethodne tri godine. Zarobljenost u siromaštvu u mlađem uzrastu može da ograniči mogućnosti mladih da ostvare pun potencijal, da utiče na njihova dostignuća na polju obrazovanja, zdravlja i blagostanja, što zauzvrat može da smanji njihovu sposobnost da žive uspešno i produktivno i poveća rizik od trajnog siromaštva. U takvoj situaciji, siromaštvo se češće generacijski prenosi, odnosno velika je verovatnoća da će deca tih osoba takođe biti izložena siromaštvu ili socijalnoj isključenosti.

Pored toga, 15,8% mladih (15–24 godine) u Srbiji trajno živi u domaćinstvima sa veoma niskim intenzitetom rada. Veća je verovatnoća da će mladi živeti u takvim domaćinstvima nego ukupno stanovništvo

(10.8%). Čini se da su ti mladi ljudi zarobljeni u domaćinstvima u kojima članovi rade veoma malo pošto 88% mladih koji su živeli u domaćinstvima sa veoma niskim intenzitetom rada u 2018, živeli su u takvim domaćinstvima i u najmanje dve od prethodne tri godine.

Slika 1. pokazuje da je 12,1% mladih u 2018. godini u Srbiji iskusilo oba rizika. Od ukupnog broja mladih koji trajno žive u domaćinstvima sa veoma niskim radnim intenzitetom, 77% istovremeno su izloženi trajnom riziku od siromaštva (Slika 1). Ovakvo preklapanje dva skupa pokazuje da nizak radni intenzitet domaćinstva kako u tekućoj, tako i tokom najmanje dve od prethodne tri godine, predstavlja bitan faktor trajno niskog dohotka domaćinstva. Međutim, samo 44% mladih koji se suočavaju sa trajnim rizikom od siromaštva, trajno živi u domaćinstvima sa veoma niskim intenzitetom rada. Za ostalih 56% mladih koji se suočavaju sa trajnim siromaštvom, nizak dohodak i nedovoljni socijalni transferi su ključni razlozi za trajno siromaštvo.



**Slika 1.** Stopa trajnog rizika od siromaštva mladih i mladi koji trajno žive u domaćinstvima veoma niskog intenziteta rada (15–24), 2018, u %

Izvor: SILC, RZS. Proračuni autora.

Rezultati pokazuju da su se mladi suočili ne samo sa višestrukim, već i sa trajnim rizicima. Većina siromašnih je iskusila trajno siromaštvo, dok je većina onih koji žive u domaćinstvima sa veoma niskim intenzitetom rada hronično zarobljena u takvim domaćinstvima. Osim toga, izgleda da je najveći broj onih koji trajno žive u takvim domaćinstvima takođe izložen trajnom siromaštvu.

## 5. ZAKLJUČAK

U ovom radu ispitivali smo ranjivost mladih, korišćenjem mikro SILC podataka u periodu 2015–2018. godine. Mladi su ranjiviji u odnosu na ukupnu populaciju, i mladi uzrasta 15–19 godina su više izloženi ranjivosti u odnosu na mlade uzrasta 20–24. Iako su u bazi podataka Eurostata dostupni brojni indikatori siromaštva i socijalne isključenosti za različite starosne grupe, bilo bi korisno da se prate još neki indikatori za mlade koje smo prikazali u ovom radu. S obzirom na to da većina mladih u Srbiji živi u domaćinstvima sa svojim roditeljima, i da se odvajanje od roditelja kasno dešava, potrebno je pratiti indikatore prema obrazovnom nivou roditelja. Obrazovni nivo roditelja utiče na aktivnost roditelja na tržištu rada, kao i na dohodak domaćinstva. Zabrinjava činjenica da je četiri petine mladih čiji roditelji su nižeg obrazovnog nivoa u riziku od siromaštva ili socijalne isključenosti u 2018. godini, kao i da se razlika u stopama prema obrazovnom nivou povećala u posmatranom periodu. U 2018. godini je bila čak 60 p.p. veća za mlade čiji su roditelji niskog obrazovnog nivoa u odnosu na visoki obrazovni nivo. Više od polovine mladih čiji su roditelji nižeg obrazovnog nivoa žive u domaćinstvima veoma niskog intenziteta rada, u poređenju sa 6,7% mladih čiji su roditelji visokog obrazovnog nivoa.

Indikatore siromaštva i socijalne isključenosti mladih je potrebno kontinuirano pratiti, i država bi trebalo da svojim politikama podstiče aktivaciju mladih na tržištu rada kao i da pruži dobar obrazovni sistem koji će mladima omogućiti znanja i veštine za brzu integraciju na tržištu rada. Visoka stopa trajnog rizika od siromaštva znači da su ti mladi zarobljeni u siromaštvu tri od četiri uzastopne godine i da se verovatnoća izlaska iz siromaštva smanjuje protokom vremena. Imajući u vidu da je siromaštvo nasledno, i da je velika verovatnoća da će osoba koja je odrasla u siromaštvu, svoju decu takođe gajiti u siromaštvu, znači da ovom problemu treba ozbiljno pristupiti u cilju prekidanja začaranog kruga siromaštva.

**LITERATURA**

- [1] Aassve, A., Iacovou, M. & Mencarini, L. (2006). Youth poverty and transition to adulthood in Europe. *Demographic Research*, 15 (July–December 2006), 21–50.
- [2] Aleksić, D. & Vuksanović, N. (2019), „Nije lako biti mlad – Ispitivanje faktora koji utiču na zapošljivost mladih u Srbiji”, u Jakšić, M. i Prašević, A. (eds.) *Ekonomska analiza interakcija tržišta, institucija i znanja – dometi i pouke u javnim politikama*, Ekonomski fakultet, Univerzitet u Beogradu, 123–141.
- [3] ETF. (2021). “Youth situation in Serbia. Employment, skills and social inclusion”
- [4] Eurostat. (2014). „Lot 1: EU-SILC (European Union Statistics on income and Living Conditions): Methodological studies and publications“, Working paper with the description of the „Income and living conditions dataset“
- [5] Eurostat. (2018). *Living conditions in Europe* (2018 ed.). Publications Office of the European Union.
- [6] Fahmy, E. (2014). „The Complex Nature of Youth Poverty and Deprivation in Europe“, in Antonucci, L., Hamilton, M., Roberts, S. (eds.) *Young People and Social Policy in Europe. Work and Welfare in Europe*, Palgrave Macmillan, London.
- [7] Krstić, G. (2017). Would an increase in low work intensity contribute to reducing poverty and inequality in Serbia. *Ekonomске ideje i praksa*, 24, 37–52.
- [8] Vukasović, N., Joksimović, Lj. & Aleksić, D. (2018). School to work transition in Serbia: returns to investment in education of youth. *Industrija*, 46 (1), 115–136.



## STATISTICAL PROPERTIES OF THE LIMIT ORDER BOOK OF THE NASDAQ STOCK MARKET

DRAGANA RADOJIČIĆ<sup>1</sup>, NINA RADOJIČIĆ MATIĆ<sup>2</sup>, ŽELJKO JOVIĆ<sup>3</sup>

<sup>1</sup> University of Belgrade – Faculty of Economics and Business, Serbia, dragana.radojicic@ekof.bg.ac.rs

<sup>2</sup> University of Belgrade, Faculty of Mathematics, Serbia, nina@matf.bg.ac.rs

<sup>3</sup> University of Belgrade – Faculty of Economics and Business, Serbia, zeljko.m.jovic@gmail.com

**Abstract:** *The stock market produces an immense amount of data and the key is to extract informative features. This work is based on limit order book data that replicates the entire NASDAQ stock market data. We focus on extracting relevant features from the dataset in order to acquire information about the dynamics of the stock market. Especially, the quoted spread, which is a pertinent component of the market liquidity, is studied. Furthermore, we incorporate technical analysis to get more insights into market behavior.*

**Keywords:** *Limit Order Book, Stock market data properties, Technical indicators, Kahan summation algorithm*

### 1. INTRODUCTION

The Limit Order Book keeps track of the evolution of the market prices and the corresponding number of orders, both on the ask and on the bid side. Thus, studying the LOB data can be beneficial to understand stock market dynamics. Therefore, there has been some interesting research concerning modeling order book dynamics. Representing the order book dynamics via Markovian queue system is done in [4]. In [8], the authors established the limit theorem when the tick size converges to zero in the model where the best ask and best bid price processes are modeled by a two-dimensional reflected Brownian motion. In [15] new feature selection methods were introduced and it was shown that they improve the Long-term memory neural networks model for the market data vector classification task.

#### 1.1. The Limit Order Book

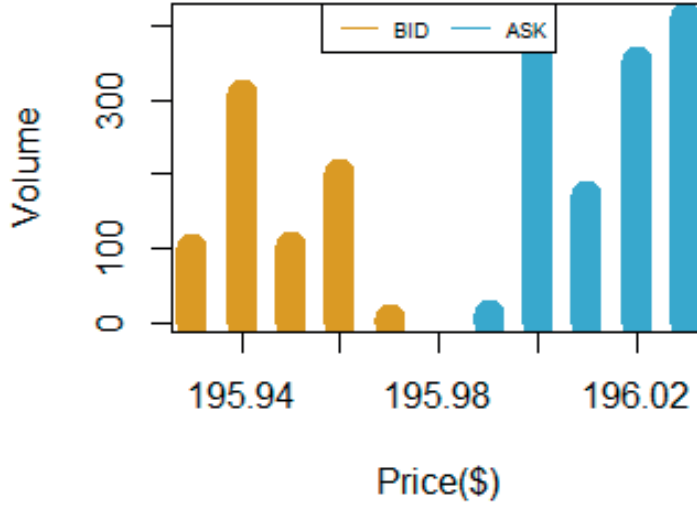
The Limit Order Book (LOB) is a file used to organize all equity, futures, derivatives, and all possible transactions on the electronic markets. The orders available to be sold are placed on the ask side, while the orders available to be bought are placed on the bid side (see Figure 1). It keeps track of all orders that can possibly be performed on the market. Furthermore, LOB records characteristics such as the price of the order, the volume of the order, the sign of the order (buy or sell), etc. Available prices are represented by the points on the price grid, and the best ask and best bid price denote the lowest price on the ask side and the highest price on the bid side. The first price below the best bid price is placed on the price level 2 on the bid side, while the lowest available price higher of the best ask price is placed on the price level 2 on the ask side, and etc. Tick is a minimal distance between two price levels (points in a discrete price grid). The main job of the LOB is to keep track of the LOB data (prices and volumes) up to the requested number of levels on both sides. For a comprehensive overview of the LOB's mathematical concept the reader is referred to Section II in [10].

#### 1.2. Notation

Let two-parameter process  $\{V(t, p) : t \in \mathcal{T}, p \in \delta\mathbb{N}\}$  represents order volume process, where  $\mathcal{T}$  is the set of all events timestamps recorded in the order book and  $\delta$  denotes the tick size. Thus,  $V(t, p)$  denotes the volume of order awaiting execution at time  $t$  at price level  $p$ . Denote by  $P_{t,1}^a$  the best ask price (i.e. price on the first level on the ask side), and by  $P_{t,1}^b$  the best bid price (i.e. price on the first level on the bid side). The order book of  $N$  levels essentially represents a snapshot that contains the best  $N$  prices on the ask side  $P_t^a = (P_{t,1}^a, P_{t,2}^a, \dots, P_{t,N}^a)$  and the best  $N$  prices on the bid side  $P_t^b = (P_{t,1}^b, P_{t,2}^b, \dots, P_{t,N}^b)$ , and also corresponding volume on both sides. Precisely, snapshot of the LOB contains information about the number of shares available on the both sides, the vector describing volume on the ask side  $V_t^a = (V_{t,1}^a, V_{t,2}^a, \dots, V_{t,N}^a)$  and the vector describing volume on the bid side  $V_t^b = (V_{t,1}^b, V_{t,2}^b, \dots, V_{t,N}^b)$ .

The *Quoted Spread*, at time  $t$  is the gap between the best bid and the best ask price, i.e.:

## Limit Order Book Volume for AAPL



**Figure 1** Snapshot of the NASDAQ limit order book for AAPL shares (July 07, 2019) at 10:16:05 for 5 levels. On the bid/ask side are placed outstanding buy/sell orders (gold/blue), and the best bid price is \$195.97 with volume of 40 orders, while the best ask price is \$195.99 with volume of 50 orders.

$$\text{QuotedSpread}_t = P_{t,1}^b - P_{t,1}^a. \quad (1)$$

The *mid-price*, which can be understood as the approximation of the real market price, represents an arithmetic average of the best bid and best ask price, i.e. the mid price at time  $t$  denoted by  $\text{MidPrice}_t$  is defined as:

$$\text{MidPrice}_t = \frac{1}{2}(P_{t,1}^a + P_{t,1}^b). \quad (2)$$

Using the measure theory concept the limit order book can be represented by a following measure  $\mu$  as

$$\mu_t = \sum_{p \in \delta\mathbb{N}} V(t, p) \varepsilon_p. \quad (3)$$

The mid-price involves during the time, in order to have a picture how the LOB shape changes with respect to the position of the current mid-price, we consider the centered order book

$$C_t(p) = V(t, \text{MidPrice} + p). \quad (4)$$

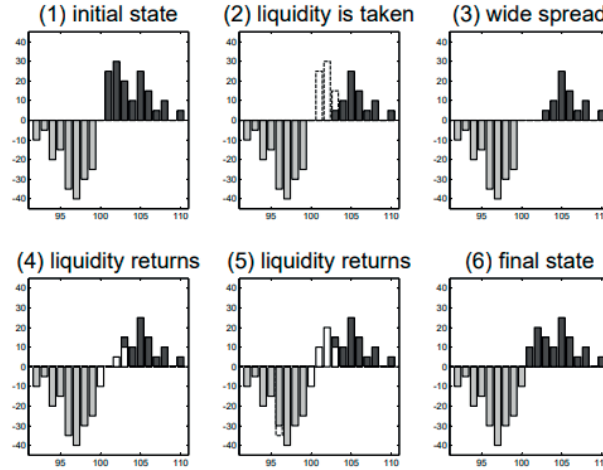
Furthermore, due to price change during time, it is convenient to capture the order volume at a distance  $i$  ticks from the mid-price. Thus, let  $X_i(t)$  be the number of shares available at time  $t$  at price  $\text{MidPrice}_t + i * \delta$ . Thus,  $X_i(t) = V(t, S_t + i\delta)$  when the mid-price is a multiple of the tick size, with  $i \in \mathbb{Z}^+$  for ask orders and  $i \in \mathbb{Z}^-$  for bid orders.

In order to also include cases when the mid price take half-integer values, the  $X_i(t)$  is defined with measure  $\mu$  by following formulas

$$X_i(t) = v_t^- \left( \left[ i\delta - \frac{\delta}{2}, i\delta + \frac{\delta}{2} \right) \right), \quad \text{for } i > 0 \quad (5)$$

$$X_i(t) = v_t^+ \left( \left( i\delta - \frac{\delta}{2}, i\delta + \frac{\delta}{2} \right] \right), \quad \text{for } i < 0. \quad (6)$$

Figure 2 (Figure 1.1. in [10]) illustrates an interaction between order book and incoming order flow, where positive (resp. negative) part corresponds to buy (resp. sell) orders. See also [1] and [2] to understand the liquidity and the interactive relation between LOB and incoming and outgoing orders.



**Figure 2** The interplay between the incoming order flow and the order book

### 1.3. The data set

This research is supported by empirical results that we have obtained by analyzing data from the NASDAQ, precisely the Limit Order Book System- Efficient Reconstructor (LOBSTER), see [11].

The 'message' file carries information of the kind of event that has updated the LOB and each row has the following columns: Time, Type of Event, Order ID, Size of the order, Price, Direction of a trade (buy/sell). The 'orderbook' file holds information of the prices and volumes at each price level of the LOB up to the requested number of levels.

LOBSTER outputs a huge amount of data for each stock symbol listed on NASDAQ. Since our goal is to derive attributes from the data set we use the framework for processing the LOBSTER dataset introduced in [14].

## 2. EXTRACTED FEATURES

By simply merging each row of the 'orderbook' file with the corresponding row from the 'message' file we obtain data consisting of the vectors in the form of

$x_t = (bidLevel_1, bidVolume_1, askLevel_1, askVolume_1, bidLevel_2, bidVolume_2, askLevel_2, askVolume_2, \dots, bidLevel_n, bidVolume_n, askLevel_n, askVolume_n, time, EventType, OrderId, Size, Price, Direction)$ . Thus, we have a dataset  $\mathcal{D} = \{x_t | 0 \leq t \leq \text{amount of events per day}\}$ .

### 2.1. Raw aggregation

During the raw aggregation part, the idea is to aggregate data with respect to 1min length intervals. Precisely, we divide data into 1min length non-overlapping intervals, and we apply different functions over those intervals to obtain a vector that encapsulates the LOB behavior during that time interval.

During the raw aggregation part, the idea is to aggregate data with respect to 60s length intervals. Precisely, we divide data into 1min length non-overlapping intervals, precisely:

$$\mathcal{P}_{60s} = \{\mathcal{T}_{j-60s}^{(j+1) \cdot 60s} | 0 \leq j < \lfloor \frac{\text{trading day duration}}{60s} \rfloor\}, \text{ where } \mathcal{T}_{j-60s}^{(j+1) \cdot 60s}.$$

Further, we apply different aggregation functions  $\mathcal{A} = \{a_i\}_{i=1}^m$  over those intervals to obtain a vector that encapsulates the LOB behavior during that time interval. Each aggregation function takes as input the part of data  $\mathcal{T}_{j-60s}^{(j+1) \cdot 60s}$ . The final output for each time interval  $\mathcal{T}_{j-60s}^{(j+1) \cdot 60s}$  is the vector  $y_j^{(j+1) \cdot 180}$  which consists of the features such as: the number of canceled limit orders, the open price, the close price, the average mid-price, the number of cancelled orders, etc.

### 2.2. Technical indicators

The technical analysis roots are found the observation of Dutch financial markets in the 17th century, see [17]. Nowadays, technical analysis is an integral part of the trading industry. A large number of new technical indicators and approaches are known, see for example [7], [19], [3]. Studying historical stock market data can

be beneficial since LOB data can be informative (see [5], [18], [13]). The basic technical analysis principle is that price reflects all relevant information and contains predictive power. Therefore, rather than looking into exogenous factors of different nature, such as economics, news, politics, etc., technical analysts study the price behavior to determine a security's trading pattern. We have incorporated technical indicators by using the open library *ta-lib* (see [16]), which contains the algorithmic implementations for some technical indicators.

### 3. QUANTITATIVE ANALYSIS

#### 3.1. Kahan summation algorithm

When it comes to summarizing a sequence of finite-precision floating-point numbers numerical errors can occur in computational calculations. Thus, in this research, we use the Kahan summation algorithm that significantly reduces the numerical error in the total compared to the obvious approach. The basic pseudocode is given as Algorithm 1, while the analysis and improvement of Kahan's summation algorithm can be found in [12].

---

#### Algorithm 1 Kahan summation algorithm

---

**Input:**  $A$  - an array

**Output:**  $sum$  - the sum of the array  $A$

```

1:  $sum = 0.0$ 
2:  $c = 0.0$ 
3: for  $i = 1$  to  $length(A)$  do
4:    $y = A[i] - c$ 
5:    $t = sum + y$ 
6:    $c = (t + sum) - y$ 
7:    $sum = t$ 
8: end for
9: return  $sum$ 

```

---

#### 3.2. Statistical analysis: The bid-ask spread behavior

The study of the quoted spread is important since the changes in the price paid by the buyer and received by the seller indicate changes in liquidity, see [9]. Authors in [6] investigated that the quoted spread is usually equal to the tick size. The liquidity is very important for the proper functioning of the market. Smaller values of the quoted spread indicates the greater the liquidity. At some stock markets, for example NASDAQ, traders provide liquidity.

For each  $t \in \mathcal{T}$  the spread is calculated, and then the average spread over all trading times is calculated using the Kahan Summation algorithm (to avoid numerical errors by summing up large numbers). Program simulation has output that the average spread is 0.011174535151268644\$ for the MU stock on the selected trading day (2019-07-01). The histogram in Figure 3 presents the frequency distribution of the bid-ask spread, so we see that the bid-ask spread is most of the time equal to one tick, which is particularly for the MU stock for the selected trading day (2019-07-01) equal to 0.01\$, i.e.  $\delta = 0.01\%$ .

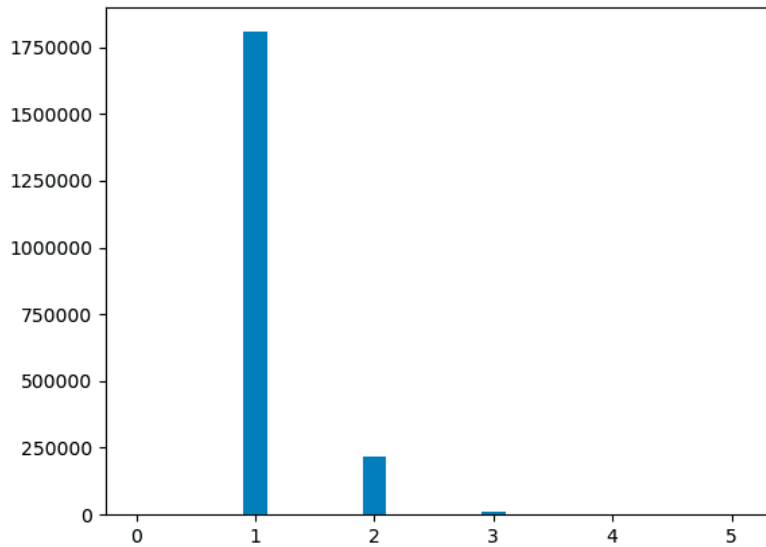
#### 3.3. Average order book profile

The average order book profile is defined as

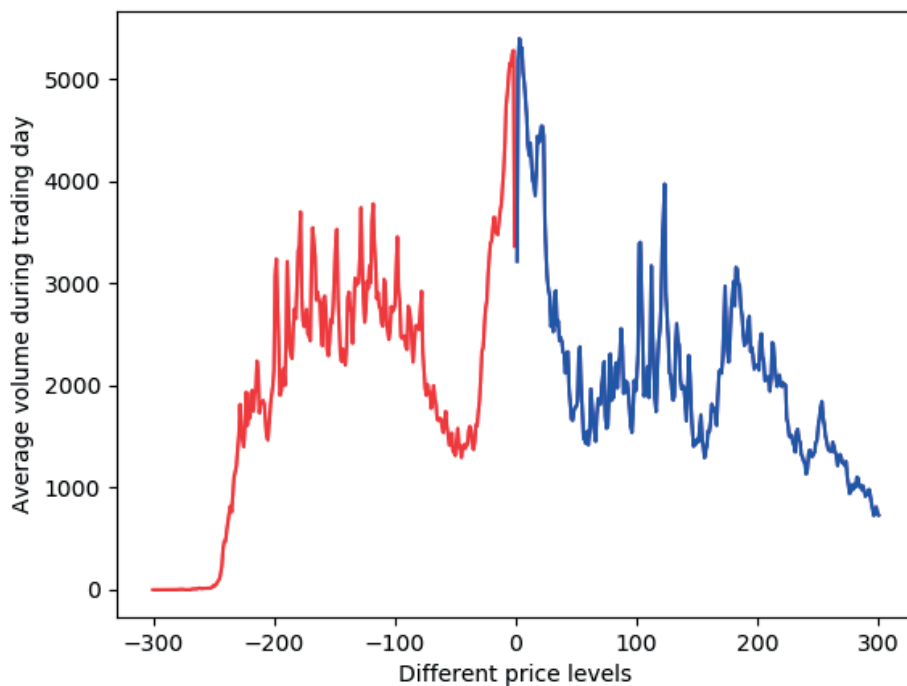
$$\bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_i(t), \quad i \in \{1, 2, \dots, d_{maxAsk}\} \wedge i \in \{-1, -2, \dots, -d_{maxBid}\}, \quad (7)$$

where  $d_{maxAsk}$ ,  $d_{maxBid}$  are the maximum distances from the mid-price at which an event has been recorded on the ask, bid side respectively.

As the average order book profile in Figure 4 shows, the choice  $N = 250$  covers  $> 90\%$  of limit orders. Henceforth we will represent the order book by its centered profile Eq. 4 with  $N = 250$  tick distances on each side. In the Figure 4, the average volume on every possible distance from the running mid-price over time is depicted. The x-axis captures all possible distances (in dollars) during one day, that is why the plot is dense.



**Figure 3** The histogram that shows frequency distribution of the spread in number of ticks for the selected trading day (2019-07-01) for the MU stock.



**Figure 4** The average order volume  $\bar{X}_k$  on  $k$  ticks from the running mid-price over time during one day (2019-07-01) for the MU stock.

In the Figure 4, the average volume on every possible distance from the running mid-price over time is depicted. The x-axis captures all possible distances (in dollars) during one day, that is why the plot is dense.

Note that in order to avoid error in numerical calculation of average value (i.e. summation of large values) the Kahan summation algorithm is used. The main idea is to keep a separate running compensation instead of calculating just a sum and dividing it by the total count.

#### 4. CONCLUSION AND FUTURE WORK

This work investigates the possibilities of extracting relevant features from the LOBSTER data set in order to potentially train neural network models on top of these features. When working with such an immense amount of data, it is crucial to choose relevant features that derives additional and useful information. Future work can incorporate social media networks to build a model that can effectively classify each market data vector and hopefully predict when is the good time to sell/buy.

#### Acknowledgement

The authors are thankful to Professor Thorsten Rheinländer for his help in doing the LOBSTER data analysis and valuable and insightful suggestions.

#### REFERENCES

- [1] Bouchaud, J. P., Mézard, M., & Potters, M. (2002). Statistical properties of stock order books: empirical results and models. *Quantitative finance*, 2(4), 251.
- [2] Burg, M. B., & Ferraris, J. D. (2008). Intracellular organic osmolytes: function and regulation. *Journal of Biological Chemistry*, 283(12), 7309-7313.
- [3] Colby, R. W. (2003). *The encyclopedia of technical market indicators*. McGraw-Hill.
- [4] Cont, R., & De Larrard, A. (2013). Price dynamics in a Markovian limit order market. *SIAM Journal on Financial Mathematics*, 4(1), 1-25.
- [5] Cont, R., Kukanov, A., & Stoikov, S. (2014). The price impact of order book events. *Journal of financial econometrics*, 12(1), 47-88.
- [6] Dayri, K., & Rosenbaum, M. (2015). Large tick assets: implicit spread and optimal tick size. *Market Microstructure and Liquidity*, 1(01), 1550003.
- [7] Gould, M. D., Porter, M. A., Williams, S., McDonald, M., Fenn, D. J., & Howison, S. D. (2013). Limit order books. *Quantitative Finance*, 13(11), 1709-1742.
- [8] Horst, U., & Paulsen, M. (2017). A law of large numbers for limit order books. *Mathematics of Operations Research*, 42(4), 1280-1312.
- [9] Muranaga, J., & Ohsawa, M. (1997). Measurement of liquidity risk in the context of market risk calculation. a BIS volume entitled *The Measurement of Aggregate Market Risk*.
- [10] Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin.
- [11] "Lobster academic research data. <https://lobsterdata.com>. Accessed: 2018-08-05."
- [12] Ozawa, K. A. Z. U. F. U. M. I. (1983). Analysis and improvement of Kahan's summation algorithm. *Journal of Information Processing*, 6(4), 226-230.
- [13] Palguna, D., & Pollak, I. (2016). Mid-price prediction in a limit order book. *IEEE Journal of Selected Topics in Signal Processing*, 10(6), 1083-1092.
- [14] Radojičić, D., & Kredatus, S. (2020). An approach for processing data from NASDAQ stock exchange database. In *Proceedings of the 10th International Conference on Information Society and Technology (ICIST 2020)*. Society for Information Systems and Computer Networks.
- [15] Radojičić, D., Radojičić, N., Kredatus, S. (2021). A multicriteria optimization approach for the stock market feature selection. *Computer Science and Information Systems*, 18(3), 749-769.
- [16] Ta-lib: Technical analysis library. <https://www.ta-lib.org/>. Accessed: 2022-05-03.
- [17] de la Vega, J. (1957). *Confusion de Confusiones [1688]: Portions Descriptive of the Amsterdam Stock Exchange (No. 13)*. Colchis Books.
- [18] Zheng, B., Moulines, E., & Abergel, F. (2012). Price jump prediction in limit order book. *arXiv preprint arXiv:1204.1381*.
- [19] Zhou, F., Zhang, Q., Sornette, D., & Jiang, L. (2019). Cascading logistic regression onto gradient boosted decision trees for forecasting and trading stock indices. *Applied Soft Computing*, 84, 105747.

## THE LINK BETWEEN TECHNICAL AND TECHNOLOGICAL PROGRESS, THE DEVELOPMENT OF NEW DIGITAL SERVICES FOR PUBLIC SECTOR GOVERNANCE AND ECONOMIC DEVELOPMENT

PETRA VUJKOVIĆ<sup>1</sup>

<sup>1</sup> University of Ljubljana – Faculty of Public Administration, [petra.vujkovic@fu.uni-lj.si](mailto:petra.vujkovic@fu.uni-lj.si)

---

**Abstract:** *Both the technology environment and the broader external environment (such as the economic, political, and social) have an impact on how e-Government evolves. The issues and requirements of society, together with technological advancements, have thus formed the various e-Government waves that have resulted from these evolutions. Similarly, in industry, we can distinguish some distinct generations of it, driven by technological evolutions. The purpose of this study is to systematize and better understand Digital Government Transformation, within the spectrum of technological innovations and in the light of economic development. In doing so, the study presents a prevailing discussion of past and recent trends that strengthen the digital transformation of government. Moving forward, we will consider the DESI index, which includes e-government services in its structure (i.e., the fourth section), and compare various EU and OECD countries based on how these services impact economic development and growth.*

**Keywords:** *Analysing data, Economic development, EU/OECD countries, Digital Government Transformation*





## ULOGA OPERACIONIH ISTRAŽIVANJA U EKONOMIJI: BIBLIOMETRIJSKA ANALIZA SRPSKIH ISTRAŽIVAČKIH REZULTATA

### THE ROLE OF OPERATIONAL RESEARCH IN ECONOMICS: BIBLIOMETRIC ANALYSIS OF SERBIAN PUBLICATIONS

MLADEN STAMENKOVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, mladen.stamenkovic@ekof.bg.ac.rs

**Rezime:** U ovom radu želimo da utvrdimo ulogu operacionih istraživanja u publikacijama vezanim za ekonomske nauke unutar srpske akademske zajednice. Uvidom u sve objavljene radove autora iz Srbije u periodu od 1996. do 2022. godine u relevantnoj citatnoj bazi, pokušaćemo da bibliometrijskom analizom ukažemo koje su to metode operacionih istraživanja bile korišćene u ovim publikacijama. Primenom mapiranja ključnih reči, prikazaćemo ulogu i značaj dve najrelevantnije familije metoda, višekriterijumskog odlučivanja i analize obavljanja podataka kao dobar komplement, standardnim ekonometrijskim metodama.

**Ključne reči:** Operaciona istraživanja, ekonomija, bibliometrijska analiza, Srbija

**Abstract:** In this paper, we aim to determine the role of operational research in publications of the Serbian research community in the field of economics. We use bibliometric analysis for all published papers of the authors with affiliation from Serbia from 1996 to 2022 in the relevant citation database to pinpoint relevance of the operational research methods in these publications. Using bibliometric maps of keywords, we show the role of two most relevant groups of methods, multiple criteria decision aiding, and data envelopment analysis, as the good complementary alternative to econometric modeling.

**Keywords:** Operational Research, Economics, bibliometric analysis, Serbia

#### 1. UVOD

Operaciona istraživanja jedna su od naučnoistraživačkih oblasti sa dugom tradicijom u srpskom akademskom prostoru. I sam Simpozijum o operacionim istraživanjima može se pohvaliti kao jedna od najdugovečnijih konferencija iz oblasti operacionih istraživanja. Interesantno, Sym-Op-Is nastaje i pre Evropske asocijacije udruženja za operaciona istraživanja (EURO) koja se formira 1975. godine i njihove konferencije EURO, danas najveće konferencije u Evropi koja je nedavno održala svoj 32. skup.

Predmet istraživanja ovog rada je da utvrdi vezu i ulogu operacionih istraživanja u publikacijama vezanim za ekonomske nauke. U tu svrhu, preuzeti su svi radovi koje su istraživači sa afilijacijom iz Srbije objavili u periodu 1996-2022. Za potrebe analize koristili smo JCR – Web of Science (JCR-WoS), odnosno Social Science Citation Index (SSCI) listu. Razlog je jednostavan. Pravilnici Republike Srbije, poput Pravilnika o kategorizaciji i rangiranju naučnih časopisa, kao i pravilnici o minimalnim uslovima za sticanje zvanja definišu međunarodni naučni časopis kao časopis koji je referisan u međunarodnim citatnim bazama Journal Citation Report i Web of Science (Science Citation Index Expanded, Social Science Citation Index, Arts and Humanities Citation Index).

Bibliometrijska analiza predstavlja dobru polaznu osnovu uvida u neki naučnu oblast. Prisutan je i veliki broj bibliometrijskih radova u oblasti operacionih istraživanja, neki fokusirani na časopise, neki na uže naučne celine od interesa [1–3]. Naravno, ni ekonomija nije izuzetak [4,5], uz napomenu da postoji časopis Journal of economic surveys, posvećen isključivo ovakvoj vrsti radova, rangiran u kategoriji M21 u oblasti Economics, sa IF2 faktorom za 2021. godinu od 4,142 (ukupno, pozicija 83 od 380 radova izlistanih u oblasti Economics na JCR-WoS listi).

## 2. RAZVOJNI PUT EKONOMSKIH NAUKA I ULOGA OPERACIONIH ISTRAŽIVANJA

U periodu od 1996 do 2022. godine objavljeno je ukupno 710 časopisa u oblasti „Ekonomija“<sup>1</sup>. Prva dekada dostupnih podataka sastoji se od manjeg broja radova, uglavnom iste grupe autora [6,7]. Veći broj radova pojavljuje se tek od 2010. godine kada se u institucionalni okvir uvodi ova lista kao relevantna, ali i sa pojavom prvih časopisa iz regiona kao što su Economic research-Ekonomska istraživanja, Panoeconomicus, Zbornik radova ekonomskog fakulteta u Rijeci-Proceedings of Rijeka faculty of economics. Upravo prva dva časopisa i imaju najviše objavljenih radova istraživača sa afilijacijom iz Srbije, 69 i 61 respektivno, što je 18,3% svih objavljenih radova.

Za nas su od interesa radovi koji se nalaze u ovom skupu podataka a koji su bazirani na metodama operacionih istraživanja. U svrhu te pretrage koristimo KeyWords Plus algoritam [8] koji uočava fraze u naslovima citiranih referenci, a ne nužno u samom naslovu rada, te na taj način povećava snagu pretraga posebno u multidisciplinarnim poljima<sup>2</sup>. Na ovaj način izdvojili smo ukupno 66 radova, odnosno 9,3% svih radova objavljenih u oblasti Ekonomija<sup>3</sup>. Bazirajući se na metodološkoj osnovi radove možemo izdvojiti dve glavne tematske celine – višekriterijumsko odlučivanje i analiza obavijenih podataka (DEA) uz neophodnu napomenu da je prisutan i veliki broj radova koji se bazira na rasplinutoj logici.

### 2.1. Višekriterijumsko odlučivanje i radovi u oblasti ekonomije

Najzastupljeniji radovi predstavljaju primenu metoda višekriterijumskog odlučivanja na razne ekonomske probleme, poput urbanog razvoja [9], turizma [10], ili finansija [11].

Radovi koji su koristili AHP/ANP metode su mnogobrojni [9,10]. Kod njih se može videti i česta multidisciplinarnost časopisa koja za posledicu ima veliki broj primena iz oblasti poput saobraćaja [12,13]. Modeli sa primenama rasplinite logike zastupljeni su kroz metode poput fuzzy AHP [11,14] ili drugim modifikacijama originalnih metoda [15]. Očekivano, prisutne su i druge poznate metode kao što je ELECTRE [16] ili PROMETHEE [17,18].

Interesantno je da od svih 710 radova u oblasti Ekonomija, u pet najcitiranijih nalazi se četiri koji koriste metode operacionih istraživanja, od kojih su tri primene višekriterijumskog odlučivanja [11,19,20], a jedna koristi analizu obavijenih podataka [21].

### 2.2. Analiza obavijenih podataka i radovi u oblasti ekonomije i ostale primene

Analiza obavijenih podataka i njena input/autput ideja predstavlja prirodno dobru polaznu osnovu za primenu u ekonomiji te ne treba da iznenadi veći broj radova koji su koristili ovu metodu [21–23], uz fokus na bankarski sektor u većem broju radova [22,24].

Svaki časopis unutar WoS razvrstava se unutar 250 različitih naučnih celina. Časopisi često imaju multidisciplinarni karakter, i rangirani su u nekoliko kategorija. Tako, na primer, imamo značajan broj radova koji se nalaze na vezi između ekonomije i saobraćaja i transporta [25–27] i mimo primena modela višekriterijumskog odlučivanja. Uočljivo je prisustvo radova koji se bave lancem snabdevanja [28,29], trenutno veoma popularnom temom.

### 2.3. Pozicija istraživača u oblasti operacionih istraživanja

Kada je reč o primena metoda operacionih istraživanja u oblasti ekonomije, ističe se jedno ime koje objavljuje u ovakvim časopisima, Dragiša Stanujkić [20,30,31] koji ima objavljenih 17 radova u oblasti. Poređenja radi, dva autora sa najviše radova u oblasti Ekonomija su Kosta Josifidis i Dejan Živkov sa malo preko 30 radova.

## 3. MAPIRANJE KLJUČNIH REČI

Uz i dalje prisutni fokus na KeyWords Plus i tako generisani skup ključnih reči izolovanih iz radova, želimo da kreiramo mapu pojmova upotrebljenih u svih 710 radova u oblasti Ekonomija. Ovo će nam pomoći da vidimo metodološke pravce u ovim radovima, i još jednom primetimo značajan udeo operacionih istraživanja

<sup>1</sup> JCR-WoS lista časopisa podeljena je na 250 oblasti. Za potrebe ove analize preuzeti su članci iz ove oblasti po njihovoj kategorizaciji.

<sup>2</sup> Za više detalja moguće je pogledati i članak dostupan na linku ispod, gde se daje odličan pregled i razvoj ideje, koja seže i do rada [35] istog autora u časopisu Science koji se smatra jednim od osnovnih radova u istoriji stvaranja citatnih indeksnih baza. [http://www.garfield.library.upenn.edu/papers/jasis44\(5\)p298y1993.html](http://www.garfield.library.upenn.edu/papers/jasis44(5)p298y1993.html)

<sup>3</sup> Preciznijom analizom, utvrđeno je još nekoliko radova koji su uvršteni u razmatranje, ali je prvobitna pretraga kroz KeyWords plus ključne reči dala 66 jasnih rezultata.

u ovoj oblasti. Za potrebe analize koristimo VOSViewer, softver otvorenog koda koji služi za mapiranje, vizualizaciju i istraživanje bibliometrijskih podataka [32]. Output softvera je dvodimenzionalna mapa, gde grane predstavljaju veze između pojmova, a veličina čvora učestalost njegovog ponavljanja. Takođe, softver generiše klasterne pojmove.

Ovom metodom softver je uočio 1329 pojmova, od kojih je 251 sa preko tri ponavljanja koliko smo postavili kao granicu. Unutar ove liste pojmova sklonili smo generičke i stavili fokus na metodološke termine. Slika 1 daje prikaz rezultata, sa ukupno devet klastera. Pokušaćemo da kvalitativno opišemo svaki, oslanjajući se na termine u njima.

Prva dva klastera vrlo jasno predstavljaju ekonometrijska istraživanja. Zamisao prilikom kreiranja mape bila je da se generišu klasteri gde ćemo imati isprepletana pojavljivanja metoda i primena. Ipak, generisani su klasteri poput trećeg koji ima i dve najčešće zastupljene reči, „performance“ i „determinants“, ali sa nazivom „Sektorski klaster“ smo želeli da sugerišemo da je on kreirao spoj svih oblasti a individualnim granama se može posmatrati koja metoda se najčešće povezuje sa kojim pojmom.

Četvrti klaster skup je vrlo očekivanih reči (videti Tabelu 1) koje bi opisale klaster povezan sa višekriterijumskim odlučivanjem. Slično se može reći i za peti klaster koji je vrlo jasno izdvojio sve povezano sa analizom obavljanja podataka. Jedino je vredno pomena da je pojam group decision making ušao u ovaj klaster. Uvidom u grane, ovaj klaster je značajno povezan sa čvorom „Performance“.

Klasteri nazvani „Finansijska tržišta“ podeljenji su u dve celine. Oba su isprepletana sa ekonometrijskim metodama i ovo ne predstavlja nikakvo iznenađenje, i veliki broj radova u oblasti ekonomije upravo je povezan sa ovim temama.

Osmi klaster sa naslovom „Korporativne finansije“ osim nekih opisnih pojmova ima u sebi efikasnost, i ovo je u potpunosti u skladu sa već spomenutom činjenicom da je primena DEA metoda često bila vezana za bankarski sektor, to se sada može primetiti i iz ovog mapiranja.

Klaster	Naziv	Glavni pojmovi
I (zeleni)	Ekonometrijska istraživanja 1	Panel data, GMM, unit root tests, decomposition economic growth, foreign direct investments
II (narandžasti)	Ekonometrijska istraživanja 2	Time series, dynamics, unit root, causality, cointegration, hypothesis, regression
III (crveni)	Sektorski klaster 1	Performance, determinants, transition, inequality, business, governance, education
IV (svetlo plavi)	Višekriterijumsko odlučivanje	Optimization, AHP, ANP, MCDM, criteria, weight
V (roze)	DEA	Data envelopment analysis, DEA, technical efficiency
VI (braon)	Finansijska tržišta 1	Markets, volatility, spillovers, monetary policy, inflation
VII (tamno plavi)	Finansijska tržišta 2	Integration, volatility spillover, crisis, stock market, structural breaks
VIII (ljubičasti)	Korporativne finansije	Market, firms, companies, corporate governance, efficiency
IX (žuti)	Sektorski klaster 2	Impact, growth, trade, sector, wages, globalization



- [2] Liao H., Tang M., Li Z., Lev B. (2018). Bibliometric analysis for highly cited papers in operations research and management science from 2008 to 2017 based on Essential Science Indicators. *Omega*, 88, 223–236.
- [3] Laengle S., Merigó J., Miranda J., Słowiński R., Bomze I., Borgonovo E., et al. (2017). Forty years of the European Journal of Operational Research: A bibliometric overview. *European Journal of Operational Research*, 262(3), 803–816.
- [4] Bahoo S, Alon I, Floreani J. (2020). Corruption in economics: a bibliometric analysis and research agenda. *Applied Economics Letters* 28(7), 565-578.
- [5] Bonilla C., Merigó J., Torres-Abad C. (2015). Economics in Latin America: a bibliometric analysis. *Scientometrics*, 105(2), 1239–1252.
- [6] Petrović P., Bogetić Ž., Vujošević Z. (1999). The Yugoslav Hyperinflation of 1992-1994: Causes, Dynamics, and Money Supply Process. *Journal of Comparative Economics*, 27(2), 335–353.
- [7] Petrović P., Mladenović Z. (2000). Money Demand and Exchange Rate Determination under Hyperinflation: Conceptual Issues and Evidence from Yugoslavia. *Journal of Money, Credit and Banking*, 32(4), 785.
- [8] Garfield E., Sher I. (1993). KeyWords Plus™ Algorithmic Derivative Indexing. *Journal of The American Society for Information Science*, 44(5), 298–299.
- [9] Stanković J., Džunić M., Džunić Ž., Marinković S. (2017). Višekriterijska evaluacija pametnih performansi Europskih gradova: Gospodarski, socijalni i okolišni aspekti. *Zbornik Radova Ekonomskog Fakulteta u Rijeci*, 35(2), 519–550.
- [10] Arsić S., Nikolić D., Živković Z. (2017). Hybrid SWOT - ANP - FANP model for prioritization strategies of sustainable development of ecotourism in National Park Djerdap, Serbia. *Forest Policy and Economics*, 80, 11–26.
- [11] Mandić K., Delibašić B., Knežević S., Benković S. (2014). Analysis of the financial parameters of Serbian banks through the application of the fuzzy AHP and TOPSIS methods. *Economic Modelling*, 43, 30–37.
- [12] Lazarević D., Dobrodolac M., Švadlenka L., Stanivuković B. (2020). A model for business performance improvement: a case of the postal company. *Journal of Business Economics and Management*, 21(2), 564–592.
- [13] Macura D., Bošković B., Bojović N., Milenković M. (2011). A model for prioritization of rail infrastructure projects using ANP. *International Journal of Transport Economics*, 38(3), 285–309.
- [14] Mandić K., Delibašić B., Knežević S., Benković S. (2017). Analysis of the efficiency of insurance companies in Serbia using the fuzzy AHP and TOPSIS methods. *Economic Research-Ekonomska istraživanja*, 30(1), 550–565.
- [15] Stević Ž., Vasiljević M., Zavadskas E., Sremac S., Turskis Z. (2018). Selection of carpenter manufacturer using fuzzy EDAS method. *Engineering Economics*, 29(3), 281–290.
- [16] Bojković N., Anić I., Pejčić-Tarle S. (2010). One solution for cross-country transport-sustainability evaluation using a modified ELECTRE method. *Ecological Economics*, 69(5), 1176–1186.
- [17] Vasić G. (2018). Application of multi criteria analysis in the design of energy policy: Space and water heating in households – City Novi Sad, Serbia. *Energy Policy*, 113, 410–419.
- [18] Radulescu M., Fedajev A., Nikolić Đ. (2017). Ranking of EU national banking systems using multi-criteria analysis in the light of Brexit. *Acta Oeconomica*, 67, 473–509.
- [19] Tzeng G., Lin C., Opricović S. (2005). Multi-criteria analysis of alternative-fuel buses for public transportation. *Energy Policy*, 33(11), 1373–1383.
- [20] Stanujkić D., Karabašević D., Zavadskas E. (2015). A framework for the selection of a packaging design based on the SWARA method. *Engineering Economics*, 26(2), 181–187.
- [21] Sheth C., Triantis K., Teodorović D. (2005). Performance evaluation of bus routes: A provider and passenger perspective. *Transportation Research Part E: Logistics and Transportation Review*, 43(4), 453–478.
- [22] Radojičić M., Savić G., Jeremić V. (2018). Measuring the efficiency of banks: The bootstrapped i-distance GAR DEA approach. *Technological and Economic Development of Economy*, 24(4), 1581–1605.
- [23] Martinović N., Savić G. (2019). Staff Assignment to Multiple Projects Based on DEA Efficiency. *Engineering Economics*, 30(2), 163–172.
- [24] Cvetkoska V., Savić G. (2017). Efficiency of bank branches: empirical evidence from a two-phase research approach. *Economic Research-Ekonomska istraživanja*, 30(1), 318–333.
- [25] Pamučar D., Macura D., Tavana M., Božanić D., Knežević N. (2022). An integrated rough group

- multicriteria decision-making model for the ex-ante prioritization of infrastructure projects: The Serbian Railways case. *Socio-Economic Planning Sciences*, 79, 101098.
- [26] Stojadinović N., Bošković B., Trifunović D., Janković S. (2019). Train path congestion management: Using hybrid auctions for decentralized railway capacity allocation. *Transportation Research Part A: Policy and Practice*, 129, 123–139.
- [27] Macura D., Kapetanović M., Knežević N., Bojović N. (2019). Rail Projects Ranking under Fuzzy Environment : Serbian Rail Projects Case Study. *International journal of transport economics*, 46(4), 91–112.
- [28] Živkov D., Durašković J., Manić S. (2019). How do oil price changes affect inflation in Central and Eastern European countries? A wavelet-based Markov switching approach. *Baltic Journal of Economics*, 19(1), 84–104.
- [29] Sremac S., Zavadskas E., Matić B., Kopic M., Stević Ž., et al. (2019). Neuro-fuzzy inference systems approach to decision support system for economic order quantity. *Economic Research - Ekonomska istraživanja*, 32(1), 1114–1137.
- [30] Ulutaş A., Popović G., Radanov P., Stanujkić D., Karabasević D. (2021). A new hybrid fuzzy psi-piprecia-cocoso MCDM based approach to solving the transportation company selection problem. *Technological and Economic Development of Economy*, 27(5), 1227–1249.
- [31] Stanujkić D., Karabašević D., Popović G., Sava C. (2021). Simplified PIVot Pairwise RElative Criteria Importance Assessment (PIPRECIA-S) Method. *Romanian Journal of Economic Forecasting*, 24(4), 141–154.
- [32] Waltman L., van Eck N., Noyons E. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629–635.
- [33] Mitrović Đ. (2015) Broadband adoption, digital divide, and the global economic competitiveness of Western Balkan countries. *Economic Annals*, 60(207), 95–115.
- [34] Petrović M., Bojković N., Stamenković M. (2018). A DEA-based tool for tracking best practice exemplars: The case of telecommunications in EBRD countries. *Economic Annals*, 63(218), 105–127.
- [35] Garfield E. (1955). Citation indexes for science. *Science*, 122(3159), 108–111.



# ELEKTRONSKO POSLOVANJE

---

## ELECTRONIC BUSINESS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business







## ELEKTRONSKO POSLOVANJE U ŽELEZNIČKOM SAOBRAĆAJU

### ELECTRONIC BUSINESS IN RAILWAY TRAFFIC

ZORAN G. PAVLOVIĆ<sup>1</sup>

<sup>1</sup>Akademija tehničko umetničkih strukovnih studija, Beograd – Odsek Visoka železnička škola, [zoran.pavlovic@vzs.edu.rs](mailto:zoran.pavlovic@vzs.edu.rs), [zoran.g.pavlovic@gmail.com](mailto:zoran.g.pavlovic@gmail.com)

**Rezime:** Razvoj i primena inovativnih naprednih internet tehnologija je sve zastupljenija u transportnim preduzećima kao i u akademskoj zajednici. Predmet istraživanja obuhvata razvoj inovativnog modela elektronskog poslovanja u železničkom saobraćaju u skladu sa svojim raspoloživim novčanim sredstvima pokušavaju da sa što manje uloženog novca steknu veće prihode u cilju povećanja kvaliteta prevozne usluge.. Jedan od načina je primena inovativnih tehnologija koje mogu da funkcionišu na osnovu računarskih sistema u preduzeću. Korisnik usluge svoje potrebe može zadovoljiti preko internet mreže, uređajima i aparatima ako preduzeće u svojoj ponudi ima takvu vrstu usluge. Osnovni cilj rada je razvoj i povezanost komponenti elektronskog poslovanja u digitalizaciji poslovnih procesa železničkog saobraćaja. U ovom radu su predstavljene osnovne komponente elektronskog poslovanja u železničkom prevozu koje su zasnovane na naprednim internet tehnologijama. Komponente obuhvataju softverske i hardverske elemente kao i njihovu međusobnu povezanost. Na osnovu navedenih komponenti stvaraju se preduslovi za upotrebu B2C forme u cilju digitalizacije korisničkih procesa i povećanja kvaliteta prevozne usluge. Rezultat ovog rada je dobro strukturirana platforma koja obuhvata komponente elektronskog poslovanja.

**Ključne reči:** digitalizacija poslovnih procesa, B2C forma, komponente inovativnog modela, napredne internet tehnologije, kvalitet usluge železnice

**Abstract:** The development and application of innovative advanced Internet technologies is increasingly prevalent in transport companies as well as in the academic community. The subject of the research includes the development of an innovative model of electronic business in railway traffic, in accordance with their available financial resources, they are trying to get more income with as little invested money as possible in order to increase the quality of the transport service. One of the ways is the application of innovative technologies that can function on the basis of computer system in the company. The user of the service can satisfy his needs through the Internet, devices and appliances if the company offers such a service. The main goal of the work is the development and connection of electronic business components in the digitalization of railway traffic business processes. This paper presents the basic components of electronic business in railway transport, which are based on advanced Internet technologies. Components include software and hardware elements as well as their interconnection. On the basis of the mentioned components, prerequisites are created for the use of the B2C form in order to digitize user processes and increase the quality of the transport service. The result of this work is a well-structured platform that includes electronic business components.

**Keywords:** digitalization of business processes, B2C form, components of innovative model, advanced internet technologies, quality of railway service

#### 1. UVOD

Elektronsko poslovanje (EP) može da se definiše kao proces realizacije aktivnosti kupovine ili prodaje koja je organizovana u sistemu povezanih računarskih mreža. Autori [1] definišu elektronsko poslovanje kao transformaciju poslovnih procesa u organizacijama upotrebom Internet tehnologija.

Prednosti primene EP u odnosu na tradicionalne načine koji su u upotrebi manifestuju se kroz:

- Povećanje kvaliteta usluge;
- Mogućnost plasiranja dodatnih usluga kako bi se zadovoljile potrebe korisnika u što kraćem vremenskom periodu;
- Smanjenje vremena koje je potrebno preduzeću za prezentaciju usluge kroz implementaciju inovativnih digitalnih kanala distribucije;
- Snižavanje cena usluge;

- Analizu prikupljenih podataka koji se ostvaruju u interakciji korisnika i servisa provajdera usluge;
- Neprekidnu digitalnu komunikaciju između korisnika i servisa provajdera usluge;
- Realizaciju novčanih transakcija putem internet tehnologija itd.

Poslovni rezultati preduzeća tj. provajdera usluge mogu se ogledati i kroz smanjenje operativnih troškova koji nastaju u tradicionalnim poslovanjima kao i smanjenje složenosti kupovine u cilju povećanja kvaliteta [2]. Primena saobraćajog inženjeringa u inteligentnim transportnim sistemima može da reši i ublaži problem upravljanja tokovima putnika u više koraka sa sveobuhvatnim razmatranjem faktora iz vremenske, izvorno-odredišne prostorne i frekventne perspektive [3]. Povećana upotreba informacionih tehnologija dovela je do više istraživanja o usvajanju inovativnih tehnologija kao i ponašanja u donošenju odluka u malim preduzećima kako bi se poboljšala upotreba informacija [4].

Računarska mreža putem Interneta omogućava povezivanje milione računarskih uređaja. Ako se internet posmatra iz ugla usluga, može se opisati kao infrastruktura koja određenim aplikacijama obezbeđuje zahtevane usluge korisniku [5, 6, 7]. U visoko razvijenim privrednim društvima, elektronsko plaćanje javnog prevoza jedan je od odgovora za rešavanje problema sa gradskim saobraćajem, kao što su zagušenja i u velikoj upotrebi je zastupljeno korišćenje privatnih vozila [8].

Na osnovu navedenog i prikaza zanačaja EP u ovom radu biće prikazane polazne i osnovne odrednice računarskih tehnologija koje se odnose na digitalizaciju poslovnih procesa u železničkom saobraćaju. Forma EP koje je aktuelna između železničkog preduzeća i korisnika (B2C) unapred reguliše i omogućava realizaciju aktivnosti putem računarskih mreža i tehnologija. Analiza primene EP u železničkom saobraćaju prouzrokuje kreiranje održivih komponenti na osnovu potreba korisnika i prevoznika kao i tehnoloških inovativnih rešenja na osnovu infrastructure (hardver, softver i internet mreža).

## 2. OSNOVNE ODREDNICE ELEKTRONSKOG POSLOVANJA

Razvoj digitalne ekonomije kao i primena elektronskog poslovanja uslovljava komercilizacija Interneta početkom devedesetih godina. Tada nastaju nove mogućnosti za poslovanje od kojih su najvažnije:

- elektronska trgovina na globalnom tržištu,
- umrežavanje i integracija sa poslovnim partnerima,
- razvoj novih poslovnih modela itd.

Politička dešavanja u svetu i ekonomska kriza motiviše i animira poslovna preduzeća da pronalaze nove načine trgovinske razmene. Osnova je obuhvatala IKT inovacije. Primena novih modela i tehnologija je uticala na smanjenje troškova poslovanja u odnosu na tradicionalne modele koji su bili zastupljeni. Primena mobilnih tehnologija je omogućila poslovanje putem računarske mreže sa bilo kog mesta i bez vremenskog ograničenja pod uslovom da je Internet konekcija aktivna. Radi brzog informisanja i kolektivne potrošnje dolazi se do društvenog umrežavanja. Primena elektronskog poslovanja utiče na razvoj inovativnih poslovnih modela koji su zasnovani na Internetu. Razni poslovni modeli u prvi plan stavljaju korisnika usluge. Velike promene se dešavaju u bankarskom sektoru gde se kreiraju inovativne usluge kako bi se zadovoljile potrebe klijenata [1].

Autori [9] napominju da razvoj IKT i pojava Interneta stvara digitalno doba, gde je omogućeno odvijanje procesa elektronskog poslovanja u virtualnom-komputerskom projektovanom prostoru. Nove mogućnosti imaju za cilj da digitalno doba predstavi novu paradigmu savremenog poslovanja gde je omogućena:

- digitalizacija i povezanost,
- primena novih tehnologija,
- stvaranje uslova za primenu novih poslovnih sistema i
- plasiranje nove usluge ili proizvoda.

Digitalizacija i povezanost uređaja predstavlja proces funkcionisanja na osnovu digitalnih informacija. Za prenos bitova iz jednog računarskog sistema ili mesta u drugi potrebna je povezanost odnosno telekomunikaciona mreža. Internet je javna računarska mreža koja povezuje razne korisnike širom sveta međusobno gde se istovremeno velikom brzinom razmenjuju informacije. Nastanak interneta vrlo brzo iz običnog komunikacionog alata pretvara se u revolucionarnu tehnologiju.

Novi poslovni sistemi naveli su mnogobrojna preduzeća da preispitaju način na koji opslužuju svoja ciljna tržišta [10, 11]. Primenom interneta u svom poslovanju stvoreni su novi uslovi i mogućnosti za prodaju roba i usluga. Upotrebom Interneta preduzeća beleže veći rast prodaje i postaju uspešnije na konkurentskom tržištu.

Rukovodioci poslovanja su sada svesni da je informaciona tehnologija predstavlja ključni strateški faktor u industriji, a ne samo sredstvo za povećanje efikasnosti poslovanja kao što je uobičajeno. Ovi događaji doveli su do brzog porasta naučnog interesovanja i aktivnosti koje su vezane za e-biznis. S obzirom na brzinu razvoja poslovnih sistema i procesa postoje veliki izazovi i mogućnosti za primenu tehnologije i

metodologije u primeni inteligentnih sistema. Zapravo, e-biznis ima nekoliko različitih karakteristika koje čine istraživanje primene tehnologije inteligentnih sistema veoma atraktivnim [12].

### 3. PRIMENA B2C FORME ELEKTRONSKOG POSLOVANJA U ŽELEZNIČKOM SAOBRAĆAJU

Dinamički razvoj IKT posebno internet servisa utiče na mnoge tehnološke inovacije u svim privrednim granama. Da bi se realizovala razmena roba i usluga banke imaju u ponudi razne modele gde se postiže efikasno, jednostavno i racionalno izvršenje finansijskih transakcija putem računarske mreže. Elektronsko plaćanje predstavlja deo procesa EP gde je korisnik usluge u stalnoj internet komunikaciji sa bankom u procesu razmenjivanja novčanih vrednosti između prodavca i kupca za prodatu robu ili pruženu uslugu. Za razliku od tradicionalnog modela plaćanja koje zahteva fizički odlazak korisnika u prostorije pružaoca usluge, elektronsko plaćanje podrazumeva transakcije koje se odvijaju digitalno putem računarske mreže pomoću platnih kartica, elektronskih čekova, mobilnih sistema plaćanja, digitalnih novčanika itd [13].

IKT imaju sve veću primenu u železničkom saobraćaju. Akcenat je stavljen na tehnologije mobilnih komunikacija gde se podrazumeva upotreba mobilnih uređaja u procesu poslovanja. Sa aspekta železničkih preduzeća potreba za implementacijom mobilnog poslovanja, korisnicima usluge pruža novi pristup i upotrebu postojećih servisa, gde se ostvaruju bolji rezultati poslovanja. Mobilno poslovanje u železničkom saobraćaju odvija se preko interneta i servisa računarskih mreža i omogućava [14]:

- Sveobuhvatnost;
- Pogodnost;
- Instant povezivanje;
- Personalizacija i lokacija.

Primenom mobilnih tehnologija korisnik usluge može da zadovolji svoje potrebe koje se ogledaju u nabavci karte. Korisnik usluge ima mogućnost da putem mobilnog uređaja slanjem poruke u mobilnoj aplikaciji naruči, plati, dobije i izvrši validaciju karte u bilo koje vreme sa bilo koje lokacije.

#### 3.1. Primena elektronskog poslovanja u železničkom saobraćaju

Za svoje lične zahteve korisnici usluge koriste Internet mogućnosti kako bi zadovoljili svoje potrebe. Internet usluge mogu da koriste putem mobilnih aparata, tableta i desktop računara. Za složenije procese moraju da imaju povezan štampač sa računarem, otvoren tekući račun u banci i platnu karticu. Platna kartica korisnika usluge omogućava transakcije u digitalnom obliku koji se koristi prilikom elektronske trgovine.

Putem interneta korisnici usluge mogu preko računara da vide razne informacije na sajtu prevoznika (red vožnje, kašnjenje voza, cenu karte na pojedinim relacijama, usputne stanice, izmene u saobraćaju itd.) ako poseduju štampač i validnu platnu karticu da naruče, plate i dobiju odštampanu potvrdu o voznoj ispravi. Model predstavlja proces u interakciji između korisnika i provajdera usluge, gde se upotrebljava internet i pametni uređaj u komunikaciji [15].

Veb stvara odlične mogućnosti za preduzeća da pruže personalizovane on-line usluge svojim klijentima. Preporučeni sistemi imaju za cilj da automatski generišu personalizovane predloge proizvoda ili usluga za kupce (preduzeća ili pojedince). Iako su sistemi preporuka dobro proučeni, još uvek postoje dva izazova u razvoju sistema za preporuke, naročito u stvarnim B2C e-uslugama:

- stavke ili korisničke profile često predstavljaju komplikovane strukture u poslovnim aplikacijama, što ne može biti i koje se obrađuju standardnim merama sličnosti i
- preferencije on-line korisnika često su nejasne i ne mogu se rešavati postojećim metodama preporučivanja.

Da bi se suočili sa oba ova izazova, autori [16] predlažu metodu za modeliranje strukturiranih korisničkih preferenci korisnika. Eksperimentalni rezultati pokazuju da predloženi korisnički preferencijalni profil odražava efekte korisničkih preferenci i preporuka pristupa pokazuje odlične performanse za strukturirane stavke, naročito u e-poslovnim aplikacijama.

Jedan od najvećih izazova u modelu EP je izdavanje i kupovina prigradskih železničkih karata. U ovom brzom rastućem svetu tehnologije, dalje stojimo u redu, kupujemo sa platnim karticama prigradske karte, što može biti frustrirajuće ponekad da stojimo u redu ili ako zaboravimo naše kartice. Ovaj model odnosi se na kupovinu prigradskih karata koji su najizazovnij u poređenju sa rezervisanjem pretplatnih karata putem "M-tiketa" (mobilnih karata) koja se ne ostvaruje sa prigradskim (lokalnim putovanjima) kartama [17].

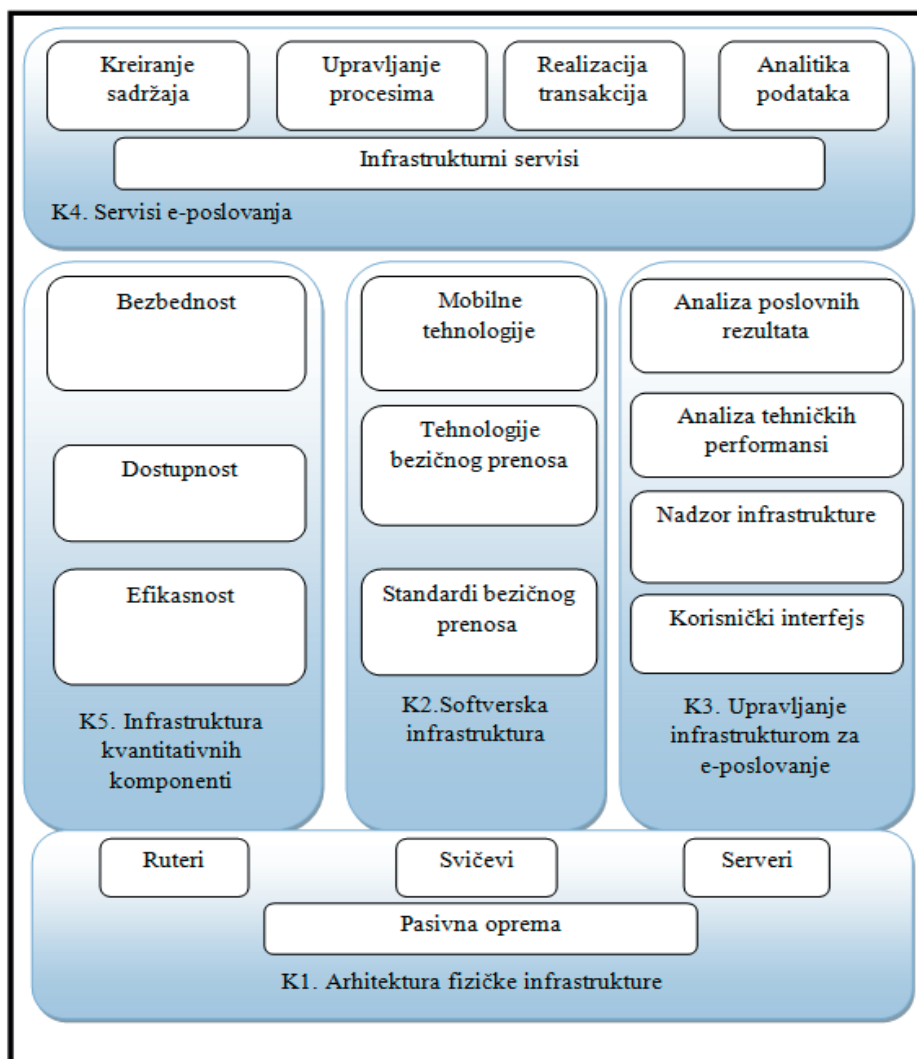
Model otvorenog pristupa internetu zasnovan na režimu centralizovanog pristupa, dizajniran je za mobilno poslovanje u vozu gde je spojen sa bežičnim pristupom i mrežnim tehnologijama. Funkcionalnosti i entiteti u svakom sloju istražuju i razmatraju tekuća pitanja prema ovom modelu. Pokazano je da se na Internetu, uključujući i mobilno poslovanje u vozu, mogu ostvariti i zadovoljiti potrebe korisnika [18].

Obim standarda EP danas prevazilazi nivo infrastrukture i uključuje šira pitanja vezana za poslovne procese. Standardi EP se sve više razvijaju putem otvorenih inicijativa, kako bi se suprotstavili formalnim standardizacijskim naporima. Za standarde usmerene na poslovne procese interesna neizvesnost prouzrokuje proces ponavljanja i usredređivanja na preciziranje zahteva. Standardi fokusirani na tehnologiju pokazuju viši nivo strukture i imaju tendenciju nižih nivoa iteracije koji su fokusirani na dizajn. Ovi nalazi omogućavaju usaglašavanje razvoja standarda i strateškog donošenja odluka, odlazak od tipične "razvojne" orijentacije korišćene u standardnoj literaturi. Nalaz naročito se drži standarda koji se odnose na poslovne procese. Otkriva se značaj "otvorenosti" kao doprinos interakcijama dokazanim u procesnim obrascima vezanim za razvoj standarda. Priznajući da je razvoj standarda složen socijalni proces studija koja se fokusira na proces aspekte fenomena. Standarde bi trebalo prilagoditi promenljivim potrebama konteksta razvoja standarda da bi se maksimalno povećala efikasnost procesa. Različiti konteksti razvoja standarda će verovatno zahtevati drugačiji pristupi za različite faze projekta [19].

### 3.2. Komponente elektronskog poslovanja u železničkom saobraćaju

Na osnovu navedenog potrebno je razviti i predstaviti komponente inovativnog transakcionog modela EP zasnovanog na naprednim Internet tehnologijama u železničkom saobraćaju. Komponente treba da obuhvate infrastrukturu koju čini hardver i softver koji ima ulogu da poveže računar i korisnika i da povezuje uređaje i komunikacione kanale za prenos podataka putem internet mreže.

Na slika 1. je prikazana detaljna struktura navedenih komponenti, kao i njihove međusobne veze.



Slika 1: Komponente modela EP

Predloženi inovativni model za realizaciju procesa EP zasnovanog na naprednim Internet tehnologijama u železničkom saobraćaju obuhvata sledeće komponente:

- K1. Arhitektura fizičke infrastrukture;
- K2. Softverska infrastruktura;
- K3. Upravljanje infrastrukturom za e-poslovanje;
- K4. Servisi e-poslovanja;
- K5. Infrastruktura kvantitativnih komponenti

U osnovi infrastruktura inovativnog modela treba da omogućiti na osnovu tehnologije (primenjene za razvoj modela), okoline (prostora za realizaciju procesa), organizacionih procesa (unapred predviđenim poslovnim pravilima) i ljudskih resursa (u železničkom preduzeću zaduženi za sigurnost, bezbednost i dostupnost resursa servisa) da usluga (u ovom slučaju e-karta u digitalnom zapisu) neprekidno upravljanje poslovnim digitalnim procesima.

Pored dizajniranja infrastrukture inovativnog modela potrebno je izvršiti i funkcionalno modeliranje poslovnih procesa za pružanje usluge [20].

## 5. ZAKLJUČAK

U ovom radu prikazane su osnove EP kao i komponente koje su neophodne za realizaciju digitalizovanih poslovnih procesa na primeru železničkog preduzeća u segmentu prevoza korisnika usluga. Inovativne tehnologije koje su zasnovane na naprednim internet tehnologijama zahtevaju istraživački pristup sa akcentom na samu upotrebu.

U ovom radu prikazane su i analizirane komponente čija je osnovna namena povezivanje korisnika i železničkog preduzeća. Navedene tehnologije u komponentama obezbeđuju bezbednost novčanih transakcija, dostupnost servisa 7/24 i efikasnost koja se odnosi na obostrano zadovoljstvo zainteresovanih strana. Na ovaj način transportno preduzeće sagledava i analizira potrebe korisnika usluge i na osnovu dobijenih rezultata pristupa modeliranju inovativnih modela. U radu su prikazane osnovne komponente inovativnog modela EP, gde je obuhvaćena arhitektura fizičke infrastrukture, softverska infrastruktura, sistemom za upravljanje infrastrukturom za EP, infrastrukturom kvantitativnih komponenti kao i servisima za realizaciju digitalnih procesa. Ujedno se kroz inovativni pristup EP proširuju mogućnosti železničkog preduzeća prvenstveno u poboljšanja kvaliteta usluge, ostvarivanju boljeg imidža kao i povećanje transportnih prihoda.

Potrebe železničkog preduzeća u cilju povećanja kvaliteta prevozne usluge su realne i ostvarljive primenom naprednih tehnologija EP. Polazne osnove u ovom radu mogu da se upotrebe u drugim inovativnim modelima sa sličnom tematikom.

## LITERATURA

- [1] Radenković, B., Despotović Zrakić, M., Bogdanović, Z., Barać, D., & Labus, A. (2015). *Elektronsko poslovanje* (1 ed.). Beograd: Fakultet organizacionih nauka Beograd.
- [2] Candeia, D., Santos, R. A., & Lopes, R. (2015). Business-Driven Long-Term Capacity Planning for SaaS Applications. *IEEE Transactions on Cloud Computing ( Volume: 3 , Issue: 3 , July-Sept. 1 2015 )*, DOI: 10.1109/TCC.2015.2424877 , 290-303.
- [3] Diao, Z., Zhang, D., Wang, X., Xie, K., He, S., Lu, X., et al. (2019). A Hybrid Model for Short-Term Traffic Volume Prediction in Massive Transportation Systems. *IEEE Transactions on Intelligent Transportation Systems ( Volume: 20 , Issue: 3 , March 2019 )*, DOI: 10.1109/TITS.2018.2841800 , 935-946.
- [4] Yang, X., & Fu, J. (2008). Review of IT/IS adoption and decision-making behavior in small businesses. *Tsinghua Science and Technology ( Volume: 13 , Issue: 3 , June 2008 )*, DOI: 10.1016/S1007-0214(08)70052-X , 323-328.
- [5] Z. G. Pavlović, "Development Of Models Of Smart Intersections In Urban Areas Based On IoT Technologies," *2022 21st International Symposium INFOTEH-JAHORINA (INFOTEH)*, 2022, pp. 1-4, doi: 10.1109/INFOTEH53737.2022.9751263.
- [6] Z. G. Pavlović, "Technologies of electronic business in traffic," *2022 21st International Symposium INFOTEH-JAHORINA (INFOTEH)*, 2022, pp. 1-4, doi: 10.1109/INFOTEH53737.2022.9751297.
- [7] Z. G. Pavlović, Z. Bundalo, M. Bursać and G. Tričković, "Use of information technologies in railway transport," *2021 20th International Symposium INFOTEH-JAHORINA (INFOTEH)*, East Sarajevo, Bosnia and Herzegovina, 2021, pp. 1-4, doi: 10.1109/INFOTEH51037.2021.9400521,

- [8] Pavlović Z, Banjanin M, Vukmirović J, Vukmirović D., (2020,05,04): *Contactless ICT Transaction Model Of The Urban Transport Service*; Research journal TRANSPORT, ISSN: 1648-4142 / eISSN: 1648-3480, Vol 35 No 5, pp 500-510.
- [9] Backović, N., Radenković, S., Đelošević, I., & Novičić, M. (2009). *Elektronsko poslovanje i Internet marketing*. Leposavić: Visoka ekonomska škola strukovnih studija Peć u Leposaviću.
- [10] Zoran G. Pavlović (2021): *Stav zaposlenih prema inovativnom IKT modelu elektronskog poslovanja*, Svet rada, Beograd, VOL 18 br. 2/2021, pp 181-190
- [11] Zoran Pavlović, Zoran Bundalo, Marko Bursać, Goran Tričković, 2021: *Analiza primene elektronskog poslovanja u železničkom saobraćaju*, YUINFO 2021, 27. naučna i biznis konferencija, 07-10. mart 2021, Kopaonik, Srbija, pp.110-114,
- [12] H. Akkermans. (2001). Intelligent e-business: from technology to value., *IEEE Intelligent Systems, Volume: 16, Issue: 4, DOI: 10.1109/5254.941352* , 8-10.
- [13] Drašković, M., Kukrić, M., & Smiljić, S. (2015). Virtuelne valute, savremeni sistemi elektronskog plaćanja roba i usluga. *SYM-OP-IS 2015: XLII International Symposium on Operations Research, 2015*, (pp. 40-43).
- [14] Vuletić P, Bojić Ž. (2015). Inteligentni uređaji i mobilne komunikacije u poslovanju železnice. *SYM-OP-IS 2015: XLII Simpozijum o operacionim istraživanjima, 2015*, (pp. 28-31).
- [15] Pavlovic, Z., & Vukmirović, A. (2016). Special offer for railwaysticket issue reserved and bought over internet., *YUINFO 2016 XXII naučna i biznis konferencija 28. februar-02. mart 2016* , Kopaonik, Srbija, ISBN978-86-85525-17-9,, (pp. 226-231).
- [16] Wu, D., Zhang, G., & Jie Lu. (2015). A Fuzzy Preference Tree-Based Recommender System for Personalized Business-to-Business E-Services, . *IEEE Transactions on Fuzzy Systems* , Volume: 23, Issue: 1, DOI: 10.1109/TFUZZ.2014.2315 , 29-43.
- [17] Karthick, S., & Velmurugan, A. (2012). Android suburban railway ticketing with GPS as ticket checker., *Advanced Communication Control and Computing Technologies (ICACCCT)*, (pp. 63-66). IEEE International Conference onYear: 2012.
- [18] Fumin, Z., & Shuling, Z. (2009). "An Open Onboard Internet System Model for the M-Commerce on Train":. *MASS '09. International Conferenceon: 2009* (pp. 1-5). Management and Service Science.
- [19] Choi, B., Raghu, T. S., Vinze, A., & Dooley, K. J. (2009). Process Model for e Business Standards Development: A Case of ebXML Standards. *IEEE Transactions on Engineering Management, Volume: 56, Issue: 3* , 448 - 467, DOI: 10.1109/TEM.2009.
- [20] Graule, A. O., Erochina, Z. N., Maiboroda, V. P., & Mizginova, M. A. (2016). Approaches to IT infrastructure modelling of electronic university. *2016 IEEE Conference on Quality Management, Transport and Information Security, Information Technologies (IT&MQ&IS)*, DOI: 10.1109/ITMQIS.2016.7751900. Nalchik, Russia: IEEE.



# FINANSIJE I BANKARSTVO

---

## FINANCE AND BANKING



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business







## ANALIZA FINANSIRANJA IZDATAKA UNIVERZITETA ODBRANE ZA NABAVKU ARTIKALA ISHRANE

### ANALYSIS OF THE UNIVERSITY OF DEFENCE FOOD ITEMS PURCHASE EXPENDITURE FINANCING

MILENA KNEŽEVIĆ<sup>1</sup>, ALEKSANDRA MITIĆ<sup>2</sup>

<sup>1</sup> Univerzitet odbrane u Beogradu – Vojna akademija, [milena.knezevic@mod.gov.rs](mailto:milena.knezevic@mod.gov.rs)

<sup>2</sup> Univerzitet odbrane u Beogradu – Vojna akademija, [anamiticle@gmail.com](mailto:anamiticle@gmail.com)

**Apstrakt:** U radu je ukazano na specifičnost finansiranja izdataka Univerziteta odbrane kao organizacionog dela Ministarstva odbrane Republike Srbije. Analizirano je učešće izdvajanja za potrebe odbrane Republike Srbije u periodu od 2010. do 2020. godine.

Cilj rada je da ukaže na značaj obezbeđenja finansijskih sredstava za potrebe nabavki artikala ishrane na Univerzitetu odbrane u Beogradu, a u krajnjem radi obezbeđenja pravilne ishrane učenika i kadeta i ostalih slušalaca.

Pored opštih naučnih metoda, s obzirom na predmet i cilj istraživanja, težišno je korišćena komparativna metoda kojom je analizirano finansiranje izdataka za nabavku artikala ishrane Univerziteta odbrane u periodu od 2010. do 2020. godine.

**Ključne reči:** Univerzitet odbrane, izdaci, artikli ishrane, pravilna ishrana

**Abstract:** The paper points out the specificity of financing the expenditures of the University of Defence as an organizational part of the Ministry of Defence of the Republic of Serbia. The participation of allocations for the needs of the defense of the Republic of Serbia in the period from 2010 to 2020, was analyzed.

The aim of this paper is to point out the importance of providing financial resources for the purchase of food items at the University of Defence in Belgrade, and ultimately to ensure proper nutrition of pupils, cadets and other students.

In addition to general scientific methods, considering the subject and goal of the research, a comparative method was mainly used, which analyzed the financing of expenditures for the purchase of food items of the University of Defence in the period from 2010 to 2020.

**Keywords:** University of Defence, expenses, food items, proper nutrition

## 1. UVOD

Univerzitet odbrane u Beogradu školuje prevashodno kadete-buduće oficire za potrebe Vojske Srbije ali i stranih oružanih sila. Obzirom da se radi o mladim ljudima i način ishrane mora biti u skladu sa njihovim potrebama a u svrhu održavanja njihovog zdravlja. U skladu sa tim, ishrana na Univerzitetu odbrane je organizovana po svim principima «pravilne ishrane». Jer, pravilna ishrana predstavlja veoma važan činilac za održavanje zdravlja, radne sposobnosti i borbene gotovosti. Ona veoma povoljno utiče na stvaranje i podizanje otpornosti organizma prema bolestima i usporavanje procesa starenja. Pravilnom ishranom mogu se izbeći bolesti koje prate nepravilnu ishranu i ublažiti bolesti degenerativne prirode [1].

Međutim, finansijska sredstva za potrebe nabavke artikala ishrane su ograničena jer je Univerzitet odbrane u Beogradu organizaciono u sastavu Ministarstva odbrane Republike Srbije. Ministarstvo odbrane predstavlja direktnog korisnika budžetskih sredstava i u skladu sa Zakonom o budžetskom sistemu finansira se sredstvima Republike Srbije raspoređenim u okviru budžeta, po namenama i izvorima finansiranja za svaku godinu.

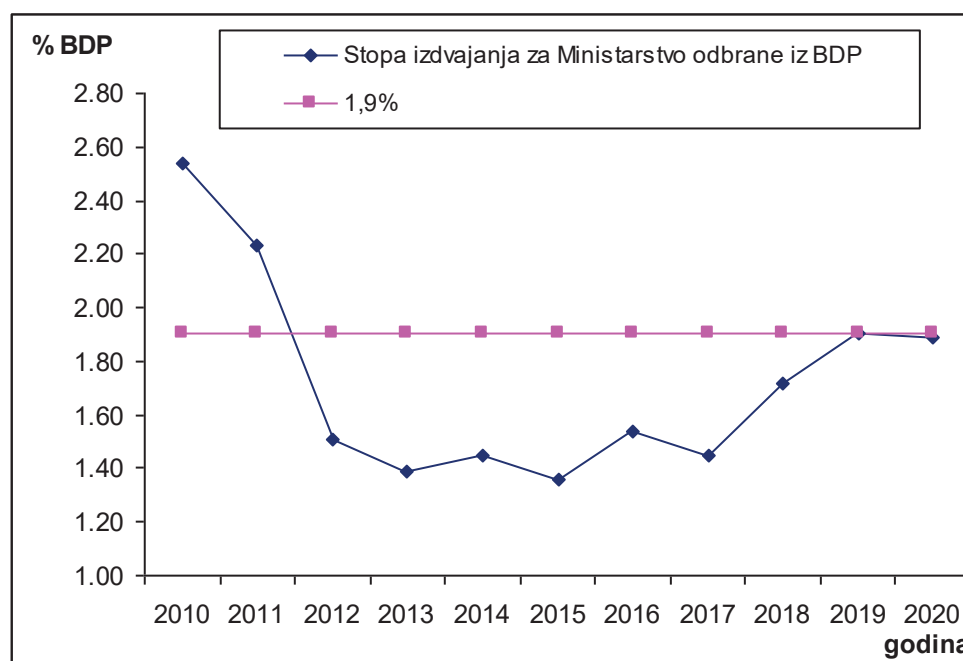
Finansijska sredstva koja se Ministarstvu odbrane opredele za nabavku artikala ishrane predstavljaju jednu poziciju u okviru ukupno odobrenih sredstava u okviru Zakona o budžetu Republike Srbije za svaku godinu.

Od tako opredeljenih sredstava jedan deo se izdvaja za finansiranje centralizovane nabavke artikala ishrane za sve jedinice Vojske Srbije a deo se opredeljuje jedinicama da nabavljaju artikle ishrane svaka za sebe pojedinačno. Univerzitet odbrane u tom smislu predstavlja samo jednu od jedinica i u skladu sa tim ima na raspolaganju finansijska sredstva za tu namenu, s tim da deo artikala ishrane isporučuje centralizovano iz Centralnog skladišta Vojske Srbije. Iako su finansijska sredstva limitirana u budžetu Republike Srbije, ishrana na Univerzitetu odbrane je organizovana po svim principima pravilne ishrane.

## 2. FINANSIRANJE IZDATAKA MINISTARSTVA ODBRANE

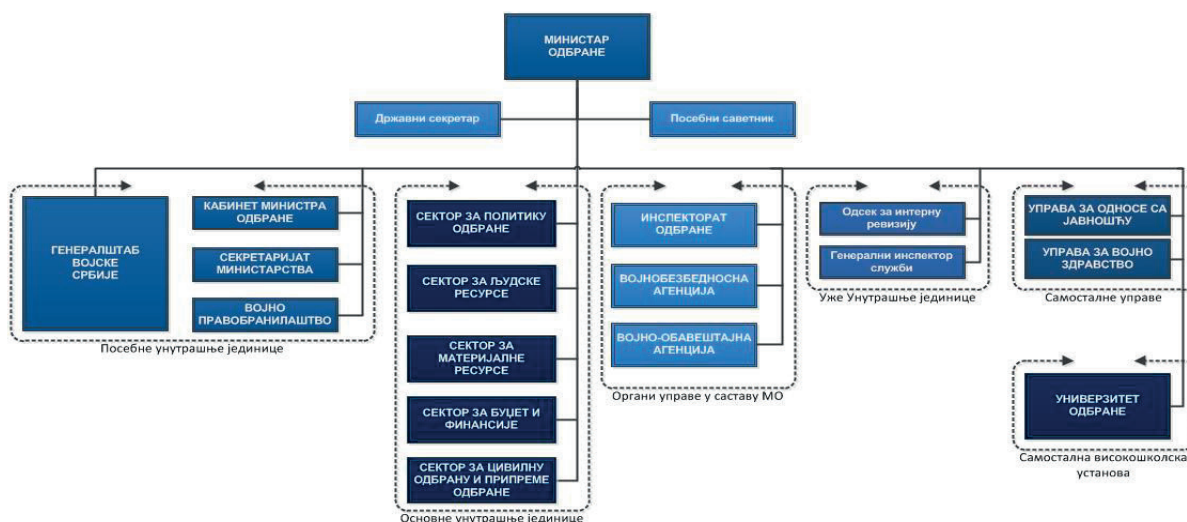
Ministarstvo odbrane Republike Srbije, kao direktni korisnik budžetskih sredstava finansira se u skladu sa Zakonom o budžetskom sistemu Republike Srbije. Sredstva opredeljena Ministarstvu odbrane Zakonom o budžetu Republike Srbije se namenski i krajnje racionalno troše u skladu sa važećim propisima u Republici Srbiji i Vojski Srbije. Država ulaže maksimalne napore da u okviru realnih mogućnosti opredeli finansijska sredstva Ministarstvu odbrane u skladu sa potrebama, za održavanje i razvoj novih sposobnosti Vojske Srbije. Održavanje i razvoj sposobnosti podrazumeva isplate za ljudstvo, opremanje, izdatke za infrastrukturu i sve neophodne trošove u skladu sa zahtevima. Međutim, razvoj ljudstva, odnosno kadra, uglavnom se realizuje na Univerzitetu odbrane u Beogradu, dok se ostale aktivnosti realizuju u ostalim jedinicama Vojske Srbije.

Za potrebe finansiranja izdataka i rashoda Ministarstva odbrane, država je u periodu od 2010. godine do danas izdvajala sredstva u skladu sa svojim mogućnostima. Dugoročnim planom razvoja sistema odbrane Republike Srbije u posmatranom periodu predviđeno je izdvajanje za Ministarstvo odbrane od 1,9% BDP. Prema Grafikonu 1, stopa izdvajanja za potrebe odbrane bila je najveća u 2010. godini od kada je beležila značajan pad do 2013. godine od kada se ta izdvajanja kreću oko 1,4% BDP do 2017. godine. Od 2017. godine država povećava procenat izdvajanja za potrebe odbrane na približno 1,9%BDP.



**Grafikon 1:** Finansiranje Ministarstva odbrane

Ministarstvo odbrane u svom sastavu ima Generalštab Vojske Srbije, Kabinet ministra odbrane, Sekreterijat Ministarstva odbrane, Vojno pravobranilaštvo, Sektor za politiku odbrane, Sektor za ljudske resurse, Sektor za materijalne resurse, Sektor za budžet i finansije, Sektor za civilnu odbranu i pripreme odbrane, Inspektorat odbrane, Vojno bezbednosnu agenciju, Vojnoobaveštajnu agenciju, Odsek interne revizije, samostalne uprave i Univerzitet odbrane, u skladu sa Slikom 1.



Slika 1. Organizacija Ministarstva odbrane Republike Srbije

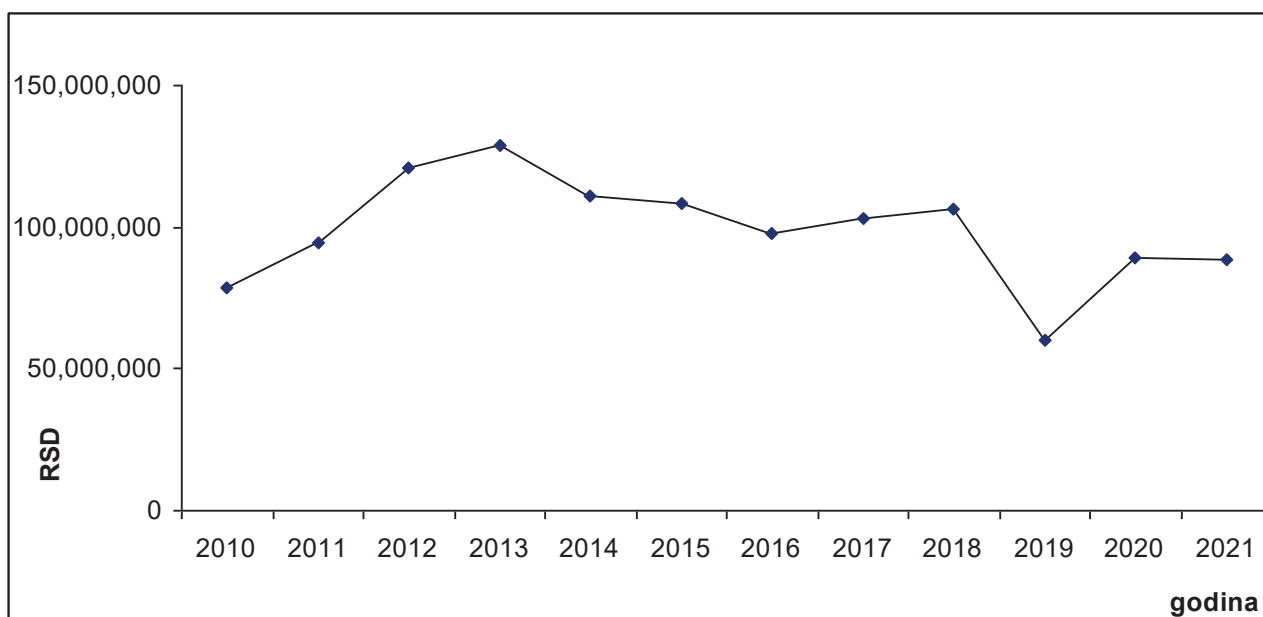
Znači, Univerzitet odbrane predstavlja samo deo Ministarstva odbrane i u skladu sa odobrenim misijama i zadacima opredeljuju mu se finansijska sredstva od ukupno odobrenih sredstava Ministarstvu odbrane za svaku budžetsku godinu, po svim namenama izuzev zarada koje se centralizovano isplaćuju na nivou Ministarstva odbrane za sva lica u MO i VS.

### 3. FINANSIRANJE IZDATAKA UNIVERZITETA ODBRANE ZA NABAVKU ARTIKALA ISHRANE

Prema većini istraživača, osnovni principi pravilne ishrane su: redovnost obroka u toku dana, raznovrsnost u izboru namirnica, odgovarajuća zastupljenost namirnica i pravilan način pripreme hrane. Na Univerzitetu odbrane u Beogradu se strogo vodi računa o poštovanju svih navedenih principa pravilne ishrane.

Kao što je navedeno u uvodnom delu rada, Univerzitet odbrane u Beogradu deo namirnica za ishranu kadeta i ostalih slušalaca dobija iz centralizovane nabavke Vojske Srbije, radi efikasnije i ekonomičnije upotrebe odobrenih finansijskih sredstava Ministarstva odbrane. Međutim, deo artikala ishrane Univerzitet odbrane samostalno nabavlja u okviru raspoloživih finansijskih sredstava, kao vansklađišne artikle.

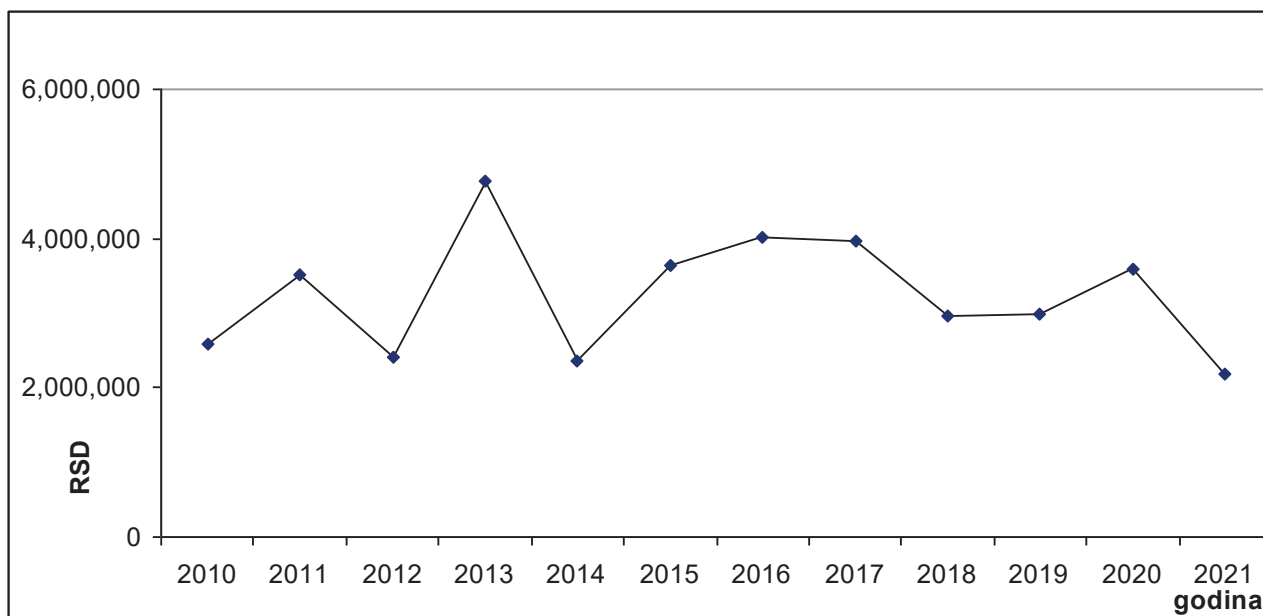
U periodu od 2010. godine do 2021. godine (Grafikon 2) Univerzitet odbrane je za nabavku vansklađišnih artikala trošio raspoloživa finansijska sredstva namenski, ekonomično i u skladu sa važećim propisima. U posmatranom periodu primetan je rast utroška finansijskih sredstava od 2010. godine do 2013. godine, pa neznatan pad do 2016. godine. Interesantna je činjenica da je zabeležen najveći pad utroška finansijskih sredstava 2020. godine, kada je proglašena pandemija COVID-19, jer, kadeti nisu bili na Univerzitetu i nije organizovana ishrana za njih.



**Grafikon 2:** Vanskladišni artikli ishrane

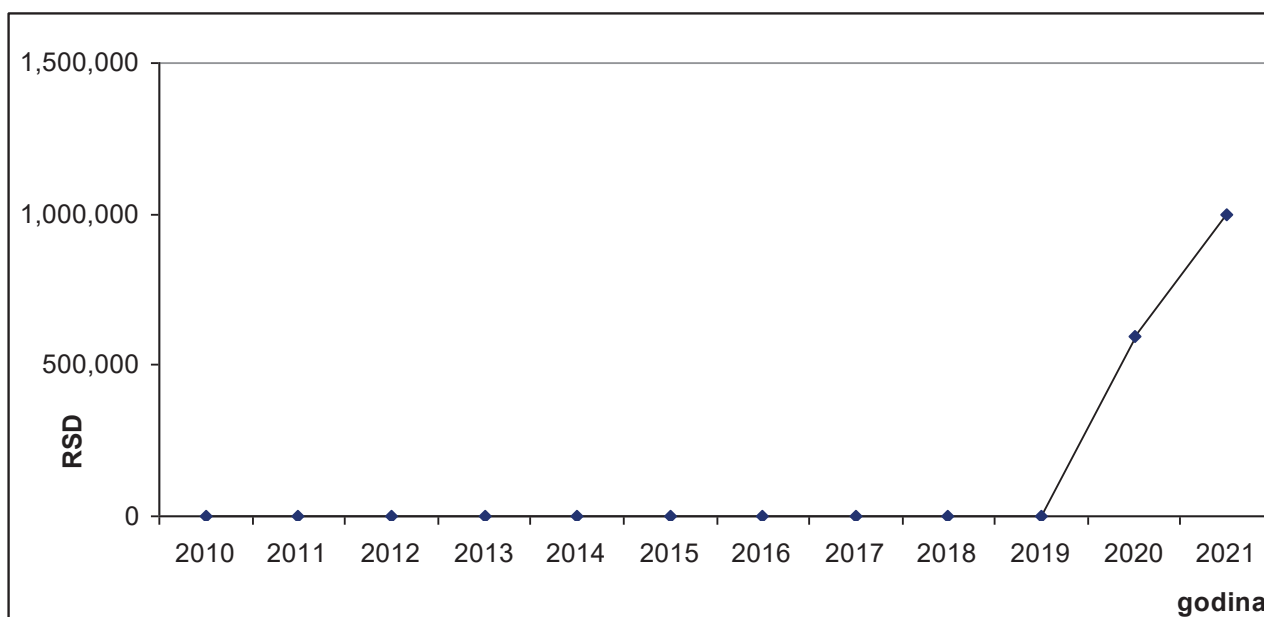
Na osnovu podataka, iz Grafikona 2, rade se predviđanja i planiranje utroška finansijskih sredstava za nabavku artikala ishrane na Univerzitetu odbrane u Beogradu za naredni period. Naravno kao veoma važan podatak uzima se brojno stanje i rast cena na malo.

Pored nabavke vanskladišnih artikala na Univerzitetu odbrane u Beogradu, a radi obezbeđenja pravilne ishrane, realizuje se nabavka polu obroka u ishrani. Na Grafikonu 3, dat je prikaz utroška finansijskih sredstava Univerziteta odbrane za nabavku polu obroka ishrane u periodu od 2010. godine do 2021. godine.



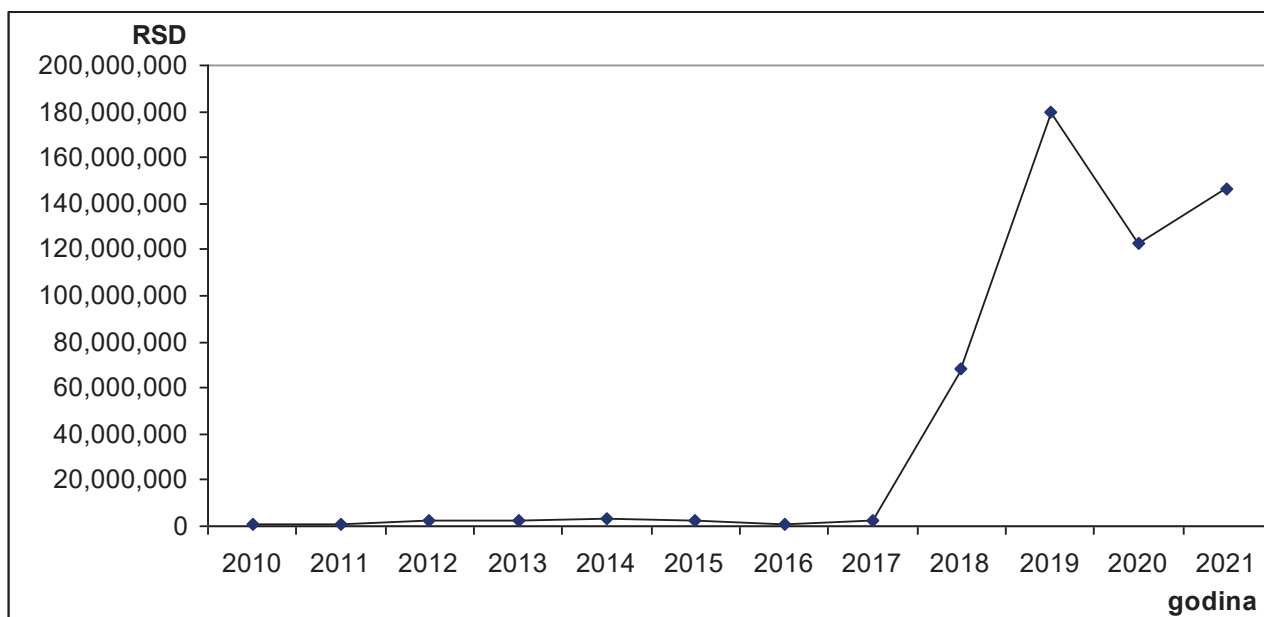
**Grafikon 3:** Polu obroci u hrani

Od 2020. godine u ishrane kadeta Univerziteta odbrane u Beogradu uvedene su žitne pahuljice. Na Grafikonu 4, dat je prikaz utroška finansijskih sredstava Univerziteta odbrane za nabavku žitnih pahuljica u posmatranom periodu.



Grafikon 4: Žitne pahuljice

Na Univerzitetu odbrane u Beogradu, pored kadeta školuju se i usavršavaju na različitim oblicima školovanja i usavršavanja, učenici u Vojnoj gimnaziji i Srednjoj stručnoj školi, polaznici Komandno štabnog i General štabnog usavršavanja kao i studenti na osnovnim, master i doktorskim studijama. Međutim, ishrana se prevashodno obezbeđuje za učenike i kadete, odnosno sve one za koje je i obezbeđen i internatski smeštaj u kasarni.



Grafikon 5: Hrana VU DEDINJE

Obzirom da su učenici Vojne gimnazije i Srednje stručne škole lokacijski smešteni u drugoj kasarni za njih je ishrana obezbeđena u celosti od 2017. godine preko VU Dedinje. Do tada je VU Dedinje samo za pojedinačne slučajeve obezbeđivao ishranu na Univerzitetu odbrane i kao što je i prikazano na Grafikonu 5, utrošak finansijskih sredstava je minimalan za t namenu do 2017. godine od kada se beleži rast sve do poerioda pandemije COVID-19, kada su učenici bili pušteni kućama, da prate nastavu *on line*.

## 5. ZAKLJUČAK

Na Univerzitetu odbrane u Beogradu se organizuje ishrana učenika, kadeta i dugih slušalaca. Obzirom da se radi o mladim ljudima, posebna pažnja se posvećuje kvalitetu i bezbednosti ishrane i to po svim principima pravilne ishrane, na osnovu naučnih normi. Odabir i nabavka artikala ishrane je strogo

kontrolisana. Obezbeđenje finansijskih sredstava za datu namenu predstavlja imperativ kako menadžmenta Univerziteta odbrane tako i upravljačkih struktura Ministarstva odbrane.

Na kraju zaključujemo, da je finansiranje izdataka Univerziteta odbrane za nabavku artikala ishrane iako u okviru limitiranih sredstava Ministarstva odbrane uvek u skladu sa potrebama učenika i kadeta i ostalih slušalaca i to po svim principima pravilne ishrane.

## LITERATURA

- [1] Stojanović, D. (2012). *Higijena sa medicinskom ekologijom*, praktikum za studente medicine, Niš, 31-45.
- [2] Rađen, S., (2012). *Ishrana*, Uloga u unapređenju zdravlja i prevenciji bolesti, Beograd
- [3] Zakon o budžetskom sistemu Republike Srbije, Sl.glasnik RS, broj 54/2009, 73/2010, 101/2010, 101/2011, 93/2012,62/2013, 63/2013, 108/2013, 142/2014, 68/2015, 103/2015, 99/2016, 113/2017, 95/2018, 31/2019, 72/2019, 149/2020 i 118/2021
- [4] [www.mod.gov.rs](http://www.mod.gov.rs)
- [5] [www.stat.gov.rs](http://www.stat.gov.rs)

## CAN A DYNAMIC CORRELATION FACTOR IMPROVE THE PRICING OF INDUSTRY PORTFOLIOS?

MILOŠ BOŽOVIĆ<sup>1</sup>

<sup>1</sup> University of Belgrade – Faculty of Economics and Business, milos.bozovic@ekof.bg.ac.rs

---

**Abstract:** *One of the notable failures of the asset pricing models is their inability to capture the behavior of 49 Fama-French industry portfolios. Since the Fama-French factors are designed to describe the most prominent patterns, such as size, value, profitability and investment, we investigate whether a common trait shared by several anomalies is better suited to characterize the average returns of industry portfolios. To achieve this goal, we include a WMS factor exploiting the spread between the weaker and the stronger correlations with the market in the standard asset pricing models. We show that it reduces the GRS statistics of these models and improves the description of the industry returns. The performance is additionally improved in periods of systemic crises such as COVID-19.*

**Keywords:** *Asset pricing, Factor models, Excess returns, Industry portfolios, Dynamic correlations*





## EFEKAT UKLJUČIVANJA INFLACIJE PRI OBRAČUNU REZERVACIJA ZA ŠTETE EFFECT OF INCLUSION OF INFLATION IN THE LOSS RESERVES CALCULATION

JELENA KOČOVIĆ<sup>1</sup>, MARIJA KOPRIVICA<sup>2</sup>, ŽELJKO JOVIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, jelena.kocovic@ekof.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, marija.koprivica@ekof.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Ekonomski fakultet, zeljko.jovic@ekof.bg.ac.rs

**Rezime:** Predmet istraživanja ovog rada je uticaj inflacije na rezervacije za štete u osiguranju od autoodgovornosti, a posledično i na mogućnost realnog obeštećenja osiguranika. Cilj rada je da se ukaže na neophodnost uzimanja u obzir inflacije prilikom obračuna rezervacija za štete i da se prikaže na koji način se inflacija može uključiti u metode njihovog obračuna. Po svojoj prirodi, osiguranje je izloženo inflaciji više nego drugi sektori, imajući u vidu da između naplate premije osiguranja i realizacije štetnog događaja koji ima za posledicu isplatu naknade štete prođe dovoljno dug vremenski period da može da se ispolji štetno dejstvo inflacije na realnost obeštećenja osiguranika. Pri tome, osiguravajuća kompanija ima zakonsku obavezu da formira rezervacije za štete, a njihov obračun je zadatak i odgovornost aktuaru. Adekvatnost rezervacija za štete je važna za solventnost osiguravača, a posledično i za sigurnost osiguranika da će biti realno i blagovremeno obeštećeni. Stoga je u inflatornim uslovima veoma važno uzeti u obzir inflaciju prilikom obračuna rezervacija za štete. U radu će na statističkim podacima jedne osiguravajuće kompanije biti prikazana primena Chain Ladder metode rezervisanja prilagođene za inflaciju i obrazloženi njeni efekti na adekvatno iskazivanje rezervacija za štete u inflatornim uslovima.

**Ključne reči:** rezervacije za štete, Chain Ladder metoda, inflacija, osiguranje od autoodgovornosti.

**Abstract:** This paper deals with the impact of inflation on loss reserves in motor third-party liability insurance, and consequently on the possibility of real indemnification of policyholders. The aim of the paper is to point out the need to take inflation into account when calculating loss reserves and to show how inflation can be included in their calculation. By its nature, insurance is more exposed to inflation than other sectors, bearing in mind that the period between the collection of the insurance premium and the realization of the harmful event resulting in the payment of compensation is long enough such that the detrimental inflation effect on the reality of the compensation can be manifested. Thereby, the insurance company has a legal obligation to form loss reserves, and their calculation is the task and responsibility of the actuary. The adequacy of loss reserves is important for the solvency of the insurer, and consequently for the security of the policyholders that they will be realistically and timely compensated. Therefore, in inflationary conditions, it is very important to take inflation into account when calculating loss reserves. The paper will present the application of the inflation-adjusted Chain Ladder reserving method on the statistical data of one insurance company and explain its effects on the adequate presentation of loss reserves in inflationary conditions.

**Keywords:** loss reserves, Chain Ladder method, inflation, motor third-party liability insurance.

### 1. UVOD

Poslovanje osiguravajućih kompanija uslovljeno je makroekonomskim okruženjem. Svetska ekonomija je na pragu nove recesije izazvane energetsom krizom zbog rata u Ukrajini. Evidentan rast cena energenata i prehrambenih proizvoda uslovljava rast inflacije. Rast inflacije ima negativne efekte na poslovanje osiguravajućih kompanija, koji se posebno ispoljavaju kroz pad realne vrednosti premija osiguranja i rast troškova koje ona treba da pokrije, a pre svega rashoda za štete. Takođe, dolazi do pada realne vrednosti rezervacija za štete, kao i do pada realne vrednosti aktive u koju su uložena sredstva ovih rezervi. Ovakve pojave mogu ugroziti solventnost osiguravajućih kompanija, a time i osnovnu funkciju osiguranja – blagovremenu isplatu naknada šteta, pa je neophodno zaštititi realnu vrednost premije i rezervacija za štete uzimanjem u obzir inflacije pri njihovom obračunu. Predmet istraživanja ovog rada je uticaj inflacije na rezervacije za štete i njeno uključivanje u metode obračuna ovih rezervacija, koje će biti potkrepljeno primerom na podacima konkretne osiguravajuće kompanije.

## 2. REZERVACIJE ZA ŠTETE I UTICAJ INFLACIJE

Izdavanjem sredstava iz tekuće tehničke premije, osiguravajuće kompanije formiraju tehničke rezerve koje su namenjene izmirenju obaveza prema osiguranicima. Ključni element tehničkih rezervi u neživotnom osiguranju su rezervacije za štete, kao procenjeni iznos obaveza osiguravača za štete koje su nastale, a nisu rešene do kraja tekućeg obračunskog perioda [7]. Od visine ovih rezervacija direktno zavisi profitabilnost, ali i solventnost osiguravača. Prema rezultatima mnogih empirijskih istraživanja, potcenjenost rezervacija za štete je jedan od osnovnih uzročnika nesolventnosti neživotnih osiguravača [1, 5, 13]. Otuda se ocena rezervacija za štete ubraja u najvažnije zadatke aktuara u neživotnom osiguranju.

Od nastanka štete do njenog rešavanja može proteći vremenski period različite dužine, u zavisnosti od vrste osiguranja. Tokom ovog perioda, brojni faktori utiču na konačan iznos šteta. Među ovim faktorima izdvaja se inflacija, koja obezvređuje iznos sredstava izdvojenih za izmirenje obaveza prema osiguranicima. Njeno dejstvo je posebno izraženo u vrstama osiguranja sa dugim repom, gde je potrebno da prođe više godina, pa čak i decenija, da bi šteta bila u potpunosti rešena. Tipičan primer vrste osiguranja sa dugim repom je osiguranje od odgovornosti zbog upotrebe motornih vozila (osiguranje od autoodgovornosti).

U svrhe ocene rezervacija za štete koriste se aktuarske metode rezervisanja, koje se obično zasnivaju na obrascu razvoja šteta u prošlosti, poput Chain Ladder metode. Istorijski podaci o štetama odražavaju prošla, ali ne i buduća inflatorna kretanja. Trouglovi rešenih iznosa šteta sadrže implicitnu inflaciju, kao inflaciju koja je zaista ispoljena tokom perioda na koji se oni odnose. Međutim, ako buduća inflacija bude veća od one koja je implicitno sadržana u trouglovima šteta, to znači da će za pokriće nastalih šteta biti potrebno više sredstava nego što je očekivano. U takvoj situaciji, rezervacije za štete će se pokazati nedovoljnim za izmirenje obaveza prema osiguranicima [4]. Otuda je, u uslovima rastuće inflacije, neophodno modifikovati metodologiju rezervisanja kako bi iznos rezervacija bio blagovremeno prilagođen i predupređena njihova neadekvatnost.

## 3. OBRAČUN REZERVACIJA ZA ŠTETE BEZ UKLJUČIVANJA INFLACIJE

Najčešće korišćena metoda određivanja rezervacija za nastale ali neprijavljene štete je Chain Ladder metoda. Polazeći od raspoloživih podataka o štetama iz prošlosti, ova metoda izvodi odgovarajuće faktore razvoja šteta, čijom se primenom u odnosu na iste podatke projektuju budući konačni iznosi šteta, potrebni za određivanje rezervacija za štete [8]. Neka slučajna promenljiva  $S_{i,j}$  označava iznos šteta nastalih u  $i$ -tom periodu, a rešenih (prijavljenih) u  $j$ -tom periodu. Ukoliko je  $1 \leq i \leq I$  i  $1 \leq j \leq J$ , inkrementalne štete, koje su poznate za kalendarske periode  $i + j \leq I + 1$ , obrazuju trougao šteta (Tabela 1).

**Tabela 1:** Trougao inkrementalnih iznosa šteta

Period nastanka	Period razvoja						
	1	2	...	$j$	...	$I-1$	$I$
1	$S_{1,1}$	$S_{1,2}$	...	$S_{1,j}$	...	$S_{1,I-1}$	$S_{1,I}$
2	$S_{2,1}$	$S_{2,2}$	...	$S_{2,j}$	...	$S_{2,I-1}$	
...	...	...	...	...	...	...	
$i$	$S_{i,1}$	$S_{i,2}$	...	$S_{i,j}$			
...	...	...	...	...			
$I-1$	$S_{I-1,1}$	$S_{I-1,2}$					
$I$	$S_{I,1}$						

Sumiranjem inkrementalnih iznosa šteta po periodima razvoja, za dati period nastanka  $i$ , dolazi se do kumulativnih iznosa šteta  $C_{i,j}$ :

$$C_{i,j} = \sum_{h=1}^j S_{i,h}, \quad (1)$$

čije su vrednosti poznate za  $i + j \leq I + 1$ . U cilju obračuna potrebnih rezervacija za štete, potrebno je oceniti vrednosti  $C_{i,j}$  za  $i + j > I + 1$ . Pretpostavka Chain Ladder metode je da postoje faktori razvoja  $f_j$  koji opisuju prirast kumulativnog iznosa šteta između  $j$ -tog i  $(j + 1)$ -og perioda razvoja, a čija je ocena:

$$\hat{f}_j = \frac{\sum_{i=1}^{I-j} C_{i,j+1}}{\sum_{i=1}^{I-j} C_{i,j}}, 1 \leq j \leq I-1. \quad (2)$$

Primenom ocenjenih faktora razvoja na iznos šteta nastalih u periodu  $i$ , a koje su rešene (prijavljene) zaključno sa tekućim kalendarskim periodom ( $C_{i,I+1-i}$ ) ocenjuje se konačan iznos šteta po pojedinim periodima nastanka šteta ( $C_{i,I}$ ):

$$\hat{C}_{i,I} = C_{i,I+1-i} \cdot \hat{f}_{I+1-i} \cdot \dots \cdot \hat{f}_{i-1}, 2 \leq i \leq I. \quad (3)$$

Potrebne rezervacije za štete koje su nastale u  $i$ -tom periodu se ocenjuju u vidu razlike:

$$\hat{R}_i = \hat{C}_{i,I} - C_{i,I+1-i}, 2 \leq i \leq I. \quad (4)$$

Primenu Chain Ladder metode možemo ilustrovati na podacima jedne osiguravajuće kompanije koja posluje u Srbiji, a koji se odnose na rešene štete u osiguranju od autoodgovornosti. Raspoloživi podaci o inkrementalnim iznosima rešenih šteta tokom prethodnog desetogodišnjeg perioda razvrstani su prema godini nastanka i godini u kojoj su date štete rešene u trougao šteta (Tabela 2). Potrebno je izračunati iznos rezervacija za štete na kraju 2021. godine.

**Tabela 2:** Trougao inkrementalnih isplata šteta (u 000 RSD)

Godina nastanka	Godina razvoja									
	1	2	3	4	5	6	7	8	9	10
2012	427.149	229.648	59.631	52.614	17.064	13.775	10.444	9.612	5.445	125
2013	505.066	310.643	65.494	41.494	37.077	36.031	12.671	8.972	3.740	
2014	565.456	266.842	52.683	22.555	8.216	7.889	5.471	888		
2015	771.509	341.808	57.371	37.972	22.088	17.879	8.204			
2016	825.745	399.948	71.691	40.174	11.128	15.975				
2017	862.641	365.555	69.055	55.641	33.038					
2018	996.430	477.949	88.537	48.422						
2019	1.040.047	459.921	114.434							
2020	1.022.582	512.884								
2021	1.037.228									

Sumiranjem inkrementalnih iznosa po godinama razvoja dobijaju se kumulativni iznosi rešenih šteta, na osnovu kojih se, prema obrascu (2), ocenjuju faktori razvoja. Pomoću ovih faktora se projektuju budući kumulativni iznosi šteta, čime se trougao kumulativnih šteta proširuje u pravougaonu formu (Tabela 3). Rezervacije za štete za svaku godinu nastanka šteta se računaju kao razlika između ocenjenog konačnog iznosa šteta (čije su vrednosti sadržane u poslednjoj koloni) i šteta rešenih zaključno sa tekućom godinom (čiji se iznosi nalaze na glavnoj dijagonali), u skladu sa obrascem (4).

**Tabela 3:** Rezultati ocenjivanja rezervacija za štete Chain Ladder metodom (u 000 RSD)

Godina nastanka	Godina razvoja										Ocenjene rezervacije za štete
	1	2	3	4	5	6	7	8	9	10	
2012	427.149	656.797	716.428	769.042	786.106	799.880	810.324	819.936	825.381	825.506	0
2013	505.066	815.709	881.203	922.697	959.774	995.805	1.008.476	1.017.448	1.021.188	1.021.343	155
2014	565.456	832.298	884.980	907.536	915.752	923.641	929.111	929.999	934.648	934.790	4.791
2015	771.509	1.113.317	1.170.689	1.208.661	1.230.749	1.248.628	1.256.832	1.265.738	1.272.065	1.272.258	15.426
2016	825.745	1.225.693	1.297.385	1.337.559	1.348.687	1.364.662	1.377.315	1.387.075	1.394.009	1.394.220	29.558
2017	862.641	1.228.196	1.297.252	1.352.893	1.385.931	1.410.140	1.423.215	1.433.300	1.440.465	1.440.683	54.752
2018	996.430	1.474.379	1.562.917	1.611.339	1.643.229	1.671.932	1.687.434	1.699.391	1.707.887	1.708.145	96.807
2019	1.040.047	1.499.969	1.614.403	1.676.176	1.709.350	1.739.208	1.755.333	1.767.772	1.776.609	1.776.878	162.475
2020	1.022.582	1.535.466	1.635.945	1.698.543	1.732.159	1.762.416	1.778.756	1.791.361	1.800.316	1.800.589	265.123
2021	1.037.228	1.534.686	1.635.114	1.697.680	1.731.279	1.761.521	1.777.853	1.790.451	1.799.402	1.799.674	762.446
<b>Ocenjeni faktori razvoja</b>	1,480	1,065	1,038	1,020	1,017	1,009	1,007	1,005	1,000	<b>Ukupno:</b>	<b>1.391.531</b>

Da bi ocena rezervacija za štete prema Chain Ladder metodi bila nepristrasna, iznosi šteta koji su nastali u različitim godinama moraju biti međusobno nezavisni [11]. Egzogeni faktori, poput rastuće stope inflacije,

čije se dejstvo odražava na dijagonalama trougla šteta, jednako utiču na štete iz većeg broja godina nastanka i dovode do narušavanja ove pretpostavke u praksi [9]. Kako se zasnivaju na podacima o štetama koji se odnose na više kalendarskih perioda, ocenjene vrednosti faktora razvoja, pored čistog razvoja šteta, implicitno odražavaju i inflatorna kretanja iz prošlosti, uz očekivanje da će se ona neizmenjeno nastaviti i u budućnosti. Ovakav pristup je opravdan ukoliko je stopa inflacije stabilna. Međutim, ukoliko stopa inflacije raste, ona mora biti eksplicitno uključena u obračun rezervacija za štete, kako bi se sprečila njihova potcenjenost [10].

#### 4. OBRAČUN REZERVACIJA ZA ŠTETE SA UKLJUČIVANJEM INFLACIJE

Da bi se u inflatornim uslovima obezbedila pouzdanost rezultata Chain Ladder metode, potrebno je najpre iz raspoloživih podataka o štetama iz prošlosti izolovati bazični razvoj šteta od uticaja inflacije, a zatim ga dalje projektovati, uz uvažavanje očekivanog budućeg delovanja inflacije. To znači da je u obračun rezervacija za štete potrebno uključiti eksternu informaciju o stopi inflacije u prošlim i budućim periodima [2].

Tokom perioda na koji se odnose podaci o štetama inflacija je bila niska i stabilna, da bi u 2021. godini došlo do naglog skoka (Tabela 4). Prosečna godišnja stopa inflacije u periodu 2013-2020. godine bila je 1,9%. U decembru 2021. godine stopa inflacije je iznosila 7,9%, nakon čega nastavlja da raste, da bi već u aprilu 2022. godine dostigla 9,6% [12]. Ovakav inflatorni rast upravo ukazuje na potrebu za uključivanjem dejstva inflacije u obračun rezervacija za štete.

**Tabela 4:** Inflacija merena procentualnim rastom potrošačkih cena

Godina	2013	2014	2015	2016	2017	2018	2019	2020	2021
Stopa inflacije na kraju godine	2,2%	1,7%	1,5%	1,6%	3,0%	2,0%	1,9%	1,3%	7,9%

Izvor: Narodna banka Srbije [12]

Kako bi iznosi šteta iz prošlosti bili iskazani u tekućim novčanim jedinicama, potrebno je uvećati inkrementalne isplate šteta primenom prikazanih stopa inflacije po dijagonalama trougla sadržanog u Tabeli 2. Uz pretpostavku da se isplate šteta vrše na kraju godine, isplate šteta izvršene u 2020. godini treba uvećati 1,079 puta, isplate izvršene u 2019. godini treba uvećati  $1,013 \times 1,079$  puta itd. Na taj način, dobija se trougao inkrementalnih isplata šteta koje su korigovane za prošlu inflaciju, tj. iskazane u novčanim jedinicama iz 2021. godine (Tabela 5).

**Tabela 5:** Trougao inkrementalnih isplata šteta korigovanih za inflaciju (u 000 RSD)

Godina nastanka	Godina razvoja									
	1	2	3	4	5	6	7	8	9	10
2012	535.139	281.513	71.876	62.482	19.951	15.635	11.622	10.501	5.875	125
2013	619.133	374.436	77.777	48.515	42.086	40.097	13.843	9.681	3.740	
2014	681.575	316.886	61.597	25.602	9.143	8.619	5.903	888		
2015	916.200	399.647	65.122	42.257	24.131	19.292	8.204			
2016	965.474	453.977	79.781	43.890	12.007	15.975				
2017	979.177	406.802	75.443	60.037	33.038					
2018	1.108.862	522.161	95.532	48.422						
2019	1.136.254	496.255	114.434							
2020	1.103.366	512.884								
2021	1.037.228									

Sumiranjem korigovanih inkrementalnih iznosa po godinama razvoja dobijaju se kumulativni iznosi rešenih šteta korigovani za inflaciju, na osnovu kojih se, prema standardnom algoritmu Chain Ladder metode, ocenjuju faktori razvoja. Ovi faktori su nešto niži od prvobitno ocenjenih faktora razvoja (sadržanih u Tabeli 3), budući da odražavaju čist razvoj šteta, bez uticaja inflacije [3]. Pomoću ovih faktora ocenjuju se buduće kumulativne isplate šteta, iskazane u tekućim novčanim jedinicama (Tabela 6).

**Tabela 6:** Raspoloživi i projektovani kumulativni iznosi rešenih šteta korigovani za inflaciju (u 000 RSD)

Godina nastanka	Godina razvoja									
	1	2	3	4	5	6	7	8	9	10
2012	535.139	816.652	888.529	951.010	970.961	986.597	998.219	1.008.720	1.014.595	1.014.720
2013	619.133	993.569	1.071.346	1.119.861	1.161.947	1.202.043	1.215.886	1.225.567	1.229.308	1.229.459
2014	681.575	998.461	1.060.058	1.085.661	1.094.804	1.103.423	1.109.326	1.110.213	1.114.991	1.115.129
2015	916.200	1.315.847	1.380.969	1.423.226	1.447.356	1.466.648	1.474.852	1.484.202	1.490.590	1.490.773
2016	965.474	1.419.452	1.499.232	1.543.122	1.555.130	1.571.105	1.584.170	1.594.213	1.601.074	1.601.271
2017	979.177	1.385.979	1.461.422	1.521.459	1.554.497	1.579.353	1.592.487	1.602.583	1.609.479	1.609.678

<b>2018</b>	1.108.862	1.631.023	1.726.554	1.774.977	1.807.567	1.836.469	1.851.740	1.863.480	1.871.500	1.871.730
<b>2019</b>	1.136.254	1.632.509	1.746.943	1.810.609	1.843.853	1.873.335	1.888.913	1.900.889	1.909.069	1.909.304
<b>2020</b>	1.103.366	1.616.250	1.717.974	1.780.584	1.813.277	1.842.270	1.857.590	1.869.367	1.877.412	1.877.643
<b>2021</b>	1.037.228	1.522.576	1.618.404	1.677.385	1.708.183	1.735.496	1.749.928	1.761.022	1.768.601	1.768.819
<b>Ocenjeni faktori razvoja</b>	1,468	1,063	1,036	1,018	1,016	1,008	1,006	1,004	1,000	

Na osnovu projektovanih kumulativnih isplata šteta iskazanih u tekućim novčanim jedinicama potrebno je izračunati buduće inkrementalne isplate šteta, koje su takođe iskazane u tekućim novčanim jedinicama (Tabela 7).

**Tabela 7:** Buduće inkrementalne isplate šteta u tekućim novčanim jedinicama (u 000 RSD)

Godina nastanka	Godina razvoja									
	1	2	3	4	5	6	7	8	9	10
2012										
2013										151
2014									4.778	137
2015								9.350	6.387	184
2016							13.065	10.043	6.861	197
2017						24.856	13.134	10.096	6.897	198
2018					32.590	28.902	15.272	11.740	8.020	231
2019				63.665	33.244	29.482	15.578	11.975	8.180	235
2020			101.724	62.610	32.693	28.993	15.320	11.777	8.045	231
2021	485.348	95.828	58.981	30.798	27.313	14.432	11.094	7.579	218	

Iako iskazani u tekućim novčanim jedinicama, iznosi prikazani u Tabeli 7 će biti isplaćeni u budućim godinama. Zbog toga ih je potrebno korigovati primenom očekivane buduće stope inflacije. Pod pretpostavkom da će prosečna godišnja stopa inflacije u narednom periodu iznositi 4%, isplate za koje se predviđa da će biti izvršene u 2022. godini treba uvećati 1,04 puta, isplate za koje se predviđa da će biti izvršene u 2023. godini treba uvećati  $1,04 \times 1,04$  puta itd. Na taj način, dobijaju se buduće inkrementalne isplate šteta korigovane za inflaciju. Njihovim sumiranjem po godinama razvoja ocenjuju se rezervacije za štete sa uključivanjem inflacije za svaku godinu nastanka šteta, i u ukupnom iznosu (Tabela 8).

**Tabela 8:** Buduće inkrementalne isplate šteta i rezervacije za štete korigovane za inflaciju (u 000 RSD)

Godina nastanka	Godina razvoja										Ocenjene rezervac. za štete sa efektom inflacije	
	1	2	3	4	5	6	7	8	9	10		
2012												
2013											158	158
2014									4.969	149		5.118
2015								9.724	6.908	207		16.839
2016							13.588	10.863	7.717	231		32.399
2017						25.850	14.205	11.357	8.068	241		59.721
2018					33.894	31.260	17.179	13.734	9.757	292		106.115
2019				66.212	35.957	33.164	18.224	14.570	10.351	310		178.787
2020			105.793	67.719	36.775	33.918	18.639	14.901	10.586	317		288.648
2021	504.761	103.648	66.346	36.029	33.230	18.261	14.599	10.372	310			787.557
											<b>Ukupno:</b>	<b>1.475.342</b>

Poređenjem rezultata ocenjivanja rezervacija za štete koji su sadržani u Tabelama 3. i 8. možemo zaključiti da su, u datom primeru, ukupne rezervacije za štete sa efektom inflacije za 6% veće od rezervacija za štete bez uključivanja inflacije. Ovo odstupanje bi bilo još izraženije ukoliko bi pretpostavljena buduća stopa inflacije bila veća. U uslovima naglašene neizvesnosti teško je sa zadovoljavajućom preciznošću predvideti kako će se stopa inflacije kretati u budućnosti. Stoga se preporučuje testiranje osetljivosti ocene rezervacija za štete na različite pretpostavljene buduće stope inflacije [6].

Ukoliko je stopa inflacije konstantna, standardna Chain Ladder metoda i Chain Ladder metoda prilagođena za inflaciju će rezultovati istom ocenom rezervacija za štete [14]. Međutim, ukoliko se očekuje da će stopa inflacije u budućnosti u proseku biti veća od stope inflacije u prošlosti, rezervacije za štete prema standardnoj Chain Ladder metodi će biti potcenjene, zbog čega je neophodno primeniti Chain Ladder metodu prilagođenu za inflaciju.

## 5. ZAKLJUČAK

U radu se predlaže primena Chain Ladder metode prilagođene za inflaciju u svrhe određivanja adekvatnih rezervacija za štete u uslovima rastuće inflacije. Kroz poređenje rezultujuće ocene rezervacija za štete sa standardnom Chain Ladder metodom na podacima o štetama jedne osiguravajuće kompanije u vrsti osiguranja od autoodgovornosti ukazano je na neophodnost eksplicitnog uvažavanja inflacije prilikom obračuna rezervacija za štete.

Kao mera inflacije u radu je korišćen procentualni rast potrošačkih cena. Ipak, treba imati u vidu da u pojedinim vrstama neživotnih osiguranja štete rastu brže od rasta opšteg nivoa cena. Takav je slučaj i sa osiguranjem od autoodgovornosti. Iznosi šteta na vozilima su prvenstveno uslovljeni kretanjem cena novih vozila, njihovih rezervnih delova i usluga servisiranja. Štete na licima, sa druge strane, prate rast cena medicinskih usluga i lekova (kada se nadoknađuju troškovi lečenja) i rast zarada (kada se nadoknađuje izgubljena zarada). Drugim rečima, za osiguravače kretanja cena u pojedinim sektorima obično imaju veći značaj od kretanja opšteg nivoa cena. Otuda je mogući budući pravac istraživanja uključivanje stvarne stope inflacije šteta u rezervacije za štete. Pored opšte stope inflacije, stopa inflacije šteta treba da uključi i dodatnu komponentu (tzv. socijalnu inflaciju) koja odražava dejstvo tehnološkog napretka i društvenih promena na štete, u zavisnosti od vrste osiguranja.

## LITERATURA

- [1] American Academy of Actuaries (1991). Study of insurance company insolvencies from 1969-87 to measure the effectiveness of casualty loss reserve opinions. *CAS Forum Winter 1991*, Arlington
- [2] Bențe, C. C., & Gavriletea, M. D. (2015). Inflation Adjusted Chain Ladder Method. *Annals of the University of Oradea, Economic Science Series*, Vol. 24, 88-89.
- [3] Booth, P., Chadburn, C., Haberman, S., James, D., Khorasane, Z., Plumb, R.H., & Rickayzen, B. (2005). *Modern Actuarial Theory and Practice*. Chapman & Hall/CRC
- [4] D'Arcy, S. P. (1982). A Strategy for Property-Liability Insurers in Inflationary Times. *Proceedings of the Casualty Actuarial Society*, 69, 163-186.
- [5] Dibra, S., & Leadbetter, D. (2007). *Why insurers fail – The dynamics of property and casualty insurance insolvency in Canada*. Toronto: Property and Casualty Insurance Compensation Corporation
- [6] Institute and Faculty of Actuaries (1997). *Claims Reserving Manual*, Vol. 1. Edinburgh
- [7] Kočović, J., & Koprivica, M. (2021). Ocena repnog faktora razvoja šteta ekstrapolacijom teorijskog oblika krive. *XLVIII Simpozijum o operacionim istraživanjima*, Zbornik radova, 123-128.
- [8] Kočović, J., Mitrašević, M., & Trifunović, D. (2019). Advantages and disadvantages of loss reserving methods in non-life insurance. *Yugoslav Journal of Operations Research*, 29(4), 553-561.
- [9] Kočović, J., Rajić, V., & Jovović, M. (2012). Prednosti i nedostaci Chain Ladder metoda za procenu rezervi za štete. *XXXIX Simpozijum o operacionim istraživanjima*, Zbornik radova, 4 strane
- [10] Koprivica, M. (2022). The impact of inflation risk on life and non-life insurers. *Development of modern insurance market - constraints and possibilities*, Kočović, J. et al. (eds.), Belgrade: University of Belgrade - Faculty of Economics, 255-271.
- [11] Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23(2), 213-225.
- [12] Narodna banka Srbije (2022). *Osnovni makroekonomski pokazatelji*. Preuzeto 7. juna 2022. sa: [https://nbs.rs/sr\\_RS/drugi-nivo-navigacije/statistika](https://nbs.rs/sr_RS/drugi-nivo-navigacije/statistika)
- [13] Sharma, P. et al. (2002). Prudential Supervision of Insurance Undertakings. *Conference of Insurance Supervisory Services of the Member States of the European Union*. Brussels: European Commission
- [14] Taylor, G. (2000). *Loss Reserving: An Actuarial Perspective*. Kluwer Academic Publishers

## EFEKTI PRIMENE DIGITALNIH TEHNOLOGIJA U OSIGURANJU

### EFFECTS OF APPLICATION OF DIGITAL TECHNOLOGIES IN INSURANCE

Nevena Ž. Petrović<sup>1</sup>

<sup>1</sup> Nacionalno telo za akreditaciju i obezbeđenje kvaliteta u visokom obrazovanju,  
[petrovicnevena97@gmail.com](mailto:petrovicnevena97@gmail.com)

---

#### Rezime

*Predmet istraživanja ovog rada je primena digitalnih tehnologija kao što su analiza podataka u Big data okruženju, internet stvari i veštačka inteligencija u osiguranju. Digitalizacija kao osnovna tekovina IV industrijske revolucije dovela je do značajnog snižavanja troškova poslovanja društava za osiguranje: smanjene su administrativne procedure, eliminisane su određene aktivnosti, itd. S druge strane, upotreba pametnih uređaja i Big data su omogućili osiguravačima pristup brojnim informacijama koje omogućuju preciznije merenje rizika i, samim tim, određivanje premije osiguranja, kao i obradu odštetnih zahteva. Takođe, primena informacione i komunikacione tehnologije i razvoj elektronskog poslovanja dovode i do pojave novih vrsta osiguranja, kao što su, na primer, osiguranje informacionih sistema, osiguranje od šteta koje nastaju kao posledica zloupotrebe podataka, itd. U radu su prikazani i konkretni primeri implementacije veštačke inteligencije u američkim osiguravajućim kompanijama gde se primena veštačke inteligencije pokazala kao izuzetno dobra praksa sa dva aspekta: procene rizika i pravilnog definisanja premija osiguranja, kao i procene šteta u slučaju realizacije osiguranog slučaja.*

*Ključne reči: osiguranje, četvrta industrijska revolucija, digitalizacija, elektronsko poslovanje*

#### Abstract

*The subject of research in this paper is the application of digital technologies such as data analysis in the Big data environment, the Internet of Things and Artificial intelligence in insurance. Digitalization as the basic achievement of the IV Industrial Revolution has led to a significant reduction in the operating costs of insurance companies: administrative procedures have been reduced, certain activities have been eliminated, etc. On the other hand, the use of smart devices and Big data have provided insurers with a wealth of information that allows for more accurate risk measurement and, thus, the determination of insurance premiums, as well as the processing of claims. Also, the implementation of information and communication technology and the development of electronic business leads to the emergence of new types of insurance within the offers of insurance companies, such as insurance of information systems, insurance against damages resulting from the misuse of data on policyholders, etc. The paper also presents concrete examples of the implementation of artificial intelligence in American insurance companies, where the use of artificial intelligence has proven to be extremely good practice in two aspects: risk assessment and proper definition of insurance premiums, and damage assessment in insured case.*

*Key words: insurance, fourth industrial revolution, digitalization, electronic business*

## 1. UVOD

Digitalizacija i povećanje korišćenja interneta i informacionih i komunikacionih tehnologija u oblasti osiguranja imaju izuzetno značajan uticaj na društveni život ljudi, ali i na

ekonomske aspekte poslovanja. Naime, proces digitalizacije je nametnuo društvima za osiguranje potrebu da na adekvatan način pristupe tržištu osiguranja, da razviju mehanizam interakcije sa osiguranicima i da prilagode karakteristike usluga osiguranja svakom pojedinačnom klijentu i njihovim preferencijama.

Predmet analize rada su najpre Big data i njena primena u osiguranju, kao i internet stvari (engl. *Internet of Things, IoT*). Takođe, biće analizirani i digitalni kanali distribucije u osiguranju kao i primena veštačke inteligencije u osiguranju.

## 2. PRIMENA BIG DATA U OSIGURANJU

Jedan od izazova četvrte industrijske revolucije jeste količina podataka koji se prenose i čuvaju preko Interneta. Kako se broj podataka povećava iz minuta u minut kreirani su novi načini prikupljanja i skladištenja velikog broja podataka u realnom vremenu koji su bazirani na *Big data* tehnologijama. Zahvaljujući Nauci o podacima (engl. *Data Science*) koja je mnogo uticala na poboljšanje interakcije između osiguravača i osiguranika kao i na komercijalizaciju usluga osiguranja, prikupljene podatke možemo obraditi i analizirati kako bismo dobili sliku o njihovom sadržaju. Sam postupak podrazumeva da pametni uređaji (telefoni, satovi, televizori i sl.) vrše prikupljanje podataka, potom se ti podaci usmeravaju na analizu i obradu, da bi nam se vratili natrag u onom obliku koji je nama razumljiv i koji možemo na adekvatan način da iskoristimo.

Zahvaljujući upotrebi ove tehnologije osiguravajućim kompanijama je omogućeno da saznaju mnogo bitnih činjenica o osiguranicima, što je uticalo da se oni sami uključe u kreiranje proizvoda i usluga preko svojih pametnih uređaja. Naime, društva za osiguranje su koristila *Big data* da bi se prevashodno pravilno odredila cena koju su osiguranici dužni da plate, zatim da bi se pravilno procenila šteta, kao i rizik kojem su osiguranik i sama osiguravajuća kompanija izloženi.

Iz Tabele broj 1 može se videti da se na tlu Evropske unije iz godine u godinu povećava broj kompanija koje koriste *Big data*. Prema podacima Eurostata za 2020. godinu, broj kompanija sa više od 10 zaposlenih, koje koriste *Big data* analizu podataka, u odnosu na ukupan broj registrovanih kompanija sa više od 10 zaposlenih iznosio je 13%. Analiza podataka u *Big data* okruženju je najviše zastupljena u Danskoj, Ujedinjenom Kraljevstvu, Belgiji, Holandiji, Islandu, Francuskoj, Nemačkoj i na Malti. Najmanju zastupljenost *Big data* od prikazanih zemalja u tabeli imaju Kipar i Republika Srbija [7]. Poredeći našu zemlju sa zemljama u regionu, možemo da zaključimo da se Republika Srbija nalazi u nezavidnom položaju. Naime, u Bosni i Hercegovini preduzeća koja imaju više od 10 zaposlenih, 2,5 puta više koriste *Big data*, dok u Severnoj Makedoniji šest puta više, a u Crnoj Gori čak 6,5 puta više. Istraživanje pokazuje da bi u narednom periodu trebalo raditi na prihvatanju analize podataka u *Big data* okruženju kao bitnog segmenta u poslovanju kompanija u Republici Srbiji, jer ovakva analiza obezbeđuje veliki dijapazon informacija koje se mogu iskoristiti sa ciljem formiranja konkurentne prednosti i unapređenja poslovanja same kompanije, ali i povećanja zadovoljstva potrošača.

U izveštaju Kanadskog instituta za osiguranje pod nazivom „Veštačka inteligencija i veliki podaci: implikacije za industriju osiguranja u Kanadi“, došlo se do zaključka da će doći do velikih promena u sektoru osiguranja zbog višestrukih mogućnosti koje digitalna tehnologija nudi, naročito kada je reč o zaključivanju ugovora o osiguranju, upravljanju odštetnim zahtevima, upravljanju investicijama, praksi zapošljavanja i marketingu u oblasti osiguranja [8].



Tabela br. 1 Procentualno učešće preduzeća u ukupnom broju preduzeća, koje koriste *Big data* analizu podataka na Evropskom tlu

Zemlja	% preduzeća koje koriste <i>Big data</i> analizu
EU-27	13
Belgija	22
Danska	24
Češka	9
Bugarska	6
Nemačka	17
Island	22
Estonija	8
Grčka	12
Španija	6
Francuska	20
Hrvatska	13
Italija	7
Kipar	3
Malta	29
Holandija	26
Ujedinjeno Kraljevstvo	25
Crna Gora	13
Severna Makedonija	12
Bosna i Hercegovina	5
Srbija	2

Izvor: [http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc\\_eb\\_bd&lang=en](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_eb_bd&lang=en)

Tokom poslednjih godina osiguravajuće kompanije su uz pomoć *Big data* analize podataka razvile nove proizvode u oblasti osiguranja automobila i to prevashodno upotrebom telematike. Telematsko osiguranje automobila podrazumeva da se vrši ugradnja uređaja (*Blackbox*) u motorna vozila ili se praćenje vozila vrši putem aplikacija na pametnim telefonima. *Blackbox* uređaji prate brzinu kretanja, kočenje, ubrzanje, krivine, doba dana kada se obavlja putovanje i svi ovi podaci se putem satelitske tehnologije prenose do osiguravača [4]. Na taj način osiguravači imaju mogućnost da pravilno procene verovatnoću nastanka osiguranog slučaja, pravilno odrede cenu svojih usluga, ali i da personalizuju svoje usluge, zavisno od ponašanja osiguranika u saobraćaju i njihovih potreba.

Telematika može da unapredi i osiguranje imovine u budućem periodu. Naime, do sada prilikom određivanja premije osiguranja, popusti su davani onim osiguranicima koji imaju ugrađene protivpožarne alarme, protivprovalne alarme i malo je primera kućne telematike kada je osiguranje u pitanju. Jedan od retkih primera primene telematike i *IoT* u ovoj oblasti je saradnja *Allianz and Deutsche Telekom*-a, gde digitalno povezane kuće automatski upozoravaju *Allianz*-ovu hitnu liniju ukoliko dođe do realizacije nekog rizika koji može da ošteti imovinu, kao što je na primer pucanje cevi [1].

U oblasti marketinga *Big data* daje brojne pogodnosti. Naime, veliki broj podataka koji se prikupi putem pametnih uređaja i obradi korišćenjem adekvatne informacione tehnologije, daje mogućnost osiguravačima

da svoju ponudu prilagode u skladu sa specifičnim zahtevima korisnika i to, kako putem direktnih, tako i putem indirektnih kanala distribucije [4].

Kada je reč o distribuciji usluga osiguranja, internet je omogućio korisnicima da lakše porede cene osiguranja. *Big data* analiza podataka omogućava osiguravačima da izvrše diverzifikaciju korisnika i prilagode im svoju ponudu. Naime, pametni uređaji i društvene mreže se već od strane nekih osiguravajućih kompanija koriste u otkrivanju lažnih podnosilaca odštetnih zahteva [4], a u budućem periodu očekuje se da upotreba ovih podataka u otkrivanju prevara u osiguranju bude još masovnija.

### 3. INTERNET STVARI U OSIGURANJU

Jedna od najcitiranijih definicija Interneta stvari jeste da on predstavlja upotrebu inteligentnih povezanih uređaja i sistema za iskorišćavanje prikupljenih podataka putem ugrađenih senzora u mašinama i drugim fizičkim predmetima.

Podaci koji se obezbeđuju putem *IoT* uređaja generišu osiguravajućim kompanijama obilje podataka, koji im daju mogućnost da vrše preciznije merenje rizika, odnosno da u potpunosti izvrše individualizaciju tarifa za svakog osiguranika. Ali svakako se mora napomenuti da do individualizacije polisa osiguranja još uvek nije došlo i da se ona, uz značajna ulaganja, naročito kada je reč o velikim igračima u industriji osiguranja očekuje, odnosno trebalo bi da polise budu prilagođene pojedinačnim korisnicima osiguranja.

*IoT* uređaji, naročito kada je osiguranje imovine u pitanju, obezbeđuju informacije koje smanjuju verovatnoću nastanka osiguranog slučaja. Takođe, pametni uređaji su omogućili pomak i u pogledu odštetnih zahteva. Naime, putem informacija koje se prenose do osiguravača, oni omogućavaju da se šteta koja nastane realnije proceni, ali i da se odštetni zahtevi obrade u kraćem roku, što jeste značajno za osiguranike, ali i za same osiguravače, u smislu da imaju niže troškove pri obradi zahteva.

*IoT* i *Big data* analiza podataka obezbeđuju brojne pogodnosti i za industriju osiguranja i za njene osiguranike, pa se u budućem periodu očekuje sve veća primena kako u svetu, tako i kod nas.

### 4. DIGITALIZACIJA KANALA PRODAJE

Kako je digitalna revolucija uzela maha i u oblasti osiguranja, osiguravajuće kompanije moraju držati korak sa tehnološkim promenama kako bi privukle što veći broj korisnika osiguranja, smanjile troškove poslovanja, povećale konkurentnost, efikasnost, fleksibilnost i agilnost operativnih modela kako bi se osvojila nova tržišta i stvorile nove mogućnosti. Osiguranici očekuju personalizovane proizvode, usluge i korisničko iskustvo koje je podržano preferiranim kanalom komunikacije (društvene mreže, veb, portali). Posledično, osiguravači treba da skrate vreme do izlaska na tržište i pruže konkurentne proizvode i usluge. Osnova za rešavanje pomenutih izazova jeste uspostavljanje *cloud* infrastrukture i rešenja. Prema Izveštaju o globalnim rizicima za 2022. godinu digitalna nejednakost se i dalje svrstava u kategoriju najvećih rizika u nerazvijenim i zemljama u razvoju, odnosno u zemljama sa niskim prihodima [3]. Zemlje koje će se najviše suočiti sa ovim problemom su zemlje Latinske Amerike i Podsaharske Afrike. Dodatna zabrinutost sa kojom se ove zemlje suočavaju je u vezi sa sajber bezbednošću. Dodatni podstrek investiranja društava za osiguranje u nove tehnologije i usvajanja rešenja u *cloud*-u prouzrokovala je kriza izazvana koronavirusom (SARS-CoV-2).

Prema istraživanju koje je 2019. godine sprovela kompanija *Accenture* osiguravajuće kuće koje konstantno idu u korak sa razvojem *cloud* tehnologija mogu da očekuju izvanredan uticaj na rast prihoda, pa čak i dupli tokom osmogodišnjeg perioda (od 2015. do 2023. godine). Međutim, sledbenici su u opasnosti da izgube 37% potencijalnog prihoda u 2023. godini. Prema ovom istraživanju lideri u osiguranju usvajaju nove tehnologije mnogo brže od sledbenika, a to naročito važi za tehnologije kao što su *Big Data* analitika, *Cloud SaaS*, veštačka inteligencija, mašinsko učenje, *Internet of Things (IoT)*, itd. Lideri u industriji osiguranja su sve više otvoreni i sve brže se kreću ka ostvarivanju potencijala *cloud* tehnologija, s jedne strane; dok su sledbenici u problemu jer imaju previše različitih sistema, ali i dalje održavaju prethodne IT strukture, što ima veliki uticaj na poslovnu efikasnost i prihode. Iz istraživanja se može zaključiti da je poslovanje u *cloud*-

u dugo putovanje koje zahteva konstantan napor, resurse i posvećene stručnjake odgovorne za definisanje i implementaciju strategija za uspešno poslovanje u *cloud* okruženju.

## 5. VEŠTAČKA INTELIGENCIJA

Primena veštačke inteligencije (engl. *Artificial Intelligence – AI*) u sektoru osiguranja doprinosi većoj brzini i tačnosti prilikom preuzimanja rizika i upravljanja odštetnim zahtevima klijenata i boljem poštovanju zakonskih i podzakonskih propisa. Osiguravajuće kompanije najviše koriste *AI* prilikom zaključivanja ugovora sa osiguranicima i prilikom obrade njihovih odštetnih zahteva. Primena robota u sektoru osiguranja kao finansijskih savetnika koji svoje usluge pružaju po nižoj ceni u odnosu na tradicionalne finansijske savetnike, a da su pritom te usluge visokog kvaliteta i višeg nivoa transparentnosti u odnosu na tradicionalne finansijske savetnike postaje sve izraženija [2]. Roboti savetnici mogu da upute potencijalne ugovarače na personalizovane proizvode osiguranja, i to predlaganjem relevantne polise osiguranja na osnovu razumevanja potreba osiguranika [4]. Jedna od najvećih američkih osiguravajućih kompanija *Liberty Mutual* je osnovala „*Solaria Labs*“ kompaniju koja je specijalizovana za inovacije u osiguranju. Društvo za osiguranje *Liberty Mutual* se preko svoje kompanije koja je specijalizovana za inovacije bavi razvojem aplikacija za telematiku zasnovane na mašinskom učenju sa sposobnošću veštačke inteligencije, a sa ciljem unapređenja bezbednosti vozača (primer je portal iz 2017. godine sa otvorenim aplikacionim programskim interfejsom - *Application Programming Interface*). Naročito je bitno istaći da njihova nova aplikacija treba da pomogne vozačima koji dožive saobraćajnu nezgodu da izvrše procenu štete na svojim vozilima u stvarnom vremenu, i to upotrebom kamere na njihovom pametnom telefonu. Sama aplikacija koristi fotografije štete radi uporedne analize štete i ima za cilj da pruži procenu troškova konkretne popravke automobila. Na duži vremenski period ovo bi moglo da dovede do smanjenja troškova procene šteta i jednostavnijeg i bržeg postupka obrade samih šteta [6].

Najveća američka osiguravajuća kuća *State Farm* od 2016. godine koristi mašinsko učenje prilikom klasifikacije sigurnosti vožnje svakog vozača koji je kod njih osiguran, a takođe i prilikom izrade i pronalaženja odgovarajuće ponude usluga za svaku klasu vozača. Ova osiguravajuća kompanija je pre nekoliko godina otvorila konkurs za najuspešnije rešenje za klasifikaciju vozača, a na osnovu fotografija koje potiču iz saobraćaja. Vozači su klasifikovani u deset kategorija počev od: sigurnih vozača; preko onih koji šalju SMS poruke u vožnji; vozača koji podešavaju radio tokom vožnje; onih koji razgovaraju preko telefona, itd... Prvo mesto na navedenom konkursu je osvojilo rešenje koje je koristilo mašinsko učenje i metodudve neuronskemreže [5]. U budućnosti se očekuje da i domaće kompanije posvete više pažnje značaju veštačke inteligencije u oblasti osiguranja.

## 6. ZAKLJUČAK

Proces digitalizacije je nametnuo odgovarajuće izazove sa kojima su se osiguravajuće kompanije suočile i to u vidu potrebe za unapređenjem iskustava korisnika njihovih usluga, poboljšanja poslovnih procesa, kreiranja novih proizvoda i usluga sa ciljem povećanja satisfakcije klijenata, kao i pojave novih konkurenata na tržištu osiguranja, koji, usled upotrebe savremene tehnologije, žestoko pariraju tradicionalnim društvima za osiguranje. Osnovni razlog sporijeg prihvatanja digitalizacije jeste kompleksnost osiguranja, ali i visok nivo pravne regulative u ovoj oblasti. Proces digitalizacije je uticao na eliminaciju velikog broja manuelnih aktivnosti i poslova koji nisu neophodni, što je za posledicu imalo rast efikasnosti i samim tim smanjenje troškova poslovanja osiguravajućih kompanija. Ovi troškovi su smanjeni i kada je prijavljivanje štete u pitanju, slanje zahteva za produženje trajanja osiguranja i sl. Takođe, pametni uređaji koji se koriste od strane klijenata obezbeđuju značajne podatke kada je obrada odštetnih zahteva u pitanju i po osnovu smanjenja prevara u sektoru osiguranja. U procesu prikupljanja i obrade podataka od potencijalnih osiguranika veliku ulogu ima upotreba pametnih uređaja i *Big data* analitike koja pomaže osiguravajućim kompanijama da se rizik pravilno odredi i premija osiguranja precizno definiše.

Primena savremene tehnike i tehnologije u sektoru osiguranja dovela je i do promene u ponudi osiguravajućih kompanija. To je pre svega prisutno u visokorazvijenim zemljama i to kod velikih igrača na

tržištu osiguranja u tim zemljama, koji su u svoju ponudu uvrstili usluge poput osiguranja informacionih sistema, osiguranja od šteta koje mogu nastati kao posledica zloupotrebe podataka koje su kompanije prikupile o svojim kupcima i sl. Bitno je napomenuti da je i u postojećim oblicima osiguranja došlo je do značajnih izmena za koje je najviše zaslužna upotreba telematike, kao spoja komunikacione i informacione tehnologije za slanje i prijem podataka. To omogućava osigravačima da preciznije procene verovatnoću nastanka osiguranog slučaja, pravilno odrede cenu svojih usluga, ali i da personalizuju svoje usluge, zavisno od ponašanja osiguranika u saobraćaju i njihovih potreba. Takođe, očekuje se sve veća primena telematike i u oblasti osiguranja imovine.

Kada je naša zemlja u pitanju, prisutni su određeni pomoci u oblasti digitalizacije u osiguranju, ali se ipak može zaključiti da su oni još uvek na nižem nivou u odnosu na razvijene zemlje. Potrebno je da društva za osiguranje shvate značaj digitalizacije kao osnovu za uvećanje svojih prihoda, ali je takođe potrebno digitalno opismenjavanje osiguranika kako bi nove trendove lakše prihvatili.

## 7. LITERATURA

- [1] Allianz SE (2014), Allianz and Deutsche Telekom enter into a digital alliance, [https://www.allianz.com/en/press/news/financials/stakes\\_investments/news-2014-06-06.html](https://www.allianz.com/en/press/news/financials/stakes_investments/news-2014-06-06.html) (accessed on 18 March 2019).
- [2] Baker, T., Dellaert, B.G.C. (2018). Regulating Robo Advice Across the Financial Services Industry, *Iowa Law Review*, Vol. 103, 713–750.
- [3] Global Risk Report 2022, WEF, [https://www3.weforum.org/docs/WEF\\_Global\\_Risks\\_Report\\_2022.pdf](https://www3.weforum.org/docs/WEF_Global_Risks_Report_2022.pdf)
- [4] OECD, The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector, OECD, 2020., 23.
- [5] Pavlović, B., Izazovi u primeni mašinskog učenja u delatnosti osiguranja, *Tokovi osiguranja*, 3/2019, 17.
- [6] Sennaar, K. (2017). How America's Top 4 Insurance Companies are Using Machine Learning, available at <https://www.techemergence.com/machine-learning-at-insurance-companies/>.
- [7] <https://ec.europa.eu/eurostat>
- [8] <https://www.insurancebusinessmag.com/ca/news/digital-age/ai-big-data-analytics-expected-to-transform-the-insurance-industry-in-the-next-10-years--iic-250194.aspx> )

## HETEROGENA OČEKIVANJA INVESTITORA I CENE AKCIJA

### HETEROGENEUS EXPECTATIONS OF INVESTORS AND STOCK PRICES

DRAGANA DRAGANAC<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, dragana.draganac@ekof.bg.ac.rs

**Rezime:** U ovom radu je dat prikaz modela koji analiziraju uticaj heterogenih očekivanja investitora i neslaganja investitora na tržišnu cenu akcija, prinos na akcije i obim trgovanja. U slučajevima kada je nivo neslaganja investitora u vezi sa nekim fundamentalnim pokazateljem mali, na tržištu učestvuju kako investitori optimisti, tako i investitori pesimisti. Tada postojanje neslaganja i činjenica da je ograničena prodaja na kratko ne utiču na cenu akcije. Međutim, kada je nivo neslaganja veliki, ograničenje prodaje na kratko uzrokuje povlačenje pesimista, cena akcije odražava verovanja samo optimističnih investitora, što rezultira precenjenošću akcije. Stepenn neslaganja investitora raste sa porastom beta koeficijenta, gde cena akcije sa visokim beta koeficijentom sadrži špekulativnu premiju, koja je rastuća funkcija beta koeficijenta. Posledično, na visokim nivoima agregatnog neslaganja, SML funkcija postaje opadajuća, što je u suprotnosti sa standardnim CAPM modelom.

**Ključne reči:** Modeli heterogenih očekivanja, Modeli neslaganja investitora, Percepcija investitora, Ograničenja prodaje na kratko, Precenjenost, Špekulativna premija, Bihevioralne finansije

**Abstract:** This paper presents a set of models that analyze the impact of heterogeneous expectations of investors and investors' disagreement on stock price, its return and turnover. If the disagreement between investors about some fundamental indicator is small, both optimistic and pessimistic investors are present in the market. In that case, the disagreement between two groups of investors and short-sale constraints do not affect the stock price. When the level of disagreement between optimists and pessimists is high, pessimists are sidelined due to short-sale constraints. In that case, the stock price reflects only the beliefs of optimistic investors which results in overpricing. The disagreement between investors increases with the rise in beta coefficient. The price of a stock with a high beta coefficient contains a speculative premium, which is the increasing function of the beta coefficient. Consequently, at high levels of aggregate disagreement, the SML function is downward sloping, which is contrary to the standard CAPM model.

**Keywords:** Heterogeneous Expectations Models, Disagreement Finance, Investors' Perception, Short-sale Constraints, Overpricing, Speculative Premium, Behavioral Finance

## 1. UVOD

Tradicionalne finansije se zasnivaju na nizu pretpostavki, kao što su simetrična informisanost tržišnih učesnika, njihova potpuna racionalnost, savršenost tržišta kapitala, nepostojanje precenjenosti ni potenjenosti finansijskih instrumenata, i sl. Mnoge od ovih pretpostavki nisu realistične niti ispunjene u praksi. Za ovaj rad je posebno interesantna pretpostavka tradicionalnih finansija o homogenim očekivanjima investitora o budućim stopama prinosa, rizicima i korelaciji prinosa na akcije. Nerealnost navedenih pretpostavki, a naročito uočene tržišne zagonetke i anomalije, sa kojima tradicionalne finansije nisu uspele da se izbore, uslovile su razvoj bihevioralnih finansija, koje su unele psihološke i sociološke koncepte u ekonomiju i finansije i tako uspele da pojasne tržišne zagonetke i anomalije. U bihevioralnim finansijama se finansijske pojave razmatraju u znatno realističnijoj postavci, gde investitori trguju u uslovima asimetričnih informacija, često imaju heterogena očekivanja, a arbitraža nije ni besplatna, ni nerizična, već ima niz ograničenja, koja mogu čak da dovedu do povlačenja arbitražera sa tržišta.

Ogroman broj transakcija na tržištima akcija, obveznica i deviznim tržištima govori u prilog činjenici da investitori često imaju heterogena očekivanja u vezi sa rizikom i prinosom na akcije. Homogena očekivanja bi, nakon uspostavljanja ravnoteže i izbora optimalnih portfolija, dovela do toga da, u stanju ravnoteže, obim transakcija bude nula. Investitori imaju različita očekivanja ne samo zato što imaju različite informacije; oni mogu da poseduju potpuno iste informacije, ali da ih različito obrade i interpretiraju. Jedna grupa može imati

optimistične prognoze, pa će se pojaviti na strani tražnje; drugi su pesimisti, pa će se pojaviti na strani ponude. Upravo ovo neslaganje u stavovima i percepcijama investitora je jedan od važnijih faktora koji generiše trgovanje, ali i utiče na cene i prinose na akcije.

U nastavku rada je dat prikaz nekoliko modela heterogenih očekivanja i neslaganja investitora. Reč je o modelima i nalazima autora *Miller (1977)*, *Chen et al. (2002)*, *Hong i Stein (2007)* i *Hong i Sraer (2016)*.

## 2. MODELI HETEROGENIH OČEKIVANJA I NESLAGANJA INVESTITORA I IMPLIKACIJE NA CENU AKCIJA

[3] ističu da modeli sa heterogenim agentima, tj. modeli „neslaganja“, imaju sledeće mehanizme u osnovi: postepen tok informacija, ograničenu pažnju i heterogena verovanja investitora. Različite prognoze mogu biti u vezi sa stopom rasta industrijske proizvodnje, stopom inflacije, kamatnom stopom, kreditnim bonitetom, i sl. Neslaganje često potiče od kognitivnih pristrasnosti investitora: samopouzdati investitori su skloni da daju preveliki ponder svojim privatnim signalima.

Neslaganje investitora na nivou akcija [2] mere preko visine standardne devijacije dugoročne stope rasta *EPS*-a, prognozirane od strane analitičara, za svaki mesec i za svaku akciju. Standardna devijacija prognozirane dugoročne stope rasta *EPS*-a ukazuje na neslaganja između investitora u pogledu prinostnog potencijala akcije. Agregatno neslaganje [2] računaju kao ponderisan prosek neslaganja na nivou akcija, gde su ponderi beta koeficijenti svake akcije. Na grafikonu 1 je dat prikaz agregatnog neslaganja investitora na *NYSE*, u vremenskom periodu od 1981. do 2014. godine.



**Grafikon 1:** Vremenska serija agregatnog neslaganja investitora u vezi sa dugoročnom stopom rasta *EPS*-a  
Izvor: [2], p. 2099

[4] i [1] su razvili model sa heterogenim očekivanjima u kome su pokazali kako se, kada je zabranjena prodaja na kratko, formira cena u slučaju kada je nivo šuma mali, a kako kada je nivo šuma veliki. Uzeto je da je nerizična stopa prinosa 0. Ponuda je standardizovana i iznosi 1. Tražnja optimista za akcijom 1 je data formulom

$$\max \left[ \eta \frac{w_1(\bar{z} + \varepsilon) - p_1}{\text{var}(d_1)}, 0 \right] \quad (1)$$

a tražnja pesimista sledećom formulom

$$\max \left[ \eta \frac{w_1(\bar{z} - \varepsilon) - p_1}{\text{var}(d_1)}, 0 \right] \quad (2)$$

gde je  $\eta$  koeficijent tolerancije prema riziku,  $w_1$  je tražnja za akcijom 1;  $\bar{z}$  je srednja vrednost agregatnog šoka u vezi sa dividendom;  $\varepsilon$  predstavlja veličinu šuma i pokazuje koliko se verovanje investitora optimista i pesimista u vezi sa srednjom vrednošću agregatnog šoka razlikuje u odnosu na pravu srednju vrednost, tj.  $\bar{z}$ ;  $p_1$  je cena akcije 1;  $\text{var}(d_1)$  je varijansa dividende. Agregatni šok u vezi sa dividendama predstavlja značajno povećanje iznosa isplaćenih dividendi u celoj privredi, koji može uticati na agregatne varijable.

Kada je nivo šuma mali, ograničenje prodaje na kratko je irelevantno, a ravnotežna cena je data formulom:

$$p_1 = w_1 \bar{z} - \frac{1}{2} \frac{1}{\eta} \text{var}(d_1) \quad (3)$$

Možemo zapaziti da, u formuli za ravnotežnu cenu, nivo šuma ne figurira, jer obe grupe investitora učestvuju na tržištu.

Ako je nivo šuma  $\varepsilon$  veliki, ograničenje prodaje na kratko je relevantno. Zbog ovog ograničenja, pesimisti odlučuju da ne učestvuju na tržištu. Ravnotežna cena iznosi:

$$p_1 = w_1(\bar{z} + \varepsilon) - \frac{1}{\eta} \text{var}(d_1). \quad (4)$$

U ovom slučaju, ravnotežna cena zavisi i od nivoa šuma. Ravnotežna cena je viša kada je nivo šuma veći. Akcija će biti procenjena, jer su na tržištu prisutni samo optimisti.

[2] su nadgradili prethodne modele razvijajući svoj model u vezi sa formiranjem ravnoteže na tržištu u uslovima heterogenih očekivanja, gde prave vezu između veličine neslaganja investitora i beta koeficijenta akcije. Na tržištu postoje dve grupe investitora: uzajamni fondovi i hedž fondovi. Uzajamni fondovi imaju heterogena verovanja, tj. među njima ima optimista i pesimista, koji imaju različita verovanja o visini  $\tilde{z}$ .<sup>1</sup> Njihov udeo u ukupnoj populaciji investitora iznosi  $\mu$  i njima nije dozvoljena prodaja na kratko. Druga grupa investitora, tj. hedž fondovi, imaju ulogu arbitražera na tržištu, mogu da prodaju na kratko i imaju homogena verovanja. Njihov udeo u ukupnoj populaciji investitora iznosi  $1 - \mu$ .

Model ima dva vremenska perioda, 0 i 1.<sup>2</sup> Nerizična stopa prinosa je egzogeno data i iznosi  $r$ . Na tržištu postoji  $N$  rizičnih finansijskih instrumenata (akcija), koja u periodu 1 isplaćuju dividendu  $\tilde{d}_i$ , koja se formira prema slučajnom procesu datom sledećom formulom:

$$\tilde{d}_i = b_i \tilde{z} + \tilde{u}_i, \quad (5)$$

gde važi:  $\tilde{z} \sim N(\bar{z}, \sigma_z^2)$ ,  $\tilde{u}_i \sim N(0, \sigma_u^2)$  i  $cov(\tilde{z}, \tilde{u}_i) = 0$ . Oznake imaju sledeće značenje:  $\beta_i$  je beta koeficijent akcije, koji ukazuje na visinu njenog sistematskog rizika;  $\tilde{z}$  pokazuje visinu agregatnog šoka u vezi sa dividendama, dok je  $\tilde{u}_i$  slučajna greška. Ukupna ponuda je standardizovana i iznosi 1.

Iz formule (5) se može zaključiti da, što je viši beta koeficijent, veći je stepen neslaganja uzajamnih fondova  $\varepsilon$ . Ako je beta koeficijent visok, tržište prvi napuštaju pesimistični uzajamni fondovi, zbog ograničenja prodaje na kratko.

U modelu u kome akteri poseduju konstantnu apsolutnu odbojnost prema riziku i slučajne promenljive imaju normalnu raspodelu (*CARA-Gauss*-ova postavka), očekivana korisnost se može predstaviti kao:

$$E[u(w)|\cdot] = \int -e^{-\gamma w} g(w|\cdot) dw, \quad (6)$$

gde je  $g(w|\cdot)$  uslovna funkcija gustine normalno raspodeljene slučajne promenljive. Moguće je dokazati da je za ovu funkciju očekivane korisnosti ekvivalent sigurnosti<sup>3</sup>:

$$E(w) - \frac{1}{2\gamma} \sigma_w^2 \quad (7)$$

Tržišni učesnici maksimiziraju sledeći ekvivalent sigurnosti:

$$\max_{\tau_i^k} \sum_{i=1}^N \tau_i^k [b_i E^k[\tilde{z}] - (1+r)p_i] - \frac{1}{2\gamma} \left[ \left( \sum_{i=1}^N \tau_i^k b_i \right)^2 \sigma_z^2 + \sum_{i=1}^N (\tau_i^k)^2 (\sigma_\varepsilon^2) \right] \quad (8)$$

uz ograničenje čišćenja tržišta:

$$\mu \left( \frac{1}{2} \tau_i^A + \frac{1}{2} \tau_i^B \right) + (1 - \mu) \tau_i^a = \frac{1}{N}. \quad (9)$$

Oznaka  $\tau_i^k$  predstavlja broj akcija tipa  $i$  kupljenih od strane investitora iz grupe  $k$  ( $k = A, B, a$ ), gde je  $a$  oznaka za arbitražere. Važi i sledeće:

$$(\bar{z} + \varepsilon) b_i - P_i(1+r) = \frac{1}{\gamma} \left( \left( \sum_{k=1}^N b_k \tau_k^A \right) b_i \sigma_z^2 + \tau_i^A \sigma_u^2 \right) \quad (10)$$

$$(\bar{z} - \varepsilon) b_i - P_i(1+r) = \frac{1}{\gamma} \left( \left( \sum_{k=1}^N b_k \tau_k^B \right) b_i \sigma_z^2 + \tau_i^B \sigma_u^2 \right) \quad (11)$$

$$\bar{z} b_i - P_i(1+r) = \frac{1}{\gamma} \left( \left( \sum_{k=1}^N b_k \tau_k^a \right) b_i \sigma_z^2 + \tau_i^a \sigma_u^2 \right) \quad (12)$$

Jednačina određivanja cene akcije bez ograničenja prodaje na kratko ( $i < \bar{i}$ ) je:

$$\bar{z} b_i - P_i(1+r) = \frac{1}{\gamma} \left( b_i \sigma_z^2 + \frac{\sigma_u^2}{N} \right). \quad (13)$$

<sup>1</sup> Broj optimista, označenih sa  $A$ , jednak je broju pesimista, označenih sa  $B$ .

<sup>2</sup> Model je proširen i na *OLG* slučaj, tj. na model sa preklapajućim generacijama (engl. *overlapping generations*). Za više informacija videti [2].

<sup>3</sup> Za dokaz videti [5].

Leva strana jednakosti predstavlja očekivan višak prinosa na akciju  $i$ ,  $\tilde{R}_i^e$ , dok desna strana jednakosti predstavlja premiju za rizik te akcije. Jednačina određivanja cene akcije sa ograničenjem prodaje na kratko (kada važi  $i \geq \bar{i}$ , kao i  $\mu_i^B = 0$ ) je:

$$\bar{z} b_i - P_i(1+r) = \frac{1}{\gamma} (b_i \sigma_z^2 + \frac{\sigma_u^2}{N}) - \pi^i. \quad (14)$$

U ovoj jednakosti  $\pi^i$  predstavlja špekulativnu premiju. Cena akcija u slučaju ograničenja prodaje na kratko je viša nego kad tog ograničenja nema, dok je prinos na tu akciju niži. Špekulativna premija je rastuća funkcija beta koeficijenta:

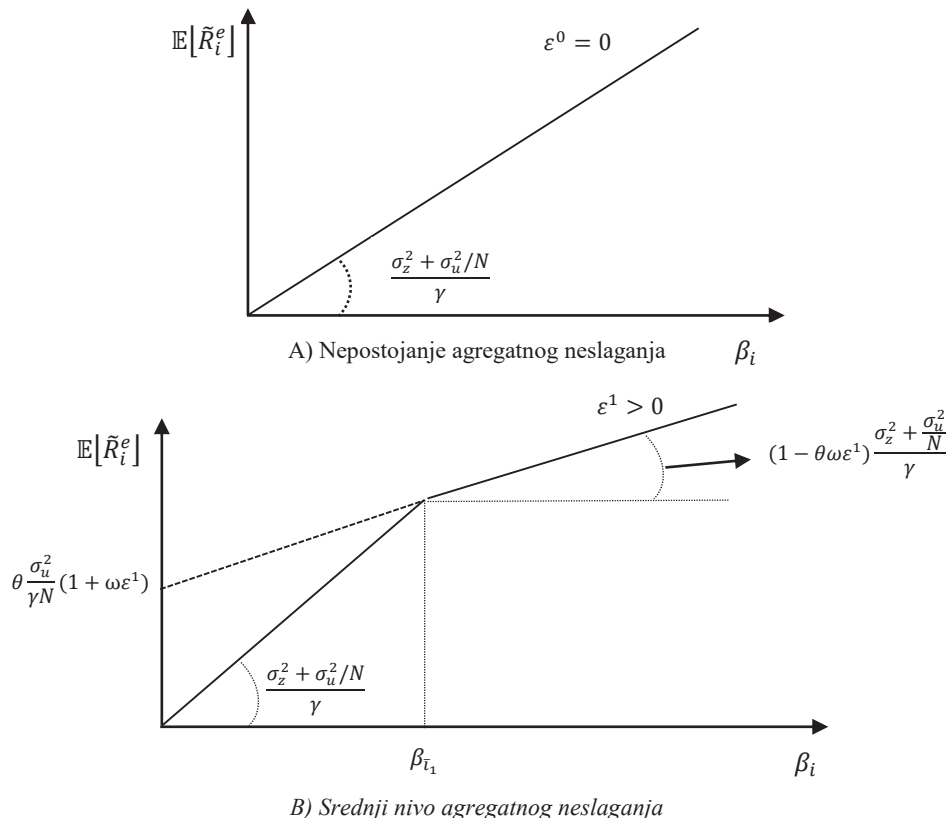
$$\pi^i = \frac{\theta}{\gamma} \left( b_i \sigma_z^2 \omega(\varepsilon) - \frac{\sigma_u^2}{N} \right), \quad (15)$$

gde je

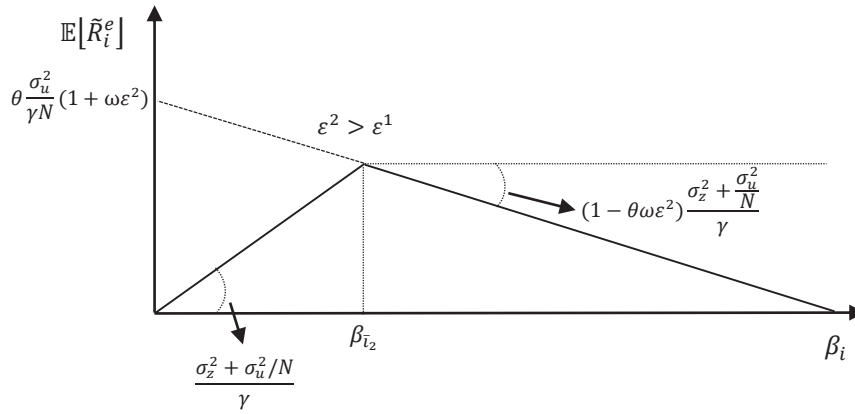
$$\theta = \frac{\frac{\mu}{2}}{1 - \frac{\mu}{2}}, \quad (16)$$

$$\omega(\varepsilon) = \frac{\varepsilon \gamma - \frac{\sigma_z^2}{N} (\sum_{i \geq \bar{i}} b_i)}{\sigma_z^2 (1 + \sigma_z^2 \sum_{i < \bar{i}} \frac{b_i^2}{b_u^2})} \quad (17)$$

Kao što je već rečeno, što je viši beta koeficijent akcija, veći je stepen neslaganja investitora u vezi sa novčanim tokovima te akcije. Nakon određene visine  $\beta$  koeficijenta ( $\beta_i \geq \beta_{\bar{i}}$ ), stepen neslaganja investitora je dovoljno veliki da bi optimalni potez pesimističnih investitora bio da prodaju akcije na kratko. Međutim, to ne mogu da urade zbog ograničenja prodaje na kratko. Sledi da akcije sa visokim beta koeficijentom sadrže špekulativnu premiju, jer njihova cena disproportionalno odražava verovanja optimista. Pored uobičajenih motiva podele rizika i likvidnosnih motiva, kod akcija sa visokim  $\beta$  koeficijentom je izražen špekulativni motiv trgovanja. Otuda će ove akcije imati višu cenu, niži prinos i viši promet. Špekulativna premija  $\pi^i$  meri stepen precenjenosti zbog ograničenja prodaje na kratko. Cene akcija sa niskim vrednostima  $\beta$  koeficijenta su korektno određene, u skladu sa *CAPM*. Na narednom grafikonu je prikazana *SML* funkcija za akcije sa različitim visinom  $\beta$  koeficijenta.







C) Visok nivo agregatnog neslaganja

**Grafikon 2:** SML linija za različite nivoe agregatnog neslaganja

Izvor: [2], p. 2105

Na delu *A* grafikona 2 prikazana je *SML* linija u slučaju nepostojanja agregatnog neslaganja, kada je  $\varepsilon^0 = 0$ . Nagib *SML* linije predstavlja premiju za tržišni rizik. Sa  $\tilde{R}_i^e$  se označava višak prinosa po akciji za akciju *i*:

$$\tilde{R}_i^e = b_i \tilde{z} + \tilde{u}_i - (1+r)P_i, \quad (18)$$

dok  $\tilde{R}_m^e$  predstavlja višak prinosa po akciji za tržišni portfolio, koji predstavljamo kao sumu viškova prinosa pojedinačnih akcija, čiji broj,  $s_j$  je normalizovan na  $1/N$ :

$$\tilde{R}_m^e = \sum_{i=1}^N s_i \tilde{R}_i^e = \sum_{i=1}^N \frac{1}{N} \tilde{R}_i^e = \tilde{z} + \sum_{i=1}^N \frac{\tilde{u}_i}{N} - (1+r)P_m. \quad (19)$$

Beta koeficijent akcije *i* jednak je:

$$\beta_i = \frac{\sigma_{i,m}}{\sigma_m^2} = \frac{b_i \sigma_z^2 + \frac{\sigma_u^2}{N}}{\sigma_z^2 + \frac{\sigma_u^2}{N}}. \quad (20)$$

Očekivan višak prinosa po akciji *i* je dat sledećim formulama:

$$\mathbb{E}(\tilde{R}_i^e) = \begin{cases} \beta_i \frac{\sigma_z^2 + \frac{\sigma_u^2}{N}}{\gamma}, & \text{za } i < \bar{i} \\ \beta_i \frac{\sigma_z^2 + \frac{\sigma_u^2}{N}}{\gamma} (1 - \theta \omega(\varepsilon)) + \theta \frac{\sigma_u^2}{\gamma N} (1 + \omega(\varepsilon)), & \text{za } i \geq \bar{i} \end{cases} \quad (21)$$

U situaciji niskog nivoa agregatnog neslaganja, nagib *SML* funkcije je je  $\frac{\sigma_z^2 + \frac{\sigma_u^2}{N}}{\gamma}$ . Očekivan prinos na akcije

zavisi samo od podele rizika. Ova situacija obuhvata i podsituaciju kada je udeo uzajamnih fondova,  $\mu = 0$ , kada ograničenje prodaje na kratko nije relevantno. Nagib prve funkcije iz formule (21) je jednak desnoj strani jednačine (13) za  $b_i = 1$ . To je situacija kad investitor drži portfolio čija je struktura jednaka strukturi tržišnog portfolija. Tada je očekivan višak prinosa na akciju *i* proporcionalan fundamentalnom, tj. sistematskom riziku, dok je nesistematski rizik u potpunosti diversifikovan.

Na delu *B* grafikona 2 je prikazan slučaj srednjeg nivoa agregatnog neslaganja, gde je agregatno neslaganje  $\varepsilon^1 > 0$ . Udeo uzajamnih fondova, koji ne mogu da prodaju na kratko, je veći od 0 ( $\mu > 0$ ). Kada se  $\beta$  koeficijent poveća, tako da je agregatno neslaganje dovoljno veliko,  $i \geq \bar{i}$ , očekivan prinos na akciju *i* zavisi od nivoa agregatnog neslaganja  $\varepsilon$ , što se može uočiti iz drugog dela formule (21). *SML* funkcija menja i nagib i odsečak. Sada na očekivan višak prinosa utiče i špekulacija, a pesimistični uzajamni fondovi su istisnuti sa tržišta. Manji nagib funkcije podrazumeva da akcije sa višim beta koeficijentom ostvaruju niži očekivani prinos, nego što je

to slučaj prema CAPM modelu. Za dati nivo beta koeficijenta očekivani prinos je manji, jer se povećava špekulativna premija. Dakle, očekivani prinos raste sa povećanjem beta koeficijenta, ali po manjoj stopi nego u situaciji kada nema agregatnog neslaganja. Odsečak na  $y$ -osi se povećava kako se povećava nivo agregatnog neslaganja  $\varepsilon$ .

Na delu C grafikona 2 prikazana je situacija sa najvišim nivoom agregatnog neslaganja  $\varepsilon^2 > \varepsilon^1$ , nastalog usled povećanja heterogenosti u očekivanjima investitora. Kada beta koeficijent pređe određenu granicu,  $\beta_{\frac{1}{2}}$ , SML linija postaje opadajuća, što je u suprotnosti sa CAPM. Sa porastom beta koeficijenta dolazi do opadanja očekivanog prinosa. Zbog daljeg povećanja špekulativne premije, za dati iznos beta koeficijenta, očekivan prinos je još manji u odnosu na situacije sa nižim nivoima agregatnog neslaganja.

### 3. ZAKLJUČAK

U radu je prikazan niz modela koji, kroz uvažavanje heterogenih očekivanja investitora i njihovog neslaganja u pogledu kretanja fundamentalnih pokazatelja, objašnjavaju formiranje cena akcija. Naime, bihevioralne finansije stoje na stanovištu da je nastanak cenovnih balona moguć, zbog toga što tržišni akteri imaju heterogena očekivanja, usled asimetrične informisanosti ili, pak, usled percipiranja i tumačenja istih informacija na različite načine. Za razliku od tradicionalnih finansija, bihevioralne finansije ne smatraju da je arbitraža nerizična i besplatna, a uvažavaju i ograničenja prodaje na kratko. Zbog činjenice da iracionalni investitori stvaraju rizik, kroz mogućnost daljeg produblivanja razlike između cene i fundamentalne vrednosti, često se dešava da se arbitražeri povlače i da na tržištu ostaju samo iracionalni investitori, koji su „kreirali prostor“ za sebe. Otuda neravnoteža, tj. pogrešne cene akcija, mogu dugo da potraju i da poprime karakteristike cenovnih balona.

### ZAHVALNICA

Autor izražava zahvalnost za finansijsku podršku od strane Ministarstva obrazovanja, nauke i tehnološkog razvoja Republike Srbije.

### LITERATURA

- [1] Chen, J., Hong, H., Stein, J. (2002). Breadth of ownership and stock returns. *Journal of Financial Economics*, 66, 171-205.
- [2] Hong, H., Sraer, D. (2016). Speculative Betas. *Journal of Finance*, 71 (5), 2095-2144.
- [3] Hong, H., Stein, J. (2007). Disagreement and the Stock Market. *Journal of Economic Perspectives*, 27 (2), 109-128.
- [4] Miller, E. M. (1977). Risk, Uncertainty, and Divergence of Opinion. *Journal of Finance*, 32 (4), 1151-1168.
- [5] [https://scholar.princeton.edu/sites/default/files/03\\_ree.pdf](https://scholar.princeton.edu/sites/default/files/03_ree.pdf) (pristupljeno 25. 6. 2022. godine)

## IDENTIFIKACIJA I METODOLOGIJA PROCENE NEMATERIJALNIH ULAGANJA U BANKAMA

### IDENTIFICATION AND VALUATION METHODOLOGY FOR THE BANKS' INTANGIBLE ASSETS

DEJAN KOKANOVIĆ<sup>1</sup>, OGNJEN VASILJEVIĆ<sup>2</sup>, NINA MILENKOVIĆ<sup>3</sup>

<sup>1</sup> KPMG d.o.o. Beograd, dkokanovic@kpmg.com

<sup>2</sup> KPMG d.o.o. Beograd, ovasiljevic@kpmg.com

<sup>3</sup> KPMG d.o.o. Beograd, nmilenkovic@kpmg.com

**Rezime:** Transakcije vezane za kapital banaka u poslednje vreme sve su češće u našem regionu. Nakon transakcije, banka sticalac kapitala dužna je (u skladu sa MSFI 3) da uradi alokaciju kupovne cene na fer vrednosti svih sredstava kupljene banke, kako ona koja su zatečena u bilansima na datum transakcije, tako i nematerijalna ulaganja koja se mogu identifikovati, ali nisu mogla interno biti knjigovodstveno evidentirana. Pored nematerijalnih ulaganja koja se mogu očekivati i u proizvodnim preduzećima, kao što su brend, odnosi sa klijentima, specijalni softveri i sl, postoje i sredstva specifična za poslovanje banaka, kao što su tzv. „core“ depoziti. Identifikacija zavisi od tipa banke, vrste klijenata, načina poslovanja, značaja konkretnog nematerijalnog ulaganja, itd. Brend se najčešće procenjuje metodom uštede na royalty-ju, odnosi sa klijentima metodom viška dobiti, a većina ostalih nematerijalnih ulaganja ili metodom viška dobiti ili metodom diskontovanog diskrecionog novčanog toka. Dodatna specifičnost kod banaka je što se i novčani tokovi i diskontna stopa odnose na sopstveni kapital.

**Ključne reči:** Nematerijalna ulaganja, alokacija kupoprodajne cene, metod uštede na royalty-ju, metod viška dobiti, metod diskontovanja diskrecionog novčanog toka..

**Abstract:** Transactions related to banks' equity have recently become more frequent in our region. In accordance with IFRS 3, the acquiring bank is obliged to perform the purchase price allocation to the fair values of all assets of the acquired bank, such as those existing in the balance sheets on the transaction date, as well as intangible assets that can be identified, but could not be internally recorded in the bank's books. In addition to intangibles that can be also expected in manufacturing companies, such as brand, customer relationships, special software, etc., there are also intangibles specific to bank's operations, e.g the core deposits. Intangibles' identification depends on the type of bank, type of clients, way of doing business, importance of specific intangible investment, etc. A brand is most often valued using the relief from royalty method, customer relationships using the Multi-period Excess Earnings Method (MEEM), and most other intangible assets either by MEEM or by the discounted discretionary cash flow method. An additional specificity for banks is that both cash flows and the discount rate refer to equity.

**Keywords:** Intangible Assets, Purchase Price Allocation, Relief from Royalty Method, Multi-period Excess Earnings Method, Discounted Discretionary Cash Flow Method



## KRATKOROČNE FLUKTUACIJE REALNOG EFEKTIVNOG DEVIZNOG KURSA ZEMALJA ZAPADNOG BALKANA KAO INDIKATOR DEVIZNOG RIZIKA U FINANSIRANJU MEĐUNARODNOG POSLOVANJA

### SHORT-TERM REER FLUCTUATIONS OF WESTERN BALKAN COUNTRIES AS AN INDICATOR OF FOREIGN EXCHANGE RISK IN FINANCING INTERNATIONAL BUSINESS

ALEKSANDRA ĐORĐEVIĆ<sup>1</sup>, IVANA POPOVIĆ PETROVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, aleksandra.djordjevic@ekof.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, ivana.popovic@ekof.bg.ac.rs

**Rezime:** U radu se analizira volatilnost realnog efektivnog deviznog kursa sa ciljem dobijanja uvida u stepen u kome su preduzeća izložena deviznom riziku koji se javlja pri finansiranju međunarodnog poslovanja. Uzorkom je obuhvaćeno pet zemalja Zapadnog Balkana za koje je izračunata volatilnost realnih efektivnih deviznih kurseva u periodu 2001-2020. godina, na bazi mesečnih podataka. Rezultati sprovedene analize ukazuju na trend smanjivanja volatilnosti realnih efektivnih deviznih kurseva svih analiziranih zemalja, osim Albanije, tokom razmatranog perioda. Iako dobijeni rezultat sugerise smanjivanje deviznog rizika kome su učesnici u međunarodnom poslovanju zemalja ovog regiona izloženi, trenutne nestabilnosti na globalnom nivou upućuju na neophodnu obazrivost i važnost razumevanja načina na koji se izloženost ovom riziku može umanjiti, o čemu se takođe govori u ovom radu.

**Ključne reči:** volatilnost, realni efektivni devizni kurs, zemlje Zapadnog Balkana, devizni rizik

**Abstract:** The paper analyzes the volatility of the real effective exchange rate in order to gain insight into the degree to which companies are exposed to foreign exchange risk that occurs when financing international business. The sample includes five Western Balkan countries for which the volatility of real effective exchange rates in the period 2001-2020 was calculated, based on monthly data. The results of the conducted analysis indicate a trend of decreasing volatility of real effective exchange rates of all analyzed countries, except Albania, during the period under review. Although the obtained result suggests a reduction in foreign exchange risk to which participants in international business of the region are exposed, current instabilities at the global level indicate the necessary caution and importance of understanding how to reduce exposure to this risk, which is also discussed in this paper.

**Keywords:** volatility, real effective exchange rate, Western Balkan countries, foreign exchange risk

#### 1. UVOD

Tokom druge polovine XX veka dolazi do značajnih promena u sagledavanju potencijalnih izvora finansiranja. Liberalizacija međunarodne trgovine praćena je i liberalizacijom kretanja faktora proizvodnje, rada i kapitala. Povećava se značaj prekograničnih kredita i inostranih kreditnih linija. Zadužujući se na tako disperziranom, međunarodnom tržištu, brojni korisnici, odnosno preduzeća, suočavaju se sa promenama deviznog kursa koje predstavljaju najvažniji neprenosivi rizik finansiranja međunarodnog poslovanja. Ovaj rizik aktuelizovan je u savremenim okolnostima koje karakterišu povećana nestabilnost i neizvesnost na tržištima širom sveta, što motiviše analiziranje kratkoročnih fluktuacija deviznog kursa kao indikatora deviznog rizika u finansiranju međunarodnog poslovanja u ovom radu. Drugim rečima, u pitanju je analiziranje *volatilnosti* deviznih kurseva imajući u vidu da, prema [14], volatilnost deviznih kurseva podrazumeva upravo kratkoročne fluktuacije deviznih kurseva, koje se posmatraju na mesečnom, nedeljnom ili čak dnevnom nivou. U radu se ispituje volatilnost *realnih efektivnih deviznih kurseva* imajući u vidu da ova vrsta deviznih kurseva predstavlja sveobuhvatni pokazatelj koji pruža jasnu sliku vrednosti domaće valute.

Analiza se odnosi na pet zemalja Zapadnog Balkana u periodu od 2001. do 2020. godine. U pitanju su: Albanija, Bosna i Hercegovina, Severna Makedonija, Srbija i Crna Gora. Praćenjem volatilnosti realnih efektivnih deviznih kurseva zemalja Zapadnog Balkana u ovom radu nastojimo da procenimo devizni rizik sa

kojim se učesnici u međunarodnom poslovanju ovih zemalja susreću, što predstavlja zanemareno pitanje u dosadašnjoj literaturi. Imajući u vidu značajnu geografsku orijentaciju srpskih izvoznika i uvoznika prema tržištu potpisnica CEFTA 2006 sporazuma, poseban doprinos ovog rada iz ugla srpskih učesnika u međunarodnom poslovanju se ogleda u analiziranju volatilnosti realnih efektivnih deviznih kurseva upravo potpisnica ovog sporazuma sa kojima Srbija ostvaruje najveći obim razmene, poput Bosne i Hercegovine, Severne Makedonije, Crne Gore, ali i Albanije, sa kojom se manje trguje [10].

S obzirom da je u međunarodnom poslovanju sve više poslova koji su praćeni kratkoročnim, kako finansiranjem, tako i kreditiranjem poslovnih partnera, praksa međunarodnog poslovanja je ukazala na brojne mogućnosti prevazilaženja rizika promene deviznog kursa, od kojih će neke biti razmatrane u ovom radu.

## 2. METODOLOGIJA I PODACI

Prilikom izračunavanja volatilnosti deviznog kursa osnovnu metodološku nedoumicu predstavlja izbor odgovarajućeg načina merenja ove vrste promena deviznih kurseva, odnosno izbor merila koje bi na najadekvatniji način odražavalo rizik, neizvesnost i troškove prilagođavanja kojima su preduzeća izložena kada nastupi promena deviznih kurseva.

I pored velikog broja teorijskih i empirijskih radova koji analiziraju volatilnost deviznog kursa, do danas nije postignuta saglasnost po pitanju najispravnijeg načina merenja ove vrste promena deviznih kurseva, što je posledica nekoliko faktora. Naime, ne postoji univerzalno prihvaćen model reakcije preduzeća na rizik koji proizilazi iz fluktuacija deviznih kurseva. Posledično, teorija nije u mogućnosti da odredi metod koji je najadekvatniji za merenje volatilnosti. Osim toga, izbor odgovarajućeg načina izračunavanja volatilnosti zavisi od empirijskog okvira istraživanja. Ukoliko su predmet analize razvijene zemlje, korišćenje podataka sa terminskih tržišta za procenu volatilnosti deviznog kursa je opravdanije nego u slučaju kada se analiza, u celini ili najvećim delom, odnosi na zemlje u razvoju. Pored toga, mora se uzeti u obzir i vremenski horizont za koji se izračunava volatilnost kursa, kao i da li se analiza sprovodi na agregiranim ili sektorskim podacima [6]. Prethodno navedeni faktori rezultirali su različitim pristupima za izračunavanje volatilnosti kursa.

U ranoj literaturi su za izračunavanje volatilnosti najčešće korišćene metode koje se zasnivaju na primeni promptnih deviznih kurseva. Tako, [13] i [1] su koristili procentualnu promenu promptnog deviznog kursa. U svom kasnijem radu, [12] koriste varijansu promptnog deviznog kursa oko predviđenog trenda. Dalje, [8] volatilnost deviznog kursa takođe izračunavaju na bazi promptnih devizni kurseva i to kao procentualnu promenu između maksimalnog i minimalnog promptnog deviznog kursa tokom određenog vremenskog perioda, uvećanu za meru neuravnoteženosti deviznog kursa, što nazivaju merom dugoročne neizvesnosti.

Za razliku od prethodno pomenutih, prema jednom od pristupa za izračunavanje volatilnosti, ova vrsta promena deviznih kurseva predstavlja izvor neizvesnosti i rizika u onom stepenu u kome se ona ne može predvideti. U tom smislu, prema ovom delu literature, predvidivost promena deviznih kurseva predstavlja važan faktor u izračunavanju volatilnosti, te pristupi koji se baziraju na ovom stanovištu nastoje da prilikom izračunavanja volatilnosti uključe predviđene vrednosti deviznih kurseva. Ovi pristupi baziraju se na primeni terminskih deviznih kurseva i ARCH i GARCH modela [9, 3]. Međutim, imajući u vidu da je predviđanje deviznih kurseva podložno kritici, što su tvrdili još [7], ovi metodi se mogu koristiti kao alternativni, a ne kao osnovni prilikom merenja deviznog rizika.

Pored prethodno navedenih, u pristupe za izračunavanje volatilnosti kursa spadaju i oni koji se zasnivaju na upotrebi standardne devijacije. Pregledom literature može se ustanoviti da su upravo oni najzastupljeniji u empirijskim radovima, posebno kada je reč o novijoj literaturi iz ove oblasti. Pristup koji je stekao široku primenu u empirijskim radovima podrazumeva da se volatilnost deviznog kursa izračunava kao standardna devijacija prve diference logaritma deviznog kursa, što se prema [11] može prikazati na sledeći način:

$$VOL_{it} = Std. dev. [\ln(e_{it,m}) - \ln(e_{it,m-1})] \quad (1)$$

gde je:

$VOL_{it}$  volatilnost realnog efektivnog deviznog kursa zemlje  $i$  u godini  $t$   
 $e_{it,m}$  mesečni realni efektivni devizni kurs zemlje  $i$

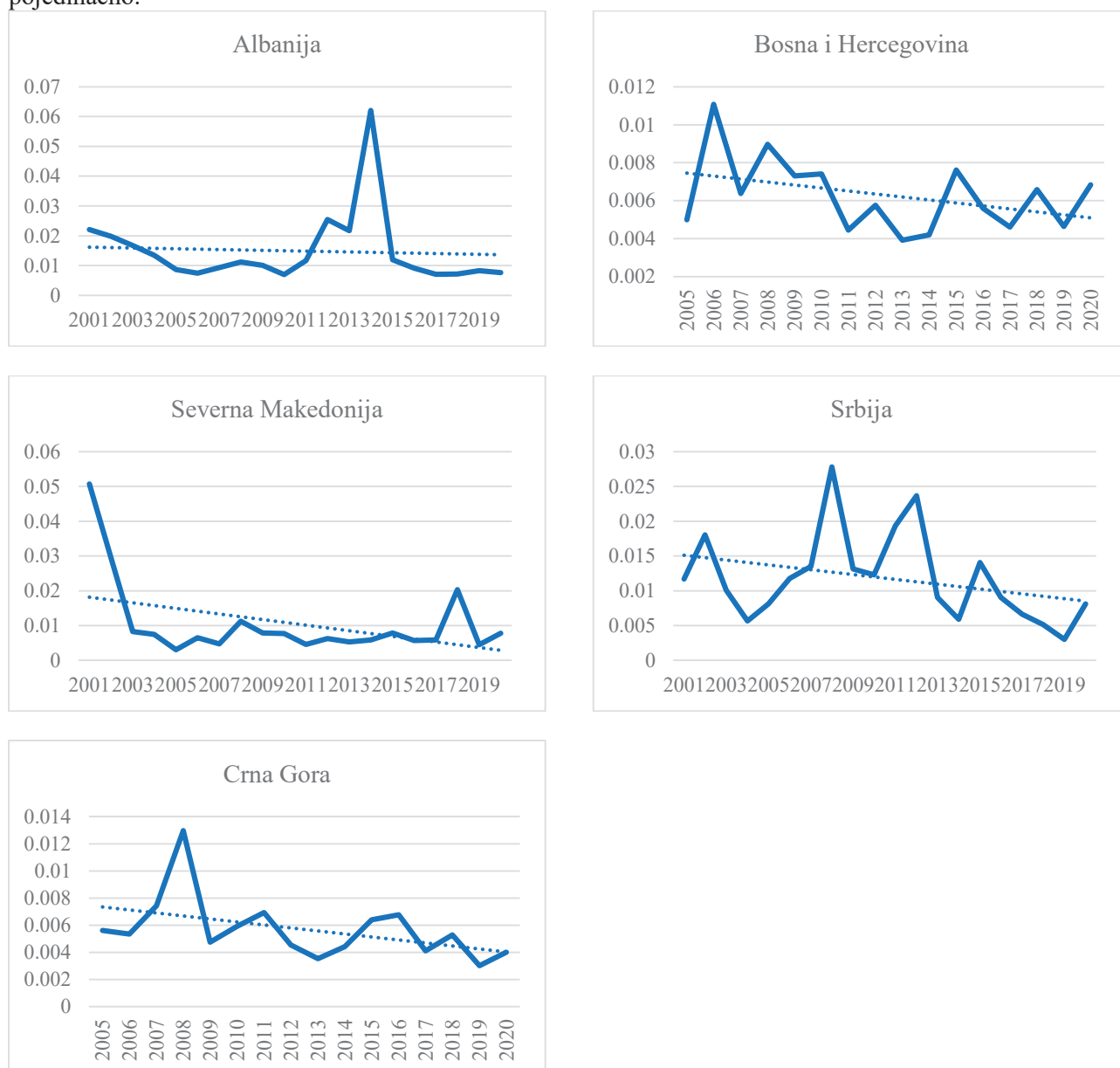
Ukoliko se, na prethodno definisan način, volatilnost računa za period od godinu dana ( $m = 1 \dots 12$ ), dobija se mera kratkoročne volatilnosti, što ovaj pristup čini pogodnim za izračunavanje volatilnosti za potrebe istraživanja u ovom radu.

Polazeći od relacije (1), za merenje volatilnost realnog efektivnog deviznog kursa zemalja Zapadnog Balkana u ovom radu je korišćena standardna devijacija prve diference logaritma realnog efektivnog deviznog kursa izračunata za period od godinu dana na bazi mesečnih deviznih kurseva.

Podaci o mesečnim realnim efektivnim deviznim kursovima u periodu januar 2001 – decembar 2020. godine za pet zemalja Zapadnog Balkana (Albanija, Bosna i Hercegovina, Severna Makedonija, Srbija i Crna Gora) prikupljeni su iz CEPII EQCHANGE baze podataka. Ova baza sadrži podatke o realnim efektivnim deviznim kursovima koji su izračunati primenom različitih pristupa u pogledu broja trgovinskih partnera koji čine korpu za obračun vrednosti valute i primenjenog sistema ponderisanja. Za potrebe analize u ovom radu izabran je realni efektivni devizni kurs koji je izračunat na osnovu korpe u kojoj se nalazi 30 trgovinskih partnera zemlje, primenom time-varying weights koji su bazirani na non-overlapping five-year average weights. Time je izabran realni efektivni devizni kurs koji je izračunat u odnosu na najvažnije trgovinske partnere svake zemlje (u poređenju sa efektivnim deviznim kursovima koji su izračunati na osnovu korpe u kojoj se nalazi 186 trgovinskih partnera), dok primena vremenski promenljivih pondera omogućava uzimanje u obzir kako trenutnih obrazaca trgovine, tako i onih iz prošlosti, što predstavlja prednost ovog pristupa.

### 3. REZULTATI

Primenom pristupa koji je predstavljen u delu 2 ovog rada dobijeni su rezultati o volatilnosti realnih efektivnih deviznih kurseva zemalja Zapadnog Balkana koji su ilustrativno predstavljeni na Grafikonu 1 za svaku zemlju pojedinačno.



**Grafikon 1:** Volatilnost realnog efektivnog deviznog kursa zemalja Zapadnog Balkana u periodu 2001 – 2020. godina

Napomene: (1) Isprekidana linija predstavlja trend liniju; (2) Za Bosnu i Hercegovinu i Crnu Goru volatilnost je izračunata za period 2005-2020. godina usled nedostupnosti podataka za period pre 2005. god.

Izvor: Kalkulacija i prikaz autora na osnovu podataka iz CEPII EQCHANGE baze podataka [15]

Na osnovu Grafikona 1 primetan je trend smanjivanja volatilnosti realnih efektivnih deviznih kurseva svih analiziranih zemalja osim Albanije tokom razmatranog perioda. Ovaj rezultat je važan jer ukazuje na smanjivanje deviznog rizika kome su izvoznici ovih zemalja izloženi. Iako dobijeni rezultat ukazuje na pozitivnu tendenciju, treba biti obazriv imajući u vidu trenutne nestabilnosti na globalnom nivou. Sa tim u vezi, važno je razumeti kako se rizici promene deviznog kursa mogu prevazići i koji su načini zaštite od ovog rizika, o čemu govorimo u nastavku.

#### **4. PREVAZILAŽENJE RIZIKA PROMENE DEVIZNIH KURSEVA U FINANSIRANJU MEĐUNARODNOG POSLOVANJA**

Rizik promene deviznog kursa pripada kategoriji neprenosivih rizika, s obzirom da se ne mogu preneti na druge učesnike u poslu, ili pak treću stranu, uglavnom osiguravače. Ovo su rizici kojima su izloženi kako finansijeri, tako i korisnici finansiranja, mada se smatra da su ipak korisnici, odnosno, dužnici, više izloženi neprenosivim rizicima, a prenosivim u manjoj meri, za razliku od finansijera [5]. Osim toga, rizik promene deviznog kursa raste sa povećanjem vremenskog roka u kom se očekuje konačna realizacija dogovorenog posla. Iz ovih razloga je važno razumeti izvore deviznog rizika u finansiranju međunarodnom poslovanju kao i načine da se u tim slučajevima rizik promene deviznog kursa prevaziđe.

Pored brojnih, opštih rizika koje donosi poslovanje u međunarodnom okruženju, u domenu deviznih rizika izdvajaju se: različitost nacionalnih valuta, zatim zastupljenost konvertibilnih valuta, kao i različitost deviznih sistema. Upravo devizni sistemi regulišu način poslovanja devizama, počev od načina obavljanja uplata i isplata između rezidenata iz različitih zemalja, preko praćenja deviznih rezervi pojedinačnih zemalja, ali i obavljanja devizne kontrole [2]. Konvertibilnim se nazivaju one valute kojima se može slobodno trgovati na deviznom tržištu, dok se na drugoj strani nalaze valute kojima se, ili ne trguje slobodno na deviznom tržištu, ili je to zabranjeno domaćim propisima, ili pak nema zainteresovanih učesnika koji bi čuvali neku vrednost u njima. Na toj strani nalaze se nekonvertibilne, ili mekane valute. Upravo ova kategorija valuta otežava realizaciju posla sa elementima inostranosti [4]. U slučaju da oba učesnika u poslu sa elementom inostranosti, raspolažu nekonvertibilnim valutama, eventualni rizik od promene deviznih kurseva mogu umanjiti ugovaranjem posla u zajednički odabranoj konvertibilnoj valuti.

Korisnici finansiranja se najviše suočavaju sa rizicima promene deviznog kursa, usled činjenice da im je odobren komercijalni kredit, ili su se odlučili za uvoz robe na kredit, a dolaze iz zemalja nekonvertibilnih valuta. Njihov najveći rizik proizilazi iz potencijalnog rasta deviznog kursa konvertibilne valute u kojoj su ugovorili posao, što bi iziskivalo izdvajanje znatno veće količine domaće valute za kupovinu inostrane valute kojom bi se otplatio kredit. Mogućnost da se uvoznik, odnosno, korisnik finansiranja zaštiti od rasta deviznog kursa jeste da pristupi hedžingu, odnosno izbegavanju rizika deviznog kursa.

U cilju izbegavanja rizika promene deviznog kursa korisnik finansiranja, odnosno, uvoznik u spoljnotrgovinskom poslu, mogao bi sopstvenim ili pozajmljenim sredstvima da kupi potreban iznos inostrane valute, po tekućem deviznom kursu, što podrazumeva trošak u vidu blokiranja sopstvenih sredstava, odnosno kamate koja se plaća na pozajmljena sredstva. Da bi se ove vrste troškova radi osiguranja od promene deviznog kursa ipak izbegle, uvoznik hedžing može da obavi na terminskom deviznom tržištu. U tom smislu, moguće je koristiti terminski ugovor. Drugim rečima, terminski devizni kursevi mogu doprineti zaštiti od rizika promene deviznog kursa. Stupanjem u terminski ugovor, uvoznik stiče mogućnost da kupi potreban iznos inostrane valute po unapred poznatom (terminskom) kursu, tačno utvrđenog dana u budućnosti. Iznos u domaćoj valute isplaćuje se u trenutnu dospeća ugovora čime se izbegava blokiranje sopstvenih sredstava. Pored ovog načina, uvoznik se može zaštiti od rizika promene deviznog kursa i pomoću kupovine fjučers ugovora na određeni, tačno definisani datum u budućnosti. Osim toga, uvoznik ima i mogućnost kupovine *call* opcije u cilju zaštite od rizika promene deviznog kursa, koja mu daje pravo ali ne i obavezu da određenog dana u budućnosti kupi tačno utvrđeni iznos strane valute po unapred dogovorenom deviznom kursu [4].

#### **5. ZAKLJUČAK**

Učesnici u međunarodnom poslovanju izloženi su brojnim rizicima. Među njima se posebno ističu rizici povezani sa promenama deviznih kurseva. Promene deviznih kurseva, kako kratkoročne, tako i dugoročne, izazivaju višestruke efekte, koji se međusobno prepliću. Kratkoročne fluktuacije deviznih kurseva, odnosno volatilnost, predstavljaju neprenosivi rizik za učesnike u međunarodnom poslovanju, čime ova vrsta promena deviznih kurseva predstavlja važan izvor neizvesnosti. Imajući to u vidu, predmet istraživanja ovog rada bilo je analiziranje volatilnosti realnih efektivnih deviznih kurseva zemalja Zapadnog Balkana sa ciljem procenjivanja deviznog rizika kome su učesnici u međunarodnom poslovanju ovih zemalja izloženi.



Na osnovu mesečnih podataka za period od januara 2001. do decembra 2020. godine, u radu je izračunata godišnja volatilnost realnih efektivnih deviznih kurseva za Albaniju, Bosnu i Hercegovinu, Severnu Makedoniju, Srbiju i Crnu Goru. Rezultati su pokazali da je u većini zemalja primetan trend smanjivanja volatilnosti u analiziranom periodu, što predstavlja pozitivnu karakteristiku i ukazuje na manji stepen neizvesnosti kada je reč o deviznim rizicima. Izuzetak je Albanija usled ekstremne vrednosti u 2014. godini. Ukoliko bi se ova ekstremna vrednost eliminisala, volatilnost realnog efektivnog deviznog kursa Albanije bi takođe imala opadajući trend.

Povećana nestabilnost na tržištima širom sveta ukazuje na moguće rizika u vidu promene trendova koji su karakterisali prethodni period, što treba imati u vidu. Povezano sa tim, učesnici u međunarodnom poslovanju moraju biti pripremljeni da se na odgovarajući način zaštite od rizika promene deviznih kurseva, o čemu je takođe bilo reči u ovom radu.

## LITERATURA

- [1] Bailey, M. J., Tavlas, G. S., & Ulan, M. (1987). The impact of exchange-rate volatility on export growth: some theoretical considerations and empirical results. *Journal of Policy Modeling*, 9(1), 225-243.
- [2] Bjelica, V., Raičević, B., Babić, B., Radmilović, S. & Radičić, M. (2001). *Finansije—teorija i praksa*. Ekonomski fakultet u Subotici, Univerzitet u Novom Sadu, Novi Sad.
- [3] Clark, P., Tamirisa, N., Wei, S. J., Sadikov, A., & Zeng, L. (2004). Exchange rate volatility and trade flows—some new evidence. *IMF Occasional Paper*, 235. International Monetary Fund, Washington D.C.
- [4] Kovačević, R. (2016). *Međunarodne finansije*. Centar za izdavačku delatnost Ekonomskog fakulteta, Beograd.
- [5] Kozomara, J. (2019). *Finansiranje međunarodnog poslovanja*. Centar za izdavačku delatnost Ekonomskog fakulteta, Beograd.
- [6] McKenzie, M. D. (1999). The impact of exchange rate volatility on international trade flows. *Journal of economic Surveys*, 13(1), 71-106.
- [7] Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample?. *Journal of international economics*, 14(1-2), 3-24.
- [8] Perée, E., & Steinherr, A. (1989). Exchange rate uncertainty and foreign trade. *European Economic Review*, 33(6), 1241-1264.
- [9] Péridy, N. (2003). Exchange rate volatility, sectoral trade, and the aggregation bias. *Review of World Economics*, 139(3), 389-418.
- [10] Popović Petrović, I. (2018). CEFTA 2006: Trade Structure, Flows and Competitiveness. In: Chizhikov Sergei, Dmitriev Andrei and Kabylnskii Boris (Eds.) (2018), *Development of Foreign Trade in Modern World: Innovations and Challenges*. (pp. 95-100). North West Institute of Management (NWIM) of Russian Presidential Academy of National Economy and Public Administration, Saint Petersburg.
- [11] Tenreyro, S. (2007). On the trade impact of nominal exchange rate volatility. *Journal of Development Economics*, 82(2), 485-508.
- [12] Thursby, J. G., & Thursby, M. C. (1987). Bilateral trade flows, the Linder hypothesis, and exchange risk. *The Review of Economics and Statistics*, 488-495.
- [13] Thursby, M. C., & Thursby, J. G. (1985). The uncertainty effects of floating exchange rates: Empirical evidence on international trade flows. In: Sven W. Amtd., Richard J. Sweeney & Thomas D. Willett (Eds.) (1985). *Exchange rates, trade and the US economy*. (pp. 153-166). Cambridge: Ballinger.
- [14] Williamson, J. (1985). *The Exchange Rate System*. Washington D.C.: Institute for International Economics.
- [15] [http://www.cepii.fr/CEPII/en/bdd\\_modele/bdd\\_modele\\_item.asp?id=34](http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=34).



## OPTIMIZACIJA PORTFOLIJA PRIME LISTING AKCIJA BEOGRADSKE BERZE

### OPTIMIZATION OF THE PORTFOLIO OF PRIME LISTING SHARES OF THE BELGRADE STOCK EXCHANGE

ZORAN POPOVIĆ<sup>1</sup>, ALEKSANDAR BOŽOVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, zoran.popovic@ekof.bg.ac.rs

<sup>2</sup> Komisija za hartije od vrednosti Republike Srbije, acobozovic95@gmail.com

---

**Rezime:** U ovom radu bavili smo se optimizacijom portfolija sačinjenog od četiri akcije kompanija koje su na prime listingu Beogradske berze i to: Aerodrom Nikola Tesla AD, Energoprojekt Industrija AD, NIS AD i Fintel Energija AD. Rizik portfolija merimo standardnom devijacijom i svaki investitor ima određeni stepen averzije prema riziku. Prinos predstavlja zapravo nagradu za rizik koji preuzimaju investitori, veće stope prinosa uglavnom nose i veće stope rizika. Optimizacija portfolija koja se ovde primenjuje odnosi se na minimizaciju rizika za dato ulaganje. Koristeći istorijske podatke za 2019. godinu, postavili smo zadatak optimizacije portfolija preko Markovicevog modela, a zatim kvadratnim programiranjem odredili optimalna ulaganja u svaku od datih akcija.

**Ključne reči:** portfolio, optimizacija, Beogradska berza, kvadratno programiranje

**Abstract:** In this paper, we have dealt with the optimization of the portfolio consisting of four shares of companies listed on the Belgrade Stock Exchange, namely: NIS, Nikola Tesla Airport, Energoprojekt and Fintel. We measure portfolio risk by standard deviation and each investor has a certain degree of risk aversion. Yield is actually a reward for the risk taken by investors, higher rates of return generally carry higher risk rates. The portfolio optimization applied here refers to minimizing the risk for a given investment. Using historical data for 2019, we set the task of portfolio optimization through Markovic's model, and then determined the optimal investments in each of the given actions by quadratic programming.

**Keywords:** portfolio, optimization, Belgrade Stock Exchange, quadratic programming



## ТИПОВИ ПРОМЕНА КОНЦЕНТРАЦИЈЕ У БАНКОВНОМ СЕКТОРУ СРБИЈЕ: ДЕКОМПОЗИЦИЈА ИНДЕКСА ХИРШМАНА-ХЕРФИНДАЛА

### TYPES OF THE CHANGES IN BANKING SECTOR OF SERBIA CONCENTRATION: DECOMPOSITION OF THE HIRSCHMAN-HERFINDAHL INDEX

РАЈКО М. БУКВИЋ<sup>1</sup>

<sup>1</sup> Нижегородский государственный инженерно-экономический университет – Княгинино (Россия), r.bukvic@yandex.com

**Резиме:** На основу података завршних рачуна банака о пет билансних величина (актива, депозити, капитал, пословни приход, кредити) обрачунати су индекси Хиршмана-Херфиндала за период 2016–2021 (I–IX). Добијене вредности декомпоноване су поступком који су предложили Бахо и Салас, чиме је добијен утицај броја банака и дисперзије њихових удела на вредности индекса концентрације. Затим је извршено класификовање промена концентрације у зависности од односа стопа промена двају фактора. Они су били неједнаки и варирали по годинама. Међу издвојеним типовима промена преодлава пад индекса изазван смањењем неједнакости тржишних удела и смањењем броја банака, при чему је прво веће. Најзад, у 2021. смањен је броја банака уз раст дисперзије њихових удела, што је с теоријског становишта јасан услов раста индекса концентрације.

**Кључне речи:** концентрација, банковни сектор, декомпозиција индекса Хиршмана-Херфиндала, број банака, дисперзија тржишних удела, типови промена степена концентрације

**Abstract:** On the basis of five balance variables in bank balances (total assets, deposits, capital, operating income, and loans), the Hirschman–Herfindahl indices for period 2016–2021 (I–IX) are calculated. The indices values are decomposed by the Bajo and Salas approach, and the impacts of number of banks and dispersion of its market shares are established. Then we classified the concentration changes, depending on relations between the rates of changes of two factors. They were unequal and vary during the years. Among the identified types of changes there was the most frequent decrease of HHI, caused by decrease of inequality of market shares and decrease of number of banks, where the first was greater. At the end, in 2021 there was the decrease of banks number with increase of dispersion of its shares, which is in theoretical sense condition for the increase of concentration index.

**Keywords:** concentration, banking sector, decomposition of the index Hirschman–Herfindahl, number of banks, dispersion of market shares, types of the concentration degree changes

## 1. УВОД

Током протеклих неколико година објављен је већи број радова наших аутора посвећених проблемима концентрације (и конкуренције), између осталог и у банковном сектору. У њима је, по правилу, за утврђивање степена концентрације коришћен Хиршман-Херфиндалов индекс (ХХИ), а он је израчунаван и навођен и у радовима који стављају акценат на друге мере концентрације, видети радове овог аутора, рецимо [2 ; 9]. Будући да је овај индекс почев од првих примена, пре више од три четврти века, стекао и по многим мишљењима потврдио репутацију идеалне мере концентрације, његова толико широка примена у истраживањима не представља изненађење; као што изненађење не представља ни његово прихватање и коришћење од стране многих антимонополских органа не само у развијеним земљама (код нас, видети [5]). При томе, занемарују се његови више пута истакнути недостаци, док се с друге стране не користе све могућности које он као аналитичко средство пружа за дубље и свестраније анализе.

У наредним редовима позабавићемо се једном од могућности анализе засноване на одликама ХХИ, која омогућава идентификовање и квантификовање фактора који делују на промене индекса током времена. Саопштење је засновано на ширем истраживању наведене проблематике [3]. Као основ за анализу послужио нам је банковни сектор Србије. Посматран је период 2016–2021, током којег овај аутор континуирано прати проблеме концентрације у наведеном сектору.

## 2. ХИРШМАН-ХЕРФИНДАЛОВ ИНДЕКС КАО ПОСЕБАН СЛУЧАЈ ИНДЕКСА ХАНА-КЕЈА И УОПШТЕНОГ ИНДЕКСА ЕНТРОПИЈЕ

Хиршман-Херфиндалов индекс одређен је као посебан случај индекса концентрације, који се у општем облику представља као

$$HNI = \sum_{i=1}^N w_i s_i \quad (1)$$

у коме су тржишни удели учесника (предузећа)  $s_i$  пондерисани самим тим уделима ( $s_i = w_i$ ). Очигледно је да, за разлику од неких касније формулисаних индекса овог типа, ХХИ даје већи значај снажнијим учесницима на тржишту, што води извесној асиметрији у распореду самих вредности индекса. Да ли је тај факт предност или недостатак индекса, питање је на које нема правог и јединственог одговора. С друге стране, минимална вредност распона у коме се крећу вредности ХХИ [ $1/N$  ; 1] зависи од величине скупа, тј. узорка (броја учесника на тржишту). То је већ недостатак индекса, али се он у већини радова занемарује. Погледајмо шта то конкретно значи. Тако ће и у идеалном случају (једнаке компетенције учесника) минимална вредност индекса (1) бити 0,5 за два учесника, 0,25 за четири, 0,10 за десет, 0,05 за двадесет учесника, итд., дакле добиће се потпуно различите информације о тржишној ситуацији у којој постоји идеална једнакост учесника. Ако се, поред тога, узме у обзир и постојање различите дисперзије тржишних удела (дакле њихове неједнакости), постаће јасно да ХХИ одликује неједнозначна информативност, односно да се за врло различите тржишне ситуације могу добити исте вредности индекса. Овај његов недостатак сасвим јасно се види из деривата основне формуле, до које се долази њеном елементарном трансформацијом

$$HNI = N\sigma^2 + \frac{1}{N} \quad (2)$$

Израз (2) јасно показује да, деловањем двају независних фактора (број учесника на тржишту  $N$  и распоред тржишних удела, тј. њихова варијанса  $\sigma^2$ ) може настати бесконачан број комбинација вредности  $N$  и  $\sigma^2$ , које могу довести до истих вредности индекса ХХИ, како то подвлаче Смарагдов и Сидорејко [4]. Тај израз експлицитно показује да концентрација на тржишту зависи од два фактора, управо онако како је то дефинисао Вотерсон [14], дакле – мера концентрације може бити развијена као функција неједнакости и броја фирми. При томе је дејство другог фактора (дисперзија тржишних удела) директно пропорционално, док дејство првог (број учесника на тржишту) не може јасно да се сагледа. Вредност индекса ће се стога очигледно смањивати са смањивањем дисперзије тржишних удела, а када она има вредност 0 (нула), тј. када сви учесници на тржишту имају једнаке уделе, вредност индекса ће се свести на  $1/N$ , као што смо већ констатовали.

У овом раду нас ће, међутим, интересовати друга консеквенца која проистиче из израза (2), на име могућности да се разграниче утицаји наведених фактора  $N$  и  $\sigma^2$  на кретање ХХИ. Сам израз (2) за тако нешто није адекватан, тако да ћемо морати да учинимо још нека методолошка објашњења.

Присетимо се најпре да је ХХИ специјалан случај индекса Хана-Кеја [11]

$$HKI = \left( \sum_{i=1}^N s_i^\alpha \right)^{\frac{1}{\alpha-1}}, \quad \alpha > 0; \alpha \neq 1, \quad (3)$$

када је  $\alpha=2$ , при чему смо изоставили формулу за случај  $\alpha=1$  који је без значаја за нашу анализу. Како су показали Бахо и Салас [6 ; 7 ; 8], индекс Хана-Кеја (3) може бити представљен и преко уопштеног индекса ентропије у облику израза (4), где је опет изостављен случај  $\alpha = 1$ :

$$HKI = \frac{[1 + \alpha(\alpha-1)GEI]^{\frac{1}{\alpha-1}}}{N}, \quad \alpha > 0; \alpha \neq 1 \quad (4)$$

где је  $GEI$  – уопштени индекс ентропије, одређен даље изразом (7), уз изостављање случаја  $\alpha = 0$  и  $\alpha = 1$ , који немају значаја за индекс Хиршмана-Херфиндала, тј. у општем случају:

$$HKI(\alpha) = \frac{\varphi[GEI(\alpha)]}{N}, \quad \forall \alpha > 0, \quad (5)$$

где је  $\varphi[GEI(\alpha)]$  – компонента неједнакости у (4), као растућа функција уопштеног индекса ентропије.

Из (5) следи да промена у концентрацији може да буде декомпонована на два дела

$$\frac{\Delta HKI(\alpha)}{HKI(\alpha)} \approx \frac{\Delta \varphi[GEI(\alpha)]}{\varphi[GEI(\alpha)]} - \frac{\Delta N}{N} \quad (6)$$

Наведимо најзад и формулу уопштеног индекса ентропије, према [10]

$$GEI = \frac{1}{N} \frac{1}{\alpha(\alpha-1)} \sum_{i=1}^N \left[ \left( \frac{S_i}{\bar{S}} \right)^\alpha - 1 \right], \quad \forall \alpha \neq 0, 1 \quad (7)$$

где  $\bar{S}$  означава аритметичке средину тржишних удела.

Ако се у (7) стави  $\alpha=2$ , добиће се:

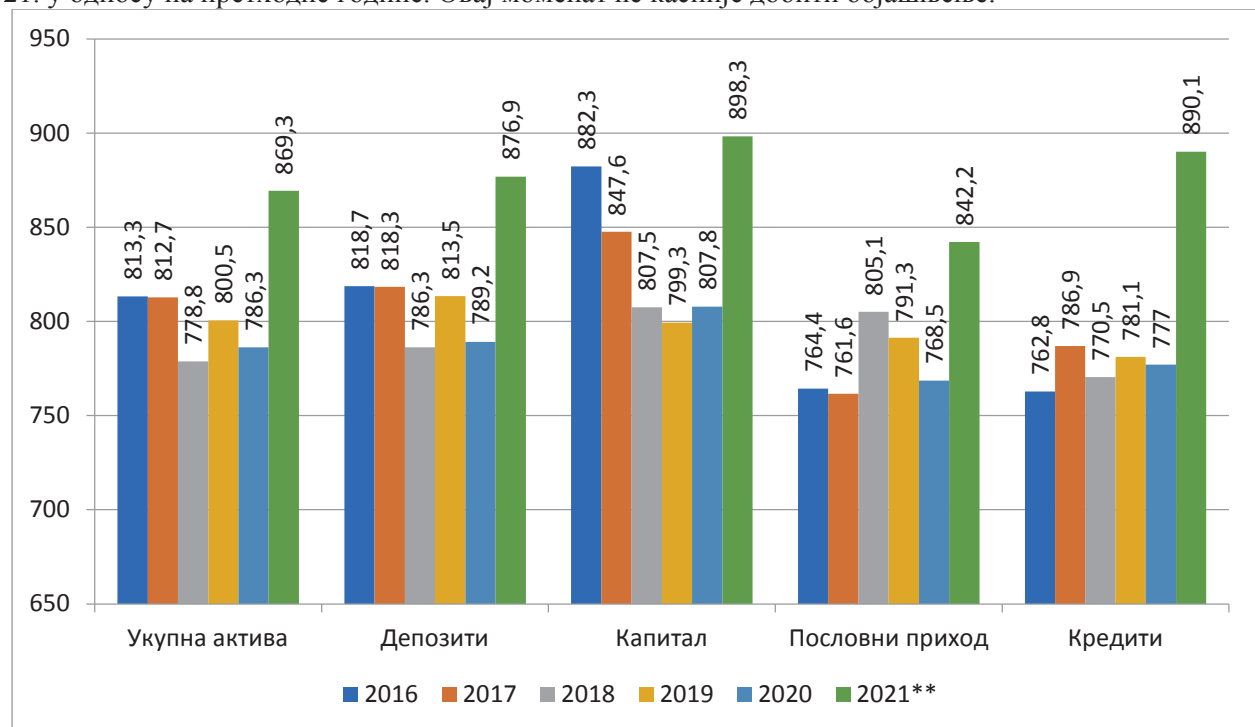
$$\frac{\Delta HHI}{HHI} \approx \frac{\Delta \varphi[GEI(2)]}{\varphi[GEI(2)]} - \frac{\Delta N}{N} \quad (8)$$

што представља полазну тачку наше наредне анализе.

Наведени методолошки поступак до сада је, зачудо, примњиван веома ретко. Бахо и Салас применили су га на великом узорку од 68 индустријских и неиндустријских сектора у Шпанији [6]. Следећа два примера односе се на финансијски сектор, конкретно – осигурање у Индији [13] и банковни сектор у Пољској [12]. Најзад, у недавном раду [3] овај аутор га је применио на банковни сектор у Србији, према тада доступним подацима, закључно са половином 2021. године.

### 3. ПРОМЕНЕ КОНЦЕНТРАЦИЈЕ У БАНКОВНОМ СЕКТОРУ СРБИЈЕ: ХИРШМАН-ХЕРФИНДАЛОВ ИНДЕКС И ЊЕГОВА ДЕКОМПОЗИЦИЈА

Опште информације о степену концентрације у банковном сектору Србије (без Косова и Метохије) даје слика 1. Приказане су вредности Хиршман-Херфиндаловог коефицијента за пет билансних променљивих, које смо примењивали и у ранијим радовима. Као и раније, коришћени су подаци из завршних рачуна банака, доступни на сајту Народне банке Србије. Прихватајући стандардну поделу, можемо закључити да је концентрација ниска (у свим случајевима мања од 900), иако и блиска умереној. Оно што посебно привлачи пажњу је значајан раст коефицијента (за све променљиве) у 2021. у односу на претходне године. Овај моменат ће касније добити објашњење.



Слика 1: Хиршман-Херфиндалов коефицијент концентрације у банковном сектору Србије\* 2016–2021\*\*

\* Без Косова и Метохије; \*\* Јануар–септембар.

Извор: Прерачунато на основу података из [1].

Како је већ истакнуто, ради одређивања утицаја наведених фактора (број учесника и дисперзија њихових удела) на кретање ХХ индекса кренућемо од израза (8), а затим ћемо применити поступак дат изразима (9) и (10), као што је то учињено у раду пољских аутора [12]:

$$I\varphi(GEI) = \frac{\frac{|\Delta\varphi[GEI(t)]|}{\varphi[GEI(t-1)]}}{\frac{|\Delta\varphi[GEI(t)]|}{\varphi[GEI(t-1)]} + \frac{|\Delta N(t)|}{N(t-1)}} \quad (9)$$

$$IN = \frac{\frac{|\Delta N(t)|}{N(t-1)}}{\frac{|\Delta\varphi[GEI(t)]|}{\varphi[GEI(t-1)]} + \frac{|\Delta N(t)|}{N(t-1)}} \quad (10)$$

На основу полазних података, узимајући у обзир идентичност индекса Хиршмана-Херфиндала (и индекса Хана-Кеја) и уопштеног индекса антропије, израчунали смо компоненте индекса НИ, тј. његових промена и допринос сваке од компонената, сходно изразима (9) и (10). Добијени резултати представљени су у табели 1. Висок проценат објашњених промена (изнад 90% у свим случајевима) потврђује квалитет извршене анализе и релевантност добијених резултата.

**Табела 1:** Декомпозиција промена индекса ХХ за банковни сектор Србије\* 2017–2021\*\*

Променљива	Година	Релативни раст			Објашњено		Допринос фактора (%)	
		НИ	$\varphi(GEI)$	N	Свега	%	$\varphi(GEI)$	N
Укупна актива	2017.	-0,001	-0,034	-0,033	-0,001	96,7	50,7	49,3
	2018.	-0,042	-0,108	-0,069	-0,039	93,1	61,0	39,0
	2019.	0,028	-0,010	-0,037	0,027	96,3	21,3	78,7
	2020.	-0,018	-0,018	0	-0,018	100	100,0	0,0
	2021.	0,106	0,021	-0,077	0,097	92,3	21,1	78,9
Депозити	2017.	0,029	-0,006	-0,033	0,028	96,7	15,4	84,6
	2018.	-0,039	-0,105	-0,069	-0,036	93,1	60,3	39,7
	2019.	0,035	-0,004	-0,037	0,033	96,3	9,8	90,2
	2020.	-0,030	-0,030	0	-0,030	100	100,0	0,0
	2021.	0,111	0,026	-0,077	0,103	92,3	25,0	75,0
Капитал	2017.	-0,039	-0,071	-0,033	-0,038	96,7	68,3	31,7
	2018.	-0,047	-0,113	-0,069	-0,044	93,1	62,1	37,9
	2019.	-0,010	-0,047	-0,037	-0,010	96,3	56,0	44,0
	2020.	0,011	0,011	0	0,011	100	100,0	0,0
	2021.	0,112	0,026	-0,077	0,103	92,3	25,6	74,4
Пословни приход	2017.	-0,004	-0,037	-0,033	-0,004	96,7	52,9	47,1
	2018.	0,057	-0,016	-0,069	0,053	93,1	18,8	81,2
	2019.	-0,017	-0,054	-0,037	-0,017	96,3	59,3	40,7
	2020.	-0,029	-0,029	0	-0,029	100	100,0	0,0
	2021.	0,096	0,012	-0,077	0,089	92,3	13,1	86,9
Кредити	2017.	0,032	-0,003	-0,033	0,031	96,7	8,3	91,7
	2018.	-0,021	-0,088	-0,069	-0,019	93,1	56,1	43,9
	2019.	0,014	-0,024	-0,037	0,013	96,3	39,3	60,7
	2020.	-0,005	-0,005	0	-0,005	100	100,0	0,0
	2021.	0,146	0,057	-0,077	0,134	92,3	42,8	57,2

\* Без Косова и Метохије; \*\* Јануар – септембар.

Извор: Сопствени прорачуни на основу података из [1].

Примењујући класификацију намењену утврђивању типова промена индекса Хиршмана-Херфиндала дату у раду Јацковича и Ковалевског, добили смо распоред типова промена индекса Хиршмана-Херфиндала (табела 2). Као што се може видети, нису заступљене све теоријске могућности, односно комбинације промена неједнакости у расподели удела и броја банака. Типови промена распоређени су не потпуно једнозначно, што се може тумачити као својеврсна потврда закључка о важности избора билансних величина за сличне анализе концентрације (и следствено конкуренције) [2]. Најчешће се појављује промена представљена у последњем реду табеле (смањење дисперзије удела и смањење броја банака, уз већу апсолутну вредност стопе раста неједнакости), тако да можемо рећи да он најбоље карактерише промене у концентрацији у банковном сектору Србије у посматраном периоду. Једини изузетак у готово униформном распореду типова промена с падом вредности индекса ХХ представља 2020. година (уз изузетак променљиве „Капитал”, за коју смо већ у претходним радовима констатовали да је најмање квалитетна за подобне анализе). Годину 2021.



(јануар-септембар) карактерише, при томе без изузетка, „чист” тип промена – повећање неједнакости у расподели удела уз смањење броја банака. Слично је у случају пада вредности ХХ индекса било у 2020. (с поменутиим изузетком индекса за капитал), с падом неједнакости у расподели удела и непромењеним бројем банака.

**Табела 2:** Типови промена индекса Хиршмана-Херфиндала у банковном сектору Србије\* у периоду 2017–2021\*\*

Промена ХХИ	$\Delta\phi$	$\Delta N$	Додатни услов	Билансне променљиве***				
				X1	X2	X3	X4	X5
Раст $r_{HHI} > 0$	>0	<0		2021	2021	2021	2021	2021
	>0	=0				2020		
	<0	<0	$ r\phi  <  rN $	2019	2017 ; 2019		2018	2017 ; 2019
Пад $r_{HHI} < 0$	<0	=0		2020	2020		2020	2020
	<0	<0	$ r\phi  >  rN $	2017 ; 2018	2018	2017 ; 2018 ; 2019	2017 ; 2019	2018

\* Без Косова и Метохије; \*\* Јануар–септембар; \*\*\* Билансне променљиве: X1 - актива; X2 - депозити; X3 - капитал; X4 - пословни приход; X5 - кредити.

Извор: Сопствени прорачуни на основу података из [1].

#### 4. ЗАКЉУЧАК

У посматраном периоду број банака у Србији (без Косова и Метохије) непрестано се смањивао, од 30 у 2016. до 24 на крају септембра 2021. (с изузетком 2020, у којој је остао непромењен према 2019). Али, без обзира на то, тенденција раста индекса концентрације се не уочава, што је супротно интуитивном (и теоријском) очекивању. Мали изузетак је 2019. година, када је забележен мањи раст индекса, али не за све променљиве, и 2021, када је дошло до доста великог раст индекса (више од 10%), при томе за све посматране променљиве.

Резултати декомпозиције индекса Хиршмана-Херфиндала показали су да су доприноси двају фактора (компонента неједнакости у распореду тржишних удела и броја учесника на тржишту) неједнаки и варирају по годинама. Међу издвојеним типовима промена преодлава пад индекса ХХ узрокован смањењем неједнакости тржишних удела и смањењем броја банака, при чему је прво по апсолутној вредности веће. С друге стране, у 2021. (јануар-септембар) код свих променљивих десило се повећање вредности изазвано повећањем компоненте неједнакости и смањењем броја банака, што је с теоријског аспекта јасна претпоставка пораста степена концентрације. У овој години први пут је, дакле, у посматраном периоду дошло до паралелног раста вредности индекса концентрације и смањења броја банака. У тумачењу овог резултата морамо се подсетити упозорења да промена броја учесника на тржишту није довољан узрок да дође до (супротне) промене вредности коефицијента концентрације, што је јасно из израза (2) и што смо већ апострофирали. Конкретно, то би значило да улазак новог учесника неће по аутоматизму довести до смањења индекса концентрације, то се, наиме, неће десити ако тај нови учесник има довољно велики тржишни удео. И обрнуто. Управо тако нешто био је случај у 2021, аквизицијом (практично, можемо рећи – интеграцијом) довољно крупних банака, које по тржишном уделу спадају у веће значајно је увећана компонента неједнакости у вредности индекса ХХ. Заправо, у анализираном периоду 2021. (првих девет месеци) десиле су се две сличне аквизиције, и то је значајно утицало на пораст концентрације.

Као врло интересантан показује се такође резултат, да општепосматрано индекс концентрације у посматраном периоду није растао, без обзира на то што је број банака опадао, и што је компонента броја банака имала у значајном броју случајева чак већи допринос у промени индекса концентрације.

#### ЛИТЕРАТУРА

- [1] Биланс стања/успеха банака [Електронски ресурс], Народна банка Србије. 2016–2021. URL: [https://nbs.rs/sr\\_RS/finansijske-institucije/banke/bilans-stanja/](https://nbs.rs/sr_RS/finansijske-institucije/banke/bilans-stanja/) (датум приступа: 10.03.2022.)
- [2] Буквић, Р. М. (2020) Шта показују индекси концентрације: пример банковног сектора Србије, XLVII Симпозијум о операционим истраживањима, SYM-OP-IS 2020, Београд 20–23 септембар 2020: зборник радова / ур. М. Видовић, К. Вукадиновић, Д. Поповић. Универзитет у Београду, 55–60.

- [3] Буквич, Р. М. (2022) Концентрация и конкуренция в современном банковском секторе Сербии: перемены и декомпозиция индекса Херфиндала – Хиршмана, *Современная конкуренция*, 16(2), 51–67.
- [4] Смарагдов, И. А. и Сидорейко В. Н. (2015). Индексы рыночной концентрации: неоднозначная информативность, *Концепт*, 9, 1–5.
- [5] Уредба о садржини и начину подношења пријаве концентрације, *Службени гласник РС*, број 5 од 25. јануара 2016.
- [6] Bajo, O.; Salas, R. (1998) Índices de concentración para la Economía Española: Análisis a Partir de las Fuentes Tributarias, *Economía Industrial*, 320, 101–116.
- [7] Bajo, O.; Salas, R. (2002). Inequality foundations of concentration measures: An application to the Hannah-Kay indices, *Spanish Economic Review*, 4(4), 311–316. DOI: 10.1007/s101080200053.
- [8] Bajo-Rubio, O.; Salas, R. (2004). Decomposing Change in Industry Concentration, *The Empirical Economics Letters*, 3(6), 311–319.
- [9] Bukvić, R. M. (2020). Measuring of concentration and competition: Serbian banking sector, *Synaxa*, 4(6–7), 65–90.
- [10] Cowell, F. A. (1977). *Measuring Inequality: Techniques for the Social Sciences*, New York: Wiley.
- [11] Hannah, L.; Kay, J. A. (1977). *Concentration in modern industry: Theory, measurement and the UK experience*, London, Basingstoke: The Macmillan Press Ltd, 1977.
- [12] Jackowicz, K.; Kowalewski, O. (2007). Dekompozycja miar koncentracji i dywersyfikacji działalności. Przypadek sektora polskich banków komercyjnych, *Bank i kredyt*, 38(5), 36–52.
- [13] Sastry, D.V.S. (2012). Concentration and Indian Insurance, *Journal of Quantitative Economics*, 10(2), 42–58.
- [14] Waterson, M. (1984) *Economic Theory of the Industry*, Cambridge: Cambridge University Press.

## TRETMAN INFLACIJE U MODELU DISKONTOVANOG NOVČANOG TOKA

### INFLATION TREATMENT IN DISCOUNTED CASH FLOW MODEL

NINA MILENKOVIĆ<sup>1</sup>, DEJAN KOKANOVIĆ<sup>2</sup>

<sup>1</sup> KPMG d.o.o. Beograd, nmilenkovic@kpmg.com

<sup>2</sup> KPMG d.o.o. Beograd, dkokanovic@kpmg.com

---

**Rezime:** Kada se procena vrednosti metodom diskontovanog novčanog toka radi u periodu stabilne niske inflacije, nije od velikog značaja da li će novčani tok biti projektovan u realnom ili nominalnom izrazu, sve dok su sve ključne veličine (projekcije poslovanja, diskontna stopa i stopa rasta u rezidualu) formalno i suštinski konzistentne. U uslovima porasta inflacije preporučuje se korišćenje nominalnog novčanog toka, ali je potrebno dodatno obratiti pažnju na odnos projektovanog rasta i stope inflacije. Ukoliko je u projekcije uključen i realni rast (tj. ako su stope rasta više od stope inflacije), neophodno je prilagoditi investicije u osnovna sredstva u odnosu na amortizaciju, a takođe i utvrditi neophodni nivo reinvestiranja profita da bi se obezbedio realni rast. Najzad, u tom slučaju potrebno je korigovati i tradicionalni Gordonov model za procenu rezidualne vrednosti.

**Ključne reči:** Diskontovani novčani tok, inflacija, diskontna stopa, rezidualna vrednost, stopa rasta u rezidualu

**Abstract:** When the discounted cash flow method is applied in valuation in a period of stable low inflation, it is not of great importance whether the cash flow is projected in real or nominal terms, as long as all key assumptions (business projections, discount rate and terminal growth rate) are formally and principally consistent. In conditions of rising inflation, it is recommended to use the nominal cash flow, but additional attention should be paid to the relationship between the projected growth and the inflation rate. If real growth is included in the projections (i.e., if growth rates are higher than the inflation rate), it is necessary to adjust investments in fixed assets in relation to depreciation, and also to determine the necessary level of profit reinvestment to ensure real growth. Finally, in that case it is necessary to adjust the traditional Gordon model for the terminal value assessment.

**Keywords:** Discounted cash flow, inflation, discount rate, terminal value, terminal growth rate



## **ZELENE OBVEZNICE U EU – ANALIZA DETERMINANTI PRINOSA**

### **GREEN BONDS IN THE EU – YIELD DETERMINANTS' ANALYSIS**

IRENA JANKOVIĆ<sup>1</sup>, VLADIMIR VASIĆ<sup>2</sup>, JELENA BASARIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, irena.jankovic@ekof.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, vladimir.vasic@ekof.bg.ac.rs

<sup>3</sup> Narodna banka Srbije, jelena.basaric@outlook.com

**Rezime:** Cilj ovoga rada jeste da predstavi važne karakteristike i specifičnosti zelenih obveznica kao brzoručne vrste dužničkih finansijskih instrumenta. Kako je osnovna namena ovih instrumenata prikupljanje finansijskih sredstava za finansiranje projekata usmerenih na očuvanje prirodne sredine i ublažavanje klimatskih promena, oni predstavljaju interesantnu investicionu alternativu za investitore koji teže diversifikaciji rizika. U radu je empirijski testiran uticaj promenljivih od značaja za kretanje prinosa zelenih obveznica emitovanih u EU. Nalazi potvrđuju značajan uticaj ročnosti, rejtinga, eksterne sertifikacije i relevantnih makropromenljivih na prinose zelenih obveznica.

**Ključne reči:** zelene obveznice, prinosi, EU

**Abstract:** The aim of this paper is to present the important features and specifics of green bonds as a fast-growing type of debt financial instruments. As the main purpose of these instruments is to raise funds to finance projects aimed at preserving the natural environment and mitigating climate change, they represent an interesting investment alternative for investors seeking to diversify risk. The paper empirically tests the influence of variables of importance for the movement of green bond yields issued in the EU. The findings confirm the significant impact of the maturity, rating, external certification and relevant macro variables on the green bond yields.

**Keywords:** green bonds, yields, EU

#### **1. UVOD**

Zelene obveznice su dužničke hartije od vrednosti koje su usmerene na finansiranje “zelenih” projekata koji za cilj imaju očuvanje prirodne sredine i ublažavanje klimatskih promena. Posledice klimatskih promena se mogu videti u naglom skoku prosečnih temperatura, povećanju kiselosti voda, podizanju nivoa mora i topljenju glečera. Kao dodatni štetni događaji, učestalo se javljaju suše i poplave. Ogromni iznosi finansijskih sredstva potrebni su za usporavanje negativnih posledica klimatskih promena. Prema skorašnjim predviđanjima, kako bi se porast globalne temperature održao ispod 2°C, što je prema Pariskom klimatskom sporazumu iz 2015. godine gornja dozvoljena granica globalnog zagrevanja koja smanjuje verovatnoću katastrofalnih ishoda, potrebno je na globalnom nivou izdvojiti preko 12 biliona američkih dolara do 2050. godine [2].

Zelene obveznice su se pojavile kao obećavajući finansijski instrumenti za borbu protiv klimatskih promena. Po svojim tehničkim karakteristikama ove obveznice su slične klasičnim obveznicama, dok se ključna razlika uočava u nameni finansijskih sredstava prikupljenih njihovom emisijom [22]. Sredstva prikupljena emisijom zelenih obveznica se usmeravaju ka adaptaciji na klimatske promene, održivim prirodnim izvorima, obnovljivoj energiji, održivoj upotrebi zemljišta, energetske efikasnosti, prečišćavanju voda, prevenciji i kontroli zagađenja. Zelene obveznice su uobičajeno strukturirane kao klasične obveznice sa fiksnim prinosom, ali mogu biti emitovane i u formi prinosa obveznica ili sekjuritizovanih instrumenata. Razlikuju se prema emitentu, kreditnom rejtingu i ročnosti.

Osnove za razvoj ovog tržišnog segmenta postavljene su 2014. godine razvojem smernica i preporuka pod nazivom *Green Bond Principles* (GBP). Dok navedeni principi definišu kriterijume za upotrebu prikupljenih finansijskih sredstava emisijom zelenih obveznica i periodično izveštavanje, sertifikacija prema *Climate Bonds Standard* (CBS) potvrđuje usaglašenost zelenih obveznica sa GBP [16, 6]. Sertifikacija suštinski potvrđuje da su projekti finansirani putem zelenih obveznica konzistentni sa okvirima Pariskog klimatskog sporazuma. Dodatno, rejting agencije procenjuju zelene emisije i nadziru upotrebu prikupljenih finansijskih sredstava. Dok sa jedne strane sertifikacija uvećava troškove emisije na kratak rok, posebno za male

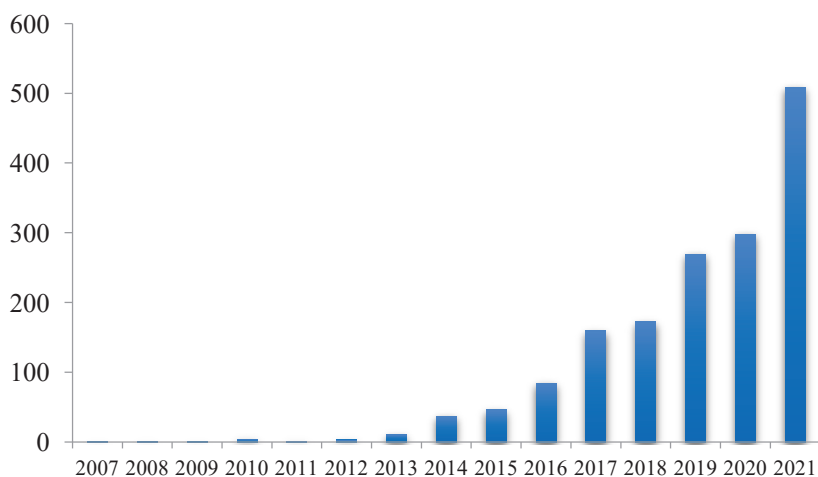
izdavaoće [13, 11], na srednji i dugi rok uvećanje transparentnosti i poverenja investitora bi trebalo da snizi zahtevani prinos na ove instrumente i da pozitivno utiče na veličinu tržišta zelenih obveznica.

Istraživanje u ovome radu usmereno je na državne zelene obveznice u EU, segment tržišta koji do sada nije zasebno ekstenzivnije izučavan. U radu se analiziraju faktori od značaja za kretanje prinosa ovih obveznica. Praktične implikacije analize ogledaju se u boljem razumevanju faktora od značaja za kretanje prinosa zelenih obveznica od strane njihovih emitenata što bi trebalo da rezultuje i povoljnijim uslovima finansiranja.

Rad je organizovan na sledeći način. Drugo poglavlje pruža prikaz trenutnog stanja na tržištu zelenih obveznica. Treći deo rada odnosi se na pregled relevantne literature. Metodološki okvir prikazan je u poglavlju četiri. Peti deo rada odnosi se na rezultate analize i prateću diskusiju. Šesti deo rada nudi zaključak celokupnog istraživanja.

## 2. PRIKAZ TRŽIŠTA ZELENIH OBVEZNICA

Prva zelena obveznica emitovana je od strane Evropske investicione banke 2007. godine. U narednim godinama tržište zelenih instrumenata se sporo razvijalo da bi prvi značajan rast doživelo u 2013. godini dostižući vrednost od 11 milijardi USD. Najnoviji podaci ukazuju na to da je vrednost ukupnih emisija zelenih obveznica dostigla vrednost od preko 508 milijardi USD u 2021. godini [7]. Iako je, nesporno, reč o značajnom rastu, ovaj tržišni segment, relativno gledano, predstavlja još uvek mali procenat ukupnog tržišta obveznica, ali sa snažnim potencijalom za nastavak rasta u narednom periodu.



**Grafikon 1:** Godišnje emisije zelenih obveznica (u milijardama USD)

Izvor: [7]

Najveći obim emisija potekao je iz razvijenih zemalja (oko 75%). Posmatrano prema regionima, emisije zelenih obveznica u Evropi dostigle su 52% ukupnih emisija na globalnom nivou u 2021. godini. Tokom iste godine emisije zelenih obveznica korporativnog sektora porasle su za 169% dok su i emisije državnih zelenih obveznica doživele značajan rast od 124% u poređenju sa prethodnom godinom. Tri najznačajnija sektora u koje se usmeravaju finansijska sredstva od emisije zelenih obveznica u Evropi jesu energetska sektor, građevinarstvo i saobraćaj. Posmatrano prema valutnoj strukturi, emisije u evrima činile su preko 72% svih emisija u 2021. godini [7].

## 3. PREGLED LITERATURE

Tokom vremena izdvojilo se nekoliko pravaca u literaturi koja se bavi analizom zelenih obveznica.

Značajan broj radova fokusira se na finansijske aspekte emisije zelenih obveznica prateći njihov uticaj na vrednost kapitala preduzeća, reputaciju i prinos na vlasničke hartije. Sveobuhvatna analiza tržišta zelenih obveznica imala je za cilj da identifikuje ključne ograničavajuće faktore za razvoj ovoga tržišnog segmenta i pruži preporuke za standardizaciju zelenih instrumenata koja bi se odrazila na veličinu i likvidnost tržišta zelenih hartija. Kao ključne izazove za razvoj ovoga tržišnog segmenta [8] navode odsustvo harmonizovanih globalnih standarda, moguća rizik zloupotrebe, visoke troškove emisija, i dr.

Sledeća grupa radova usmerena je na vrednovanje i analizu prinosa zelenih obveznica. Veliki broj radova u ovoj grupi poredi prinose zelenih i klasičnih obveznica istog emitenta. Oni uvode pojam tzv. zelene premije koja označava negativnu razliku u prinosima zelenih i klasičnih obveznica. Tada se zelene obveznice prodaju uz cenovnu premiju. Pri tome, prva podgrupa autora pronalazi dokaze koji idu u prilog postojanju zelene premije [2,12,25,14,1,10]. Sa druge strane, druga podgrupa autora ne nalazi čvrste dokaze koji idu u prilog postojanju zelene premije [17,21,18,23]. Oprečni nalazi dve podgrupe autora najverovatnije su posledica različitog uzorkovanja, razlika u analiziranim regionima i posmatranim vremenskim periodima, primenjenim metodologijama i karakteristikama emitentata i samih analiziranih obveznica.

Poseban pravac u literaturi o zelenim obveznicama bavi se pitanjem faktora koji utiču na ponudu ovih instrumenata kao i kretanje njihovih prinosa. Tako [9] u svom istraživanju zaključuju da regulatorni okvir i adekvatna supervizija pozitivno utiču na ponudu zelenih obveznica. Autori [5] zaključuju da kuponska stopa ima negativan dok kreditni rejting i kolateral imaju pozitivan uticaj na veličinu emisija i posebno evro-denominovane emisije zelenih obveznica.

Analizirajući zelene obveznice u Kini [24] zaključuju da eksterna certifikacija, visok kreditni rejting, veličina emisije i ročnost utiču na smanjenje prinosa zelenih obveznica i troškove finansiranja za emitente. Sličan nalaz se može videti i u radu [19] gde se potvrđuje da visok kreditni rejting i certifikacija snižavaju troškove finansiranja za listirane emitente zelenih obveznica. Analizom zelenih obveznica G-20 zemalja u periodu od 2007-2016. godine [20] ukazuju na to da viši nivo usaglašenosti sa Principima zelenih obveznica rezultuje većom tražnjom za ovim instrumentima. Ako su obveznice emitovane od strane države to umanjuje negativan uticaj niske usaglašenosti sa pomenutim principima. [15] potvrđuju da sertifikovane hartije nose niže prinose od necertifikovanih. [4] su putem analize modela panela proveravali uticaj likvidnosti i kreditnog rejtinga na spredove prinosa zelenih obveznica i uočili negativan uticaj ovih promenljivih na spredove prinosa.

#### 4. METODOLOGIJA, OPIS PROMENLJIVIH I ISTRAŽIVAČKE HIPOTEZE

Kako bismo ispitali uticaj relevantnih objašnjavajućih promenljivih na kretanje prinosa državnih zelenih obveznica u EU ocenili smo sledeći model sa fiksnim efektima za podatke u formi balansiranog panela:

$$Y_{it} = \alpha + \beta X_{it} + \mu_i + \vartheta_{it} \quad (1)$$

za  $t = 1, \dots, T$  i  $i = 1, \dots, N$ , gde je  $Y_{it}$  zavisna promenljiva,  $\alpha$  odsečak,  $\beta$  koeficijenti,  $X_{it}$  objašnjavajuće promenljive,  $\mu_i$  individualni specifični efekti, a  $\vartheta_{it}$  preostali deo slučajne greške [3]. Prednost ovakve specifikacije ogleda se u tome što uzima u obzir specifičnosti pojedinačnih jedinica posmatranja.

Putem navedenog modela ispitujemo uticaj preostale ročnosti obveznice, 6-mesečne stope *Euribor*, *VIX* indeksa, kreditnog rejtinga i eksterne certifikacije na prinose državnih zelenih obveznica u EU. U uzorku se nalaze dnevni podaci za posmatrane promenljive u periodu od 31.03.2021-30.09.2021. godine preuzeti sa platforme *Bloomberg*. Uzorak se sastoji od 13 obveznica (dve emitovane od strane Francuske, dve od strane Nemačke, po jedna u Italiji, Irskoj, Litvaniji, Holandiji, Mađarskoj, Belgiji i tri u Poljskoj).

**Tabela 1:** Zavisna i objašnjavajuće promenljive

Naziv promenljive	Definicija	Izvor
Prinos	Prinos na državne zelene obveznice u EU	<i>Bloomberg Professional Software Service</i>
Ročnost	Preostala ročnost obveznice	Obračun autora
<i>VIX</i>	Indeks volatilnosti (S&P500)	<i>Bloomberg Professional Software Service</i>
<i>6M EURIBOR</i>	Šestomesečna <i>EURIBOR</i> kamatna stopa	<i>Bloomberg Professional Software Service</i>
Kreditni rejting	Uzima vrednost jedan ako je Moody's rejting Aaa ili viši i vrednost 0 u suprotnom slučaju	Kreirana od strane autora
Odsustvo eksterne certifikacije	Uzima vrednost 1 ako obveznica nije eksterno sertifikovana, a vrednost 0 u suprotnom slučaju	Kreirana od strane autora

Analizu baziramo na sledeće dve istraživačke hipoteze:

- Indeks volatilnosti *VIX*, 6-mesečna stopa *Euribor*, preostala ročnost i odsustvo eksterne certifikacije imaju pozitivan uticaj na prinose zelenih obveznica.
- Zelene obveznice koje imaju viši kreditni rejting imaju niži prinos.

## 5. REZULTATI ANALIZE I DISKUSIJA

U Tabeli 2. prikazane su deskriptivne statistike za analizirane promenljive.

**Tabela 2:** Deskriptivne statistike

Promenljiva	Broj opservacija	Srednja vrednost	Standardna devijacija	Min.	Maks.
Prinos	1.690	0,2747373	0,6443952	-0,868	1,912
Ročnost	1.690	13,24438	7,366417	4	27
VIX	1.690	18,10946	2,173647	15,07	27,59
6M EURIBOR	1.690	-0,5182385	0,0061522	-0,531	-0,505

Minimalni prinos imala je nemačka 5-godišnja zelena obveznica rejtinga Aaa, dok je najviši prinos imala italijanska obveznica rejtinga Baa3. Prosečna ročnost posmatranih obveznica iznosila je 13 godina. Prosečna vrednost *VIX* indeksa iznosila je 18 što ukazuje na to da se nije očekivala visoka volatilnost u narednom periodu. Šestomesečna stopa *Euribor* bila je prilično stabilna i kretala se oko nivoa od -0.52%.

Nakon što je potvrđeno odsustvo multikolinearnosti kod objašnjavajućih promenljivih (niska korelisanost i *VIF* u rasponu od 1,02 (za *VIX*) i 1,17 (za preostalu ročnost)). Primenom *Hausman* testa na nivou značajnosti od 5% potvrdili smo da je specifikacija sa fiksnim efektima odgovarajuća (nasuprot alternativnoj specifikaciji sa slučajnim efektima) ( $\chi_2 = 5,99, p = 0,0144$ ).

Problemi koje smo identifikovali u analizi jesu međuzavisnost između jedinica posmatranja, prisustvo autokorelacije i heteroskedastičnosti kod reziduala. Adresirali smo poslednji problem kod specifikacije sa fiksnim efektima. Kako bi smo rešili sva tri identifikovana problema ocenili smo dodatno koeficijente u modelu primenom metode uopštenih najmanjih kvadrata (kod koga smo specifikovali prisustvo heteroskedastičnosti i korelisanosti između reziduala u odnosu na panele, kao i prisustvo autokorelacije reziduala). Rezultati ocenjenih modela prikazani su u Tabeli 3.

**Tabela 3:** Testiranje uticaja objašnjavajućih promenljivih na prinose zelenih obveznica – panel model sa fiksnim efektima i model ocenjen metodom uopštenih najmanjih kvadrata

	(1) <i>FE</i>	(2) <i>GLS</i>
Ročnost	0,0319 (0,0201)	0,0428*** (0,000944)
<i>VIX</i>	0,0105*** (0,00189)	-0,000342 (0,000276)
<i>6M Euribor</i>	10,24*** (1,334)	0,0216 (0,147)
Kreditni rejting		-0,931*** (0,0137)
Odsustvo eksterne certifikacije		0,0171 (0,0192)
Konstanta	4,965*** (0,804)	0,631*** (0,0817)
N	1690	1690
r <sup>2</sup>	0,449	
r <sup>2</sup> <sub>o</sub>	0,602	
r <sup>2</sup> <sub>b</sub>	0,620	
r <sup>2</sup> <sub>w</sub>	0,449	
sigma <sub>u</sub>	0,493	
sigma <sub>e</sub>	0,0758	
rho	0,977	

Standardne greške su navedene u zagradama.

\* p<0,05 \*\* p<0,01 \*\*\* p<0,001



Poređenjem rezultata dobijenih analizom, možemo zaključiti da preostala ročnost pozitivno utiče na prinose državnih zelenih obveznica u EU. Što je duža ročnost hartija to investitori zahtevaju dodatnu kompenzaciju u formi višeg prinosa. Što je viša očekivana volatilnost na finansijskim tržištima, merena *VIX* indeksom, to su viši prinosi zahtevani na zelene obveznice kao kompenzacija za rastuće rizike. Kretanje 6-mesečne *Euribor* stope kao referentne kamatne stope za međubankarske pozajmice u evrima pozitivno utiče na prinose državnih zelenih obveznica u EU. Obveznice koje nisu eksterno certifikovane nose viši prinos jer suštinski nose i viši rizik da sredstva prikupljena njihovom emisijom neće biti korišćena za finansiranje ekološki održivih projekata. Obveznice sa nižim kreditnim rejtingom zahtevaju viši prinos u poređenju sa obveznicama čiji je kreditni rejting viši. Dobijeni rezultati idu u prilog potvrdi definisanih istraživačkih hipoteza u ovome radu.

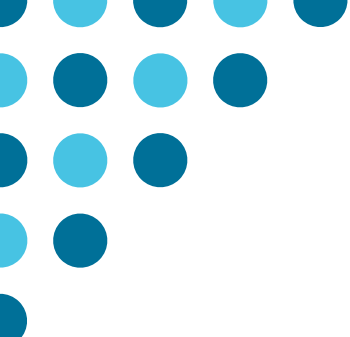
## 6. ZAKLJUČAK

Cilj analize u ovome radu bio je da se ispituju ključni faktori od značaja za kretanje prinosa državnih zelenih obveznica u EU. Rezultati analize potvrđuju pozitivan uticaj preostale ročnosti obveznica, kretanja *VIX* indeksa i 6-mesečne *Euribor* stope na kretanje prinosa. Niži kreditni rejting ovih obveznica rezultuje višim prinosom kao i odsustvo njihove eksterne certifikacije koja bi potvrdila usaglašenost sa Principima zelenih obveznica. Dalja unifikacija standarda u ovoj oblasti kreirala bi jako uporište za dalji razvoj tržišta zelenih obveznica.

## LITERATURA

- [1] Bachelet, M. J., Becchetti, L., & Manfredonia, S. (2019). The green bonds premium puzzle: The role of issuer characteristics and third-party verification. *Sustainability*, 11(4), 1098.
- [2] Baker, M., Bergstresser, D., Serafeim, G., & Wurgler, J. (2018). *Financing the response to climate change: The pricing and ownership of US green bonds* (No. w25194). National Bureau of Economic Research.
- [3] Brooks, C. (2019). *Introductory Econometrics for Finance*. Cambridge University Press.
- [4] Chang, K., Feng, Y. L., Liu, W., Lu, N., & Li, S. Z. (2021). The impacts of liquidity measures and credit rating on corporate bond yield spreads: evidence from China's green bond market. *Applied Economics Letters*, 28(17), 1446-1457.
- [5] Chiesa, M., & Barua, S. (2019). The surge of impact borrowing: The magnitude and determinants of green bond supply and its heterogeneity across markets. *Journal of Sustainable Finance & Investment*, 9(2), 138-161.
- [6] Climate Bonds Initiative. (2019). *Climate Bonds Standard*. London: Climate Bonds Initiative.
- [7] Climate Bonds Initiative. (2022). *Database*. <https://www.climatebonds.net/market/data/>
- [8] Deschryver, P., & De Mariz, F. (2020). What future for the green bond market? How can policymakers, companies, and investors unlock the potential of the green bond market?. *Journal of Risk and Financial Management*, 13(3), 61.
- [9] Dou, X., & Qi, S. (2019). The choice of green bond financing instruments. *Cogent Business & Management*, 6(1), 1652227.
- [10] Fatica, S., Panzica, R., & Rancan, M. (2021). The pricing of green bonds: are financial institutions special?. *Journal of Financial Stability*, 54, 100873.
- [11] Forsbacka, K., & Vulturius, G. (2019). A Legal Analysis Of Terms and Conditions For Green Bonds: Focus on the Financial Markets in the Nordics. *Europarättslig tidskrift*, 3, 397-442.
- [12] Gianfrate, G., & Peri, M. (2019). The green advantage: Exploring the convenience of issuing green bonds. *Journal of cleaner production*, 219, 127-135.
- [13] Hachenberg, B., & Schiereck, D. (2018). Are green bonds priced differently from conventional bonds?. *Journal of Asset Management*, 19(6), 371-383.
- [14] Hyun, S., Park, D., & Tian, S. (2020). The price of going green: the role of greenness in green bond markets. *Accounting & Finance*, 60(1), 73-95.
- [15] Hyun, S., Park, D., & Tian, S. (2021). Pricing of green labeling: A comparison of labeled and unlabeled green bonds. *Finance Research Letters*, 41, 101816.
- [16] International Capital Market Association. (2018). *Green Bond Principles*. Zurich: International Capital Market Association.

- [17] Karpf, A., & Mandel, A. (2018). The changing value of the 'green' label on the US municipal bond market. *Nature Climate Change*, 8(2), 161-165.
- [18] Larcker, D. F., & Watts, E. M. (2020). Where's the greenium?. *Journal of Accounting and Economics*, 69(2-3), 101312.
- [19] Li, Z., Tang, Y., Wu, J., Zhang, J., & Lv, Q. (2020). The interest costs of green bonds: Credit ratings, corporate social responsibility, and certification. *Emerging Markets Finance and Trade*, 56(12), 2679-2692.
- [20] Nanayakkara, K. G. M., & Colombage, S. (2021). Does compliance with Green Bond Principles bring any benefit to make G20's 'Green economy plan' a reality?. *Accounting & Finance*, 61(3), 4257-4285.
- [21] Partridge, C., & Medda, F. R. (2020). The evolution of pricing performance of green municipal bonds. *Journal of Sustainable Finance & Investment*, 10(1), 44-64.
- [22] Reboredo, J. C. (2018). Green bond and financial markets: Co-movement, diversification and price spillover effects. *Energy Economics*, 74, 38-50.
- [23] Tang, D. Y., & Zhang, Y. (2020). Do shareholders benefit from green bonds?. *Journal of Corporate Finance*, 61, 101427.
- [24] Wang, Q., Zhou, Y., Luo, L., & Ji, J. (2019). Research on the factors affecting the risk premium of China's green bond issuance. *Sustainability*, 11(22), 6394.
- [25] Zerbib, O. D. (2019). The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking & Finance*, 98, 39-60.



# GEOINFORMACIONI SISTEMI

---

# GEOINFORMATION SYSTEMS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business



## KARTOGRAFSKA GENERALIZACIJA NASELJENIH MESTA NA TOPOGRAFSKIM KARTAMA

### CARTOGRAPHIC GENERALIZATION OF SETTLEMENTS ON TOPOGRAPHIC MAPS

MARKO SIMIĆ<sup>1</sup>, MARKO STOJANOVIĆ<sup>2</sup>, RADOJE BANKOVIĆ<sup>3</sup>, JASMINA M. JOVANOVIĆ<sup>4</sup>, TANJA M. JANKOVIĆ<sup>5</sup>, MARIJA STOJANOVIĆ<sup>6</sup>

<sup>1</sup> Vojnogeografski institut – „General Stevan Bošković“, Beograd, marko.simic.simke1996@gmail.com

<sup>2</sup> Vojnogeografski institut – „General Stevan Bošković“, Beograd i Univerzitet odbrane – Vojna akademija, Beograd, stojanovicm80@yahoo.com

<sup>3</sup> Vojnogeografski institut – „General Stevan Bošković“, Beograd i Univerzitet odbrane – Vojna akademija, Beograd, radojebankovic@yahoo.com

<sup>4</sup> Univerzitet u Beogradu – Geografski fakultet, jasminda.jovanovic@gef.bg.ac.rs

<sup>5</sup> Univerzitet u Beogradu – Geografski fakultet, tanja.jankovic115@gmail.com

<sup>6</sup> Vojnogeografski institut – „General Stevan Bošković“, Beograd, marijasmaki996@gmail.com

**Rezime:** Zbog velikog privrednog, kulturnog, političkog i vojnog značaja, naseljena mesta su jedan od najznačajnijih elementata na topografskim kartama. U zavisnosti od karakteristike naseljenog mesta i razmere topografske karte razlikuje se način prikaza naseljenog mesta i ispis toponima. U eri digitalne kartografije velika pažnja se posvećuje automatizaciji procesa kartografske generalizacije i razvoju alata za automatsko uopštavanje sadržaja. U ovom radu su objašnjeni modeli i procedure za kartografsku generalizaciju naseljenih mesta prilikom izrade topografskih karata kompletnog razmernog niza.

**Ključne reči:** Kartografska generalizacija, Naseljena mesta, Topografske karte.

**Abstract:** Due to their great economic, cultural, political and military significance, settlements are one of the most important elements on topographic maps. Depending on the characteristics of the settlement and the scale of the topographic map, the way of displaying the settlement and toponyms differs. In the era of the digital cartography more attention is paid to the automation process of cartography generalization and developing tools for automatic generalization of content. This paper explains the models and procedures for cartographic generalization of settlements during the production a complete scale of topographic maps.

**Keywords:** Cartographic generalization, Settlements, Topographic maps.

## 1. UVOD

Potreba čoveka da opiše svoje stanište i grafički ga prestavi datira još od prvih preteča karata (crteži na zidovima pećina, koži, kori drveta, kamenu i dr.). Karte predstavljaju osnovni kartografski dokument, koji sadrži tačan, detaljan i vizuelno najpregledniji prikaz elemenata sadržaja geoprostora. Načini kartografske vizuelizacije su se usavršavali i menjali u odnosu na stepen razvoja društva i nivo tehničkih mogućnosti. Potreba da se što vernije kartografski predstavi položaj i specifičnosti naseljenih mesta uslovljena je njihovim značajem u svim geoprostornim istraživanjima brojnih nauka. U savremenim uslovima, tipologije naselja smatraju se korisnim instrumentom za sprovođenje socio-ekonomskih razvojnih programa, odnosno referentnim okvirom za kreiranje nacionalnih i regionalnih razvojnih politika. Imaju veliki privredni, kulturni, politički i vojni značaj. Predstavljaju osnovni pokazatelj razmeštaja stanovništva, centre saobraćaja, privrednih i kulturno-istorijskih sadržaja. U odnosu na svoj geografski položaj, fizionomiju i značaj ustanova koje se u njima nalaze, naseljena mesta se razlikuju po svom tipu i funkcijama (gradska i seoska; politički, administrativni, kulturni i ekonomski centri). Zbog svog velikog značaja, ona su jedan od najvažnijih elemenata geografske karte i kao takva obavezno se predstavljaju na topografskim kartama (TK). Kartografskom vizuelizacijom elemenata sadržaja naseljenih mesta omogućava se njihovo kontinuirano, uniformno i standardizovano prikazivanje na TK celokupnog razmernog niza. Digitalna kartografska vizuelizacija integriše kartografske metode i principe u Geografski informacioni sistem (GIS), što omogućava istraživačke analize geopodataka. Olakšano je dobijanje georeferenciranih podataka, prostorno istraživanje, analiza, sinteza i prezentacija informacija na sistematizovan i transparentan način.

## 2. NASELJENA MESTA

Oblikovanje jedinstvenog, prostorno i vremenski postojanog modela diferenciranja urbanih i ruralnih naselja otežano je iz više razloga. Pre svega, to je posledica opšteg društvenog, političkog, privrednog i tehničko-tehnološkog, odnosno civilizacijskog razvoja, ali i evolucije socijalne misli i shvatanja urbanog, odnosno ruralnog [1]. Na prostoru Republike Srbije, raznovrsni istorijski, geografski i ekonomski uslovi nastanka i razvoja naselja, različiti demografski tokovi i niz drugih činilaca, rezultirali su veoma izraženim razlikama među naseljima, kako u pogledu njihove populacione veličine, tako i u demografskim i socio-ekonomskim karakteristikama stanovništva koje u njima živi. Evidentna raznolikost baznih karakteristika naselja obavezuje na njihovo adekvatno tretiranje, identifikovanje gradskih i prigradskih naselja kao i ruralnih centara, odnosno definisanje kategorije seoskih naselja.

Metodologija kartografskog uopštavanja osnovnih tipova naseljenih mesta olakšava njihovo prikazivanje na TK. Karakteristike naseljenih mesta koje se uzimaju kao osnova za njihovu klasifikaciju su tip, veličina, politički i administrativni značaj [2]. Najvažnija je klasifikacija naseljenih mesta po tipu, jer se ona koristi kao osnova za dalje klasifikovanje po nekom drugom osnovu. Prema tipu, dele se na naseljena mesta gradskog tipa i naseljena mesta seoskog tipa.

Naseljena mesta gradskog tipa karakteriše specifična struktura, uslovljena gustom stanovništvom, izgrađenošću stambenim, privrednim i kulturnim objektima, komunikacijama i drugim urbanim specifičnostima. Ona predstavljaju kulturne i političko-administrativne centre većih područja, njihova specifična struktura se odražava u: uređenosti ulica, trgova i parkova; formiranju blokova zgrada zbog zbijenosti prostora, naročito u centru grada; jasnom ocrtavanju spoljašnjih ivica grada; i razgranatosti komunikacija u raznim pravcima. Naseljena mesta seoskog tipa se po svojoj strukturi i veličini razlikuju od gradskih. Gustina naseljenosti sela je znatno manja nego u gradovima. Oblik i fizionomija seoskih naselja pod velikim je uticajem geografskog položaja i prirodnih specifičnosti. Seoska naselja karakteriše: nedostatak uređenih ulica, trgova i parkova; pojedinačni razmeštaj objekata na većem ili manjem rastojanju u zavisnosti od prirodnih uslova i komunikacija; slabije ocrtana kontura; slabije razgranati prilazni putevi koji su najčešće lošijeg kvaliteta i dr. Sva seoska naselja u odnosu na svoju strukturu, mogu se svrstati tipski u seoska naselja razbijenog tipa i seoska naselja zbijenog tipa. Seoska naselja razbijenog tipa karakteriše veće rastojanje između kuća ili grupa kuća, koje su locirane na velikoj površini, bez nekog reda. Dok su kod seoskih naselja zbijenog tipa kuće zbijene jedna do druge, grupisane pravilno ili nepravilno, u zavisnosti od zemljišta na kome se nalaze i od same razvijenosti naselja. Veličina naseljenih mesta koja se prikazuju na TK data su na slici 1 [3].

Називни насељених места:	
Гргегег	Насељено место до 100 становника и мањи засеок
Кремна	Насељено место од 100 до 1 000 становника и већи засеок
Рудник	Насељено место од 1 000 до 2 000 становника
Врдник	Насељено место од 2 000 до 5 000 становника
КЛАДОВО	Насељено место од 5 000 до 10 000 становника
СЈЕНИЦА	Насељено место од 10 000 до 25 000 становника
КИКИНДА	Насељено место од 25 000 до 50 000 становника
КРАЉЕВО	Насељено место од 50 000 до 100 000 становника
НОВИ САД	Насељено место од 100 000 до 500 000 становника
БЕОГРАД	Насељено место са више од 500 000 становника

**Slika 1:** Upotreba slova za klasifikaciju naseljenih mesta po broju stanovnika na TK25 [3]

Na TK se, pored četiri osnovne karakteristike naseljenih mesta koje se prikazuju na svim geografskim kartama (tip naseljenog mesta, veličina u odnosu na broj stanovnika, značaj i odnos prema ostalim elementima sadržaja), prikazuju i oblik, dimenzije i unutrašnja struktura naseljenog mesta, jer krupniji razmer to dozvoljava. Na krupnorazmernim TK naseljena mesta se prikazuju u ortogonalnoj projekciji, a na sitnorazmernim pomoću uslovnih znakova u vidu kružića ili kontura gradova.

Pored individualnih karakteristika naselja koje se prikazuju u zavisnosti od razmera i namene karte, na TK se posebno ističu detalji koji se mogu koristiti kao orijentiri. Tu se svrstavaju zgrade koje se vide iz daleka izvan naseljenog mesta i markantni detalji koji su lako vidljivi u samom mestu. Oni se predstavljaju, tako da na karti budu uočljivi, znakom koji je tačno pozicioniran [2].

### 3. KARTOGRAFSKA GENERALIZACIJA

Kartografska generalizacija je proces uopštavanja pri projektovanju i sastavljanju sadržaja geografskih karata. Manifestuje se kao uopštavanje kvalitativnih i kvantitativnih karakteristika objekata, zamene pojedinačnih opštim, apstrahovanje od pojedinosti i detalja radi jasne slike glavnih obeležja prostorne distribucije [4]. Suština procesa je da se na karti predstave glavne, tipične karakteristike objekata. Primenjen postupak kartografske generalizacije određuje kvalitet karte. Sadržajna i grafička generalizacija je složen proces koji primarno zavisi od razmere i namene karte, kao i specifičnosti i karakteristika teritorije koja se kartira.

U postupku generalizacije pojedinosti nestaju, a najbitnije karakteristike objekta se jasnije ističu. Proces generalizacije je teže formalizovati i automatizovati od drugih kartografskih procesa. Ne mogu se algoritmovati sve faze i procedure, niti se mogu svi kriterijumi nedvosmisleno propisati. Kvalitet generalizacije umnogome zavisi od kartografskog razumevanja suštine sadržaja prikazanih objekata i pojava, kao i sposobnosti kartografa da indentifikuje glavne i tipične osobine. Iskustvo pokazuje da automatizacija kartografske generalizacije treba da se zasniva na interaktivnim postupcima koji obezbeđuju aktivno učešće kartografa [5].

Teško je uspostaviti trajna pravila za generalizaciju. Pravila utvrđena u jednom slučaju, za jednu teritoriju, mogu biti neprikladna u drugom slučaju. Međutim, neophodno je uspostaviti osnovne principe kartografske generalizacije kako bi se u svakom pojedinačnom slučaju moglo naći pravo rešenje i kako bi postojala ujednačena vrednost i standardni kvalitet na celoj teritoriji kartiranja.

Faktori koji utiču na stepen generalizacije karata su: namena karte (ukoliko je namena karte šira utoliko stepen generalizacije pojedinih elemenata treba da bude manji i obrnuto), razmer karte (ograničava broj elemenata kartirane teritorije koji može biti prikazan na karti), karakteristike teritorije kartiranja (analizom različitih geografskih područja na istoj karti određuje se stepen generalizacije koji obezbeđuje adekvatno prikazivanje posebnih karakteristika tih sredina), kartografski izvori (bogat i dobar kartografski izvor doprinosi pravilnom određivanju stepena generalizacije) i prag čitljivosti karte (jedna od polaznih osnova za rešavanje veličina, oblika, boje i intenziteta kartografskih znakova u odnosu na način korišćenja karte.) [6].

Razvojem GIS-a i digitalne kartografije došlo je do promene postupka izrade karte i načina skladištenja geoprostornih podataka. Oni se ne čuvaju više na kartama već u bazama podataka [7].

### 4. KARTOGRAFSKA GENERALIZACIJA NASELJENIH MESTA NA TOPOGRAFSKIM KARTAMA

Na krupnorazmernim kartama generalisanje naseljenih mesta ima za cilj da što bolje sačuva njihovu tipičnu strukturu. Izborom važnijih objekata koje treba predstaviti i postepenim uopštavanjem njihovih pojedinih karakteristika dolazi se do tog cilja. Naseljena mesta su jako osetljiva na promenu razmera, ona se mogu prikazati kao površinski objekti, sa konturama u razmeru karte do razmere 1:100.000. Do ove razmere prikazuje se i struktura naselja koju čine glavne ulice, izgrađeni i neizgrađeni delovi, industrijski, stambeni i rekreativni delovi, glavni objekti i prilazi naselju u vezi sa ostalim elementima karte.

Pri predstavljanju naselja neophodno je očuvati odnos izgrađenih i neizgrađenih površina kako bi se pravilno predstavila njihova struktura. Na topografskoj karti razmere 1:25.000 (TK25) ulice su prikazane šire nego u prirodi, tako da se zgrade do puta pomeraju ka unutrašnjosti bloka. Većina zgrada se prikazuje veća od svojih dimenzija u prirodi, dok se neki objekti i detalji izostavljaju. Na topografskoj karti razmere 1:50.000 (TK50) nastavlja se širenje ulica zbog čega dolazi do daljeg potiskivanja zgrada ka unutrašnjosti. Kako bi se očuvao odnos izgrađenih i slobodnih površina (crno-beli odnos) dolazi do izostavljanja sporednih ulica, spajanja manjih blokova u veće i redukcije prikazanih zgrada. Na topografskoj karti razmere 1:100.000 (TK100) ulice zauzimaju još više prostora u odnosu na veličinu u prirodi. Zgrade su prikazane minimalnim veličinama i zauzimaju četiri puta više prostora. Zbog očuvanja odnosa crno-belih površina dolazi do dalje redukcije prikazanih objekata [8].

U narednim primerima je prikazana generalizacija naselja gradskog i seoskog tipa (razbijenog i zbijenog) na TK u razmernom nizu od 1:25.000 do 1:250.000.

#### 4.1. Generalizacija naseljenog mesta gradskog tipa na primeru Vranja

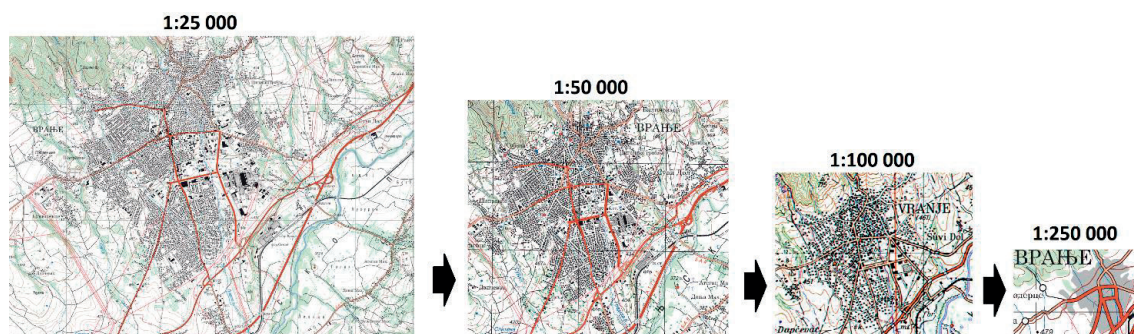
Vranje je naseljeno mesto gradskog tipa, ima 83.524 stanovnika (popis stanovništva 2011. god.) i kategoriše se kao grad srednje veličine. Na TK25 prikazani su svi objekti od javnog značaja (najveći objekti površinskim znakom, a ostali uslovnim), kao i većina kuća i zgrada. Kuće i zgrade se na TK25, TK50 i TK100 prikazuju istim uslovnim znakom. U ovoj razmeri, kao i na TK50 i TK100, je moguće prikazati

tipičan raspored zgrada (u nizu, u blokovima, raštrkane itd). Prikazani su svi fabrički kompleksi sa svim, ili većinom, objekata u okviru kompleksa. Prikazani su svi magistralni putevi koji prolaze kroz grad i sve ulice u gradu kategorisane po značaju i širini. Prikazani su svi dalekovodi, kao i njihovi stubovi, koji provode struju iznad 35 kV i deo antenskih stubova. Naziv grada je na celom razmernom nizu (TK25, TK50, TK100 i TK250) ispisan velikim slovima na obodu grada. Prikazani su i nazivi manjih naselja koja se nalaze u okviru grada. Prikazane su skraćenice koje se odnose na objekte od javnog značaja, mostove i puteve, kao i kote i trigonometrijske tačke u okviru grada koliko opterećenost karte sadržajem to dozvoljava. Prikazuju se stalni i povremeni tokovi koji prolaze kroz grad.

Na TK50 veliki broj kuća i zgrada je redukovan. Prikazani su svi fabrički kompleksi sa redukovanim brojem objekata u okviru kompleksa. Redukovan je broj prikazanih antenskih stubova i dalekovoda jer se prikazuju samo oni značajniji. Prikazuju se glavne i najveće ulice, a sporedne se redukuju. Dolazi do redukcije skraćenica, dok se stubovi dalekovoda i njihove skraćenice ne prikazuju. Nazivi i nadmorske visine crkava su redukovani. Prikazuju se samo visine odrđenih crkvi, sa visinama koje se ispisuju ispod imena naselja. Takođe su i kote i trigonometrijske tačke redukovane.

Na TK100 izvršena je velika redukcija kuća, zgrada i objekata od javnog značaja. Prikazuju se samo najveći i najbitniji fabrički kompleksi sa znatno redukovanim brojem objekata u okviru njih. Dalekovodi i antenski stubovi u naselju se ne prikazuju, dok se oni na obodu prikazuju. Magistralni putevi koji prolaze kroz grad se proširuju, i zbog toga dolazi do suženja prostora za objekte i ulice. Ostaju prikazane samo glavne i najveće ulice u naselju. Popunjenim blokovima, zgrade se objedinjuju u gotove gradske blokove, dok se u blokovima gde nisu sve zgrade izgrađene ostavlja prazan prostor, kako bi ostao utisak odnosa između izgrađene i neizgrađene površine. U blokovima sa zgradama koje su razmaknute jedna od druge čuva se ta karakteristika tako što se jedne redukuju, a druge verno orijentišu i prikažu. Nazivi manjih naselja u okviru grada se ne prikazuju. Skraćenice se ne prikazuju, osim ako je objekat na obodu grada tako da ima mesta za skraćenicu. Prikazuje se samo visina najbitnije crkve ispod naziva naselja. Kote i trigonometrijske tačke su znatno redukovane, kao i većina mostova.

Na TK250 redukovani su svi osim najbitnijih objekata koji se prikazuju uslovnim znakom. Na primeru Vranja prikazana je samo železnička stanica. Umesto objekata, vanrazmerno se prikazuje površina i struktura naselja pomoću konturnih linija. Magistralni putevi koji prolaze kroz grad se prikazuju i proširuju, dok su sve ulice u naselju redukovane i ne prikazuju se. Skraćenice, visine crkava, kote i trigonometrijske tačke se ne prikazuju. Prikazuju se samo stalni tokovi koji prolaze kroz grad, dok su povremeni tokovi redukovani.



Slika 2: Prikaz naseljenog mesta gradskog tipa, na isečku TK25, TK50, TK100 i TK250 [3,9,10,11]

#### 4.2. Generalizacija naseljenog mesta zbijenog seoskog tipa na primeru Zlatokopa

Zlatokop je naseljeno mesto zbijenog seoskog tipa, ima 795 stanovnika (popis stanovništva 2011. god.) i kategoriše se kao selo srednje veličine. Na TK25 prikazani su svi objekti od javnog značaja, kao i većina kuća. Kuće se na TK25, TK50 i TK100 prikazuju istim uslovnim znakom. Na karti ove razmere, kao i na TK50, moguće je prikazati tipičan raspored kuća. Prikazani su svi dalekovodi, kao i njihovi stubovi, koji provode struju iznad 35 kV. Prikazani su svi putevi koji prolaze kroz selo, sve ulice u naselju i većina kolskih puteva. Naziv sela je na celom razmernom nizu ispisan sa prvim velikim slovom (ostala su mala) i nalazi se na obodu sela. Date su sve skraćenice. Prikazane su sve kote i trigonometrijske tačke sa svojim visinama. Prikazana je većina voćnjaka i deo usamljenih stabala koji se nalaze u selu. Takođe su prikazani i stalni i povremeni tokovi koji prolaze kroz selo.

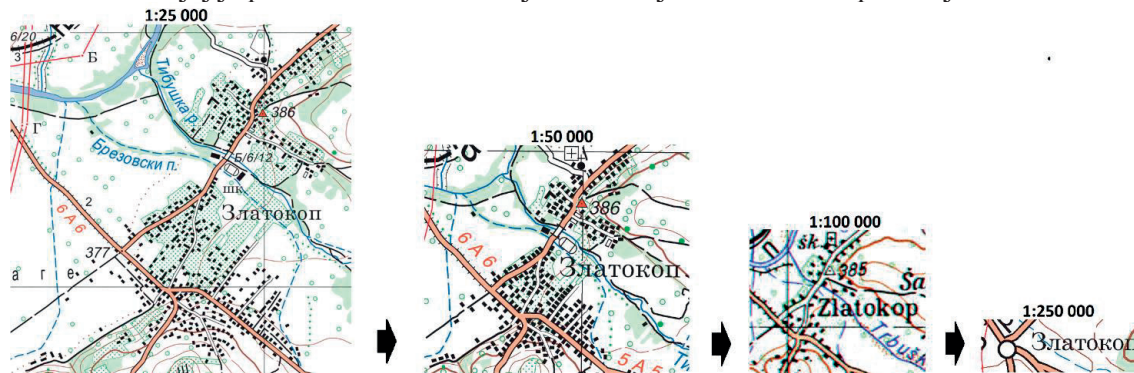
Na TK50 kuće su prikazane dosta zbijenije, a deo njih je redukovan. Takođe, redukovan je deo objekata od javnog značaja. Prikazuju se samo najbitniji dalekovodi, dok se njihovi stubovi ne prikazuju. Prikazani su svi magistralni putevi koji prolaze kroz selo, kao i većina ulica, dok se deo kolskih puteva redukuje. Skraćenice su potpuno redukovane i ne prikazuju se. Kote i trigonometrijske tačke su prikazane, ali



redukovano. Prikazani su stalni i povremeni tokovi koji prolaze kroz selo. Voćnjaci koji se nalaze u selu su redukovani, kao i većina usamljenih stabala.

Na TK100 izvršena je velika redukcija kuća i objekata od javnog značaja. Dalekovodi i antenski stubovi u naselju se ne prikazuju, dok se oni na obodu prikazuju. Magistralni putevi koji prolaze kroz grad se proširuju i zbog toga dolazi do suženja prostora za objekte i ulice. Ostaju prikazane samo glavne i najveće ulice, a takođe je redukovana i većina kolskih puteva. Kuće nisu više prikazane zbijeno, već se ostavlja prazan prostor između njih, kako bi se naglasio manji broj stanovnika u odnosu na gradove. Kote i trigonometrijske tačke su redukovane. Mostovi u selu su potpuno redukovani i ne prikazuju se. Voćnjaci i usamljena stabla koji se nalaze u selu su još više redukovani.

Na TK250 redukovani su svi objekti, a selo je predstavljeno uslovnim znakom. Prikazuju se samo dalekovodi koji prolaze pored naselja. Pošto je selo na raskrsnici puteva, uslovni znak je stavljen na preseku komunikacija. Ulice i kolski putevi u selu se ne prikazuju. U ovoj razmeri toponim znatno više opterećuje kartu nego samo naseljeno mesto, koje je predstavljeno uslovnim znakom. Kote i trigonometrijske tačke u selu se ne prikazuju. Prikazuju se samo stalni tokovi, ako kroz selo prolazi tok, a uslovni znak se stavlja sa one strane toka na kojoj je pretežni deo sela. Voćnjaci i usamljena stabla se ne prikazuju.



Slika 3: Prikaz naseljenog mesta zbijenog seoskog tipa, na isečku TK25, TK50, TK100 i TK250 [3,9,10,11]

#### 4.3. Generalizacija naseljenog mesta razbijenog seoskog tipa na primeru Klisurice

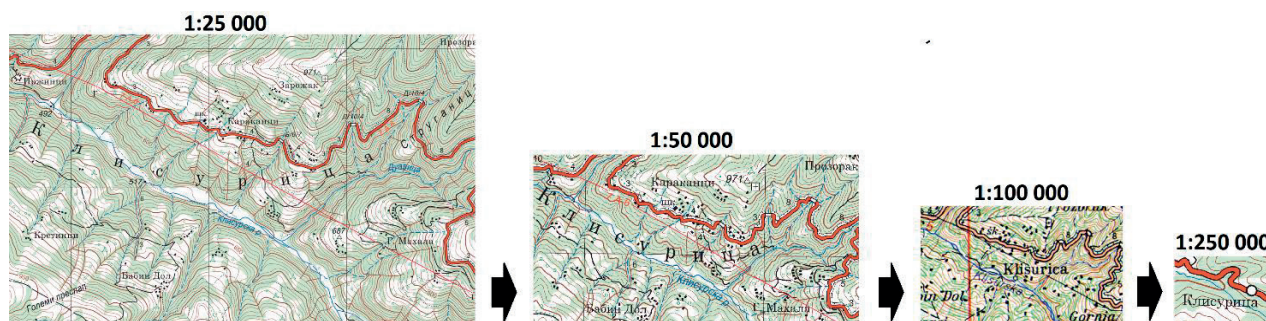
Klisurica je naseljeno mesto razbijenog seoskog tipa, ima 173 stanovnika (popis stanovništva 2011. god.) i kategoriše se kao selo male veličine. Na TK25 prikazani su svi objekti od javnog značaja, kao i većina kuća. Kuće se na TK25, TK50 i TK100 prikazuju istim uslovnim znakom. U ovoj razmeri, kao i na TK50 i TK100, je moguće prikazati tipičan raspored kuća. Prikazani su svi dalekovodi, kao i njihovi stubovi, koji provode struju iznad 35 kV. Prikazani su svi putevi koji prolaze kroz selo, kao i većina kolskih puteva. Ovo selo nema ulica, kao ni većina sela razbijenog tipa. Naziv sela je ispisan razbijeno sa prvim velikim slovom (ostala su mala) i predstavljen je kao površinski naziv koji je rasprostranjen preko teritorije celog sela. Takođe su dati nazivi manjih mesta u okviru sela. Prikazane su sve skraćenice, kote i trigonometrijske tačke. Prikazana je većina voćnjaka i deo usamljenih stabala koja se nalaze u selu. Takođe su prikazani stalni i povremeni tokovi koji prolaze kroz selo.

Na TK50 deo kuća je redukovano. Ostale kuće koje su prikazane nisu zbijane kako bi bio dat utisak razbijenosti sela. Prikazani su svi objekti od javnog značaja u selu. Prikazuju se samo najbitniji dalekovodi, dok se njihovi stubovi ne prikazuju. Prikazan je magistralni put koji prolazi kroz selo, dok se deo kolskih puteva redukuje. Naziv sela je ispisan razbijeno sa prvim velikim slovom i površinski rasprostranjen preko teritorije celog sela. Dati su nazivi manjih mesta u okviru sela. Skraćenice su delimično redukovane, dok se stubovi dalekovoda i njihove skraćenice ne prikazuju. Prikazuju se kote i trigonometrijske tačke, ali su redukovane. Voćnjaci koji se nalaze u selu su redukovani, kao i većina usamljenih stabala.

Na TK100 izvršena je redukcija kuća i objekata od javnog značaja. Dalekovodi i antenski stubovi u naselju se ne prikazuju, dok se oni na obodu prikazuju. Magistralni put koji prolazi kroz grad se prikazuje i proširuje. Većina kolskih puteva u naselju se redukuje. Naziv sela je ispisan zbijeno sa prvim velikim slovom i nije više rasprostranjen preko teritorije celog sela, već se nalazi u centru teritorije koju obuhvata selo i pokriva znatno manju površinu. Nazivi manjih mesta u okviru sela se ne prikazuju. Skraćenice su skoro potpuno redukovane. Većina kote i trigonometrijskih tačaka je redukovana. Prikazani su stalni i povremeni tokovi, sa mostovima, koji prolaze kroz selo. Voćnjaci i usamljena stabla u selu potpuno su redukovani i umesto njih prikazan je površinski znak za šumu.

Na TK250 redukovani su svi objekti, a selo je predstavljeno uslovnim znakom. Dalekovodi su potpuno redukovani. Pošto se selo nalazi na magistralnom putu, uslovni znak je stavljen na sam put. Kolski putevi u selu se ne prikazuju. Naziv sela je ispisan zbijeno sa prvim velikim slovom i nalazi se pored uslovnog znaka.

U ovoj razmeri toponim znatno više opterećuje kartu nego uslovni znak za selo. Prikazuju se samo stalni tokovi, ako kroz selo prolazi tok uslovni znak se stavlja sa one strane na kojoj je pretežni deo sela.



Slika 4: Prikaz naseljenog mesta razbijenog seoskog tipa, na isečku TK25, TK50, TK100 i TK250 [12,15,14,15]

## 5. ZAKLJUČAK

Potpunost, detaljnost i tačnost slike područja na karti zavisi pre svega od njene razmere. Krupnorazmerne karte prikazuju manju oblast sa većom količinom detalja, a sitnorazmerne karte prikazuju veću oblast sa manje detalja. Savremena vizuelizacija velikog broja podataka geografskog sadržaja, podržana kompjuterskom tehnologijom i GIS-om, omogućava brže i sveobuhvatnije dobijanje informacija, u cilju boljeg razumevanja i izražavanja fenomena i zakona geografskog sveta, integriše prenos informacija (komunikaciju) i omogućava kognitivne analize.

Kako bi naseljena mesta, sa svim svojim karakteristikama, bila što vernije predstavljena u svakoj razmeri, uprkos smanjenju prostora na listu karte za njihov prikaz, koji dolazi sa smanjenjem razmere, neophodno je izvršiti pravilnu generalizaciju. Kako bi se izvršio pravilan izbor objekata koji će biti prikazani i pravilno izvršilo njihovo uopšteno i grafički uprošćeno prikazivanje neophodno je poštovati postupke kartografske generalizacije. Oni omogućavaju očuvanje bitnih i tipičnih karakteristika naselja, dobijanje najvažnijih informacija o naselju, prilikom prelaska na sitniji razmer, i zadržavaju čitljivost sadržaja karte na zahtevanom nivou. Procena i izbor kartografskog metoda, uz poštovanje unapred razrađenih kriterijuma kartografske generalizacije, kako bi se adekvatno definisao sadržaj karte i sačuvala sva njena svojstva, neophodno obuhvata geografsko poznavanje i proučavanje teritorije kartiranja. U postupku kartografske generalizacije dolaze do izražaja naučna, teorijsko - metodološka znanja iz kartografije, stvaralačke mogućnosti, grafičko umeće i veštine kartografa.

## LITERATURA

- [1] Živanović, Z. (2018). Prilog diskusiji o tipologiji naselja Srbije. *Demografija*. No 15, str. 33-49.
- [2] Peterca, M., Radošević, N., Milisavljević, S. i Racetin, F. (1974). *Kartografija*. Vojnogeografski institut, Beograd.
- [3] Vojnogeografski institut, (2017). *TK25 NK34-5/6-2-4 Vranje*. Vojnogeografski institut, Beograd.
- [4] Берлянт, А. М., Востокова, А. В., Кравцова, В. И., Лурье, И. К., Сваткова, Т. Г., Серапинас, Б. Б. (2003). *Картоведение*. Москва: Издательство Аспект Пресс.
- [5] Берлянт, А. М. (2002). *Картография*. Издательство Аспект Пресс, Москва.
- [6] Салищев, К. А. (1990). *Картоведение*. Издательство Московского Университета, Москва.
- [7] Đorđević, M. (2016). *Primena GIS-a u kartografskoj generalizaciji kategorijskih karata*. Doktorska disertacija. Prirodno-matematički fakultet, Novi Sad.
- [8] Frančula, N. (2003). *Kartografska generalizacija - skripta*. Geodetski fakultet, Zagreb.
- [9] Vojnogeografski institut, (2022). *TK50 NK34-5/6-2 Vranje*. Vojnogeografski institut, Beograd.
- [10] Vojnogeografski institut, (1987). *TK100 Leskovac\_632*. Vojnogeografski institut, Beograd.
- [11] Vojnogeografski institut, (2020). *TK250 NK34-5 Priština*. Vojnogeografski institut, Beograd.
- [12] Vojnogeografski institut, (2017). *TK25 NK34-6/4-1-3 Bujkovac*. Vojnogeografski institut, Beograd.
- [13] Vojnogeografski institut, (2022). *TK50 NK34-6/4-1 Kriva Feja*. Vojnogeografski institut, Beograd.
- [14] Vojnogeografski institut, (1987). *TK100 Vlasotince\_633*. Vojnogeografski institut, Beograd.
- [15] Vojnogeografski institut, (2020). *TK250 NK34-6 Sofija*. Vojnogeografski institut, Beograd.

## KORIŠĆENJE MAŠINSKOG UČENJA ZA PROCENU VREDOSTI STANOVA U GIS OKRUŽENJU

### USING MACHINE LEARNING TO ESTIMATE THE VALUE OF APARTMENTS IN A GIS ENVIRONMENT

STANISLAVA BOSIOČIĆ<sup>1</sup>, ZORAN SRDIĆ<sup>1</sup>, DRAGAN SAVIĆ<sup>2</sup>

<sup>1</sup> АТУСС, одсек Висока грађевинско-геодетска школа, Београд, stanislava98@gmail.com, zoran.m.srdic@gmail.com

<sup>2</sup> Univerzitet Singidunum, ojlerovafunkcija@gmail.com

**Rezime:** Procena vrednosti stanova predstavlja jedan od zahtevnijih zadataka u proceni vrednosti nepokretnosti, naročito kada su u pitanju stanovi na teritoriji glavnog grada Republike Srbije - Beogradu, naročito u uslovima intenzivne i raznovrsne stanogradnje. U radu su opisane metode mašinskog učenja, na primeru najmnogoljudnije opštine Grada Beograda, opštine Novi Beograd, kao jednog od mogućih rešenja primenjenih u procenjivačkoj praksi nekretnina u Srbiji. Data je i analiza predikcije različitih modela regresora koji se koriste u tehnikama regresije mašinskog učenja.

**Ključne reči:** procena vrednosti stanova, metode mašinskog učenja, regresija

**Abstract:** Valuation of apartments is one of the most demanding tasks in real estate valuation, especially when it comes to apartments in the capital of the Republic of Serbia - Belgrade, especially in conditions of intensive and diverse housing. The paper describes the methods of machine learning, on the example of the most populous municipality of the city of Belgrade, the municipality of New Belgrade, as one of the possible solutions applied in real estate appraisal practice in Serbia. An analysis of the prediction of different models used in machine learning regression techniques is also given.

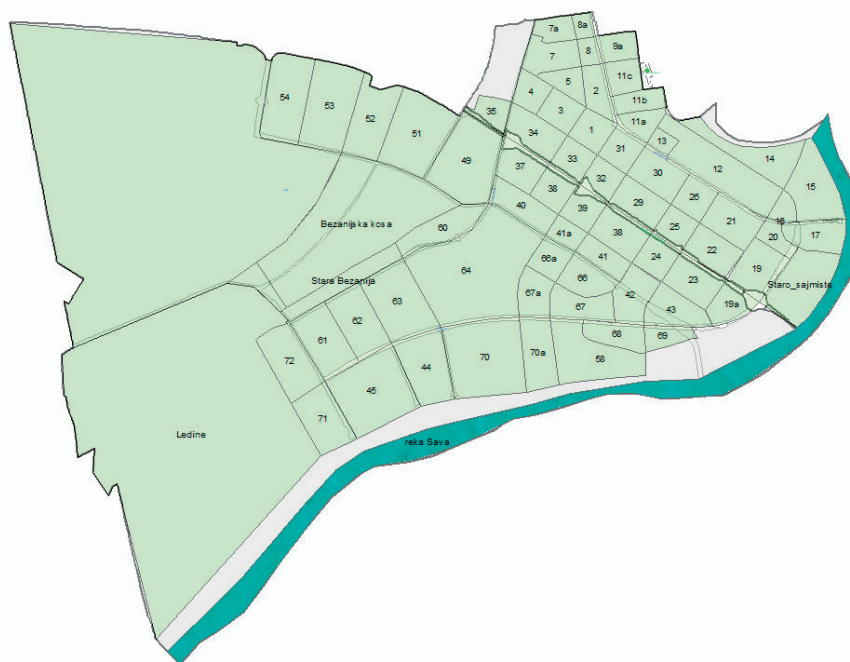
**Keywords:** valuation of apartments, machine learning methods, regression

#### 1. UVOD

Republički geodetski zavod (RGZ) je 2011. godine shodno Zakonu o državnom premeru i katastru, u svoju nadležnost uvrstio procenu i vođenje vrednosti nepokretnosti. U pitanju je sistem masovne procene vrednosti nepokretnosti i njegova glavna uloga je u obračunavanju poreza na imovinu, a bazira se na postulatima koji obezbeđuju pravičnost, jednakost i uniformnost. Za potrebe masovne procene vrednosti nepokretnosti, RGZ vodi registar cena nepokretnosti. Podaci iz registra dostupni su profesionalnim korisnicima (agencijama za promet nekretnina, javnim beležnicima, proceniteljima, bankama i drugim zainteresovanim korisnicima). Baza podataka registra cena nepokretnosti sadrži podatke iz ugovora o prometu i ugovora o zakupu nepokretnosti. Dostavu ugovora obavljaju organi nadležni za overu ugovora - javni beležnici i osnovni sudovi. Masovna procena vrednosti nepokretnosti u RGZ se radi po Pravilniku o proceni vrednostui nepokretnosti, a kalibracija modela za vrednovanje nepokretnosti, se vrši uz određivanja koeficijenata u modelu statističkim testiranjem [3]. Procenom nekretnina se mogu baviti i licencirani procenitelji, a tu oblast reguliše Zakon o proceniteljima vrednosti nepokretnosti. Ispit za procenitelje se polaže pri Ministarstvu finansija, posle čega procenitelj dobija licencu i može vršiti procene vrednosti nepokretnosti [2]. Tržište stanova je najrazvijenije podržište nepokretnosti u Republici Srbiji. Broj registrovanih kupoprodaja u kojim su učestvovali stanovi na teritoriji Republike Srbije u 2021.godini iznosio je 54.008, od toga u Beogradu 21.833, a na Novom Beogradu 2918 (1.937 starogradnje i 845 novogradnje) [6].

Opština Novi Beograd je jedna od 17 beogradskih opština. Nalazi se na levoj obali reke Save i desnoj obali Dunava. Naselje Novi Beograd je izgrađeno nakon Drugog svetskog rata i razvilo se u moderno naselje. Specifičnost Opštine Novi Beograd je relativno mala površina u odnosu na druge opštine, površina je 4.074 ha i deset puta je manja od površine najveće opštine u Gradu Beogradu - Opštine Palilula koja iznosi 44.661 ha. Što se tiče broja stanovnika daleko je najmnogoljudnija opština u Gradu Beogradu sa 217.773 stanovnika (Opština Palilula ima 155.900 stanovnika) po popisu iz 2002. godine.

Specifičnost stanogradnje na Novom Beogradu je da su po lokaciji stambene zgrade razvrstane u blokove (Slika 1), koje imaju svoje specifičnosti, a teritorija opštine je podeljena i u mesne zajednice kojih ima 19. Za razliku od Bežanije i Starog Sajmišta, ostale četvrti nemaju istorijska ili tradicionalna imena. Blok 1 se naziva još i „Fontana“, sagrađen 1962. godine kod bivšeg bioskopa Fontana i to je prvi formirani blok zgrada na Novom Beogradu. Takodje blokovi 45, 44, 70 i 70a se nazivaju Savski blokovi jer su smešteni uz Savsko šetalište, izgrađeni su 1979. godine i spadaju u starogradnju. Na području Novog Beograda izdvajaju se i lokacije novih stambenih kompleksa sagrađenih 2021. i 2022. godine (novogradnja) West65, Savada 3, Sakura park, New Minel, Plato Bežanija, V park, Elikzir garden i drugo.



**Slika 1.** Mapa Opštine Novi Beograd sa oznakama lokacije stanova po blokovima

Prema smernicama Ministarstva trgovine Republike Srbije za posrednike u prometu nepokretnosti [4], faktori koji imaju veći ili manji uticaj na konačnu cenu konkretne nepokretnosti su:

- Lokacija utiče na bruto cenu u proseku oko 25%;
- Sledeći faktor po značaju je sprat na kome se nepokretnost nalazi. Ukoliko se stan nalazi u suterenu cena je manja za 10 do 15% od regularne cene stana koji se nalazi između prvog i četvrtog sprata. Ukoliko se stan nalazi na višim spratovima cena takođe pada i po tome se naše tržište nepokretnosti razlikuje u odnosu na tržišta u razvijenim zemljama;
- Grejanje je faktor koji utiče na cenu oko 7% mada treba uvažiti činjenicu da i tu ima razlike u zavisnosti od vrste i kvaliteta grejanja (grejanje na struju, gas, solarno, daljinsko centralno, etažno i sl.);
- Očuvanost stana utiče na cenu oko 10%. To znači da je potrebno oko 10% od vrednosti stana da bi se isti renovirao. Razlika od 10% je razlika u ceni novogradnje i stanova starih više godina;
- Struktura (višesobnost) stana je faktor koji utiče oko 4% na cenu stana. Najskuplji po metru kvadratnom su mali stanovi do 35m<sup>2</sup> i preko 135m<sup>2</sup>. Za male stanove mala je i suma novca koja je potrebna za kupovinu te je i broj potencijalnih kupaca koji raspolažu potrebnim sredstvima veći. Za velike stanove je karakteristično da ih nema puno, te je i njihova cena po metru kvadratnog veća;
- Stanje sanitarija (kupatilo, kuhinja) utiče na cenu oko 5% jer za većinu adaptacija potrebno je menjati kompletnu instalaciju;
- Stolarija i bravarija utiču na cenu takođe sa oko 5% jer upotrebna vrednost iste tokom godina opada;
- Telefon utiče sa oko 3% pri čemu je intenzivnijim korišćenjem mobilne telefonije ovaj faktor drastično izgubio na značaju;
- Postojanje ili ne postojanje terase utiče na cenu sa oko 3%. Supstitut prethodnom su lođa, veranda, francuski balkon itd.;
- Višespratna ili niska zgrada utiče na cenu sa oko 4%;
- Lift u konstelaciji sa spratnošću utiče na cenu sa oko 3%;

- Orientacija stana (okrenutost prema severu npr. i sl.), osunčanost utiču na cenu sa oko 3%;
- Blizina prevoza utiče na cenu stana sa oko 3%, međutim ima i izuzetaka. Deo Beograda, Dedinje, Topčider i sl. gde prevoz nije prisutan drži visoku cenu, tako da se navedeni faktor tiče pre svega klasične stanogradnje;
- Faktor „čovjek“ je u principu najuticajnije i utiče sa 10% u proseku i najmanje je ekonomski opravdan. Struktura i osobine komšiluka time imaju veliki značaj;
- Ostali faktori utiču sa oko 5%. U ove faktore se ubrajaju broj parking mesta, blizina parka, vrtića, škole i sl. Problemi sa ograničenim brojem parking mesta doveli su do toga da je cena garaža oko 50% od cene kvadratnog metra stana, cena garažnog mesta u kolektivnoj garaži 35% od cene kvadrata stana, a da je cena parking mesta oko 25% od cene metra kvadratnog stana na toj lokaciji.

Da bi se napravila mapa lokacija za Novi Beograd (Slika 1) i utvrdila pripadnosti stana određenoj lokaciji korišćeni su javni podaci dostupni na internet stranici Urbanističkog zavoda Beograda, u biblioteci planova [5].

## 2. MAŠINSKO UČENJE

U radu je metodama mašinskog učenja analizirano tržište stanova na Novom Beogradu. Pri primeni metode prvi korak je formiranje tabele sa atributima, koji reprezentuju karakteristične osobenosti podataka koji se modeliraju. Postoji mnogo parametara koji utiču na cenu formiranja stana. Ti parametri variraju od države do države, pojedine oblasti ili gradovi imaju svoje specifičnosti, a takođe i određene mikrolokacije. U mašinskom učenju ti parametri su predstavljeni atributima koji mogu biti: kategorijski, diskretni i kontinualni.

Pri odabiru atributa uzete su specifičnosti mikrolokacije Novi Beograd (Tabela 1): prostorno dosta homogeno područje, svi sadržaji koji se smatraju atraktivnim su dostupni za sve lokacije (škola, vrtić, prodavnice, tržni centri, pošte, sportski sadržaji, domovi zdravlja, zabavni sadržaji), velika disproporcija kvaliteta i cene stanova starogradnje i novogradnje u susednim ili čak u istom bloku zgrada.

U mnogim radovima stranih autora [1] parametri koji se navode da utiču na cenu stana su udaljenost od regionalnog puta, železnice i metroa. Specifičnost Novog Beograda je dobra razvijenost putne mreže, dobra mreža javnog gradskog prevoza (autobuskog i tramvajskog). Specifičnost predstavlja naselje Ledine u kojem je ekspanzija potpuno novih i funkcionalnih stanova, sa lokacijom ne mnogo daljom od novobeogradskih blokova, ali bez potrebne infrastrukture (trotoar, neasfaltirani poprečni putevi, nedostatak škole, prodavnice, kulturnih i sportskih sadržaj). Ovaj parametar nije uključen u model.

**Tabela 1:** Nazivi korišćenih atributa sa objašnjenjima

atributi	objašnjenje	vrednosti
cena	Cena stana izražena u eurima (cena po kojoj se stan prodaje na tržištu)	0 - 700000
kv	Kvadratura stana bez terase (kvadratura stana na kojoj je stan uknjižen u katastru, a veliki broj stanova imaju neko dodatno proširenje koje je evidentirano u napomeni)	0 - 190
lokacija	Lokacija stana izražena kao pripadnost bloku	1,2,3,...11c,...65,..70a, 71,72, Stara Bezanija, Bezanijska kosa
broj soba	Broj postojećih prostorija u stanu (neki stanovi su izvorno bili na pr. 2 sobni a sadašnje stanje je 2,5)	0,5 - 5
spratnost	Sprat na kojem se stan nalazi	suteren, prizemlje, visoko prizemlje, 1- 40
CG	Stan poseduje centralno grejanje	Da/Ne
lift	Zgrada u kojoj se stan nalazi ima lift (neke više zgrade imaju i dva lifta, ne postoji podaci o svim zgradama o postojanju lifta, specifičnost je što i neke zgrade koje spadaju u novogradnju, a niske su spratnosti, do 4 sprata, nemaju lift)	Da/Ne
terasa	Stan poseduje terasu ili lođu (neke zgrade imaju i dve terase, ali taj podatak nije uzet u obzir jer ne postoji podatak za sve stanove)	Da/Ne

Jedna od specifičnih karakteristika stana koji utiču na cenu, kako na celoj teritoriji Republike Srbije, tako i na svakoj njenoj opštini je i da li je stan u prizemlju ili na poslednjem spratu. Inače je stan na poslednjem spratu u evropskim metropolama zbog mogućnosti penthousa i otvorenog velepnoog pogleda na grad pogodnost i plus koji povećava cenu, a u Beogradu se smatra manjkavošću i snižava cenu stana zbog pretpostavke mogućih problema u vezi održavanja krova i održavanja rada liftova. Taj parametar takođe iako značajan nije uziman u razmatranje zbog nedovoljno podataka.

Godina renoviranja stana, iako značajna karakteristika stana, nije uzimana u obzir zbog nepotpunih podataka za sve nepokretnosti.

Na osnovu podataka o karakteristikama stanova i poznatih cena stanova, primenom metoda regresione analize razmatrani su mogući modeli predviđanja nepoznate cene stana.

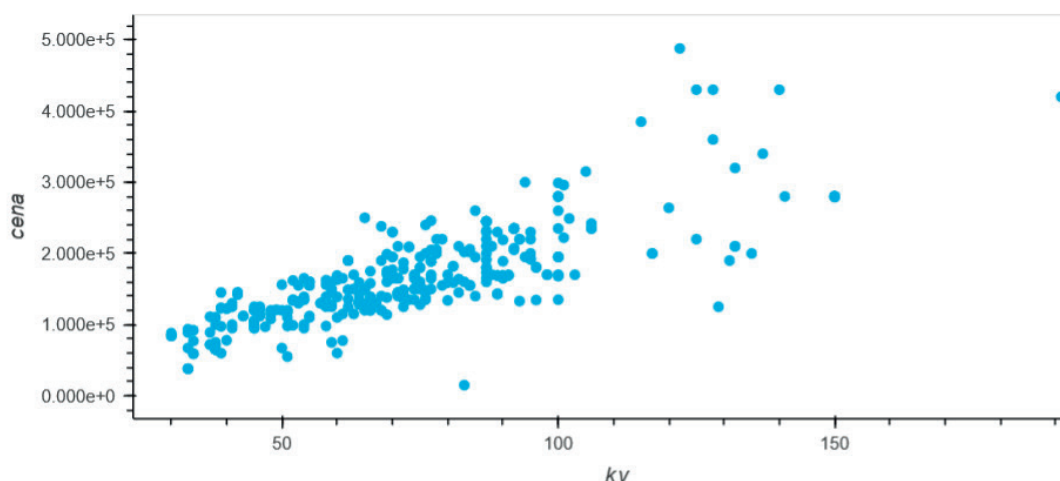
Za obradu podataka, formiranje i validaciju modela i poređenje modela korišćena je platforma *Google Colab* u *Python* programskom kodu (biblioteke *pandas*, *numpy*, *matplotlib* i *sklearn*) [7].

### 3. EKSPERIMENT

U zadatku su korišćeni podaci o 400 stanova starogradnje na teritoriji opštine Novi Beograd, objavljenih za prodaju u maju 2022. godine. Podaci o stanovima su uzeti sa interneta, sa sajta o nekretninama (nekretnine.com) i na osnovu tih podataka formirana je tabela sa obeležjima stanova koji utiču na cenu stana.

Podaci (obeležja stanova), koji utiču na formiranje cene stana i koriste se za mašinsko učenje pre formiranja modela su prošli postupak preprocesiranja koji podrazumeva:

- proveru da li ima praznih polja sa nepostojećim podacima,
- pretvaranje tekstualnih podataka u numeričke,
- proveru postojanja korelacije među podacima,
- skaliranje podataka.

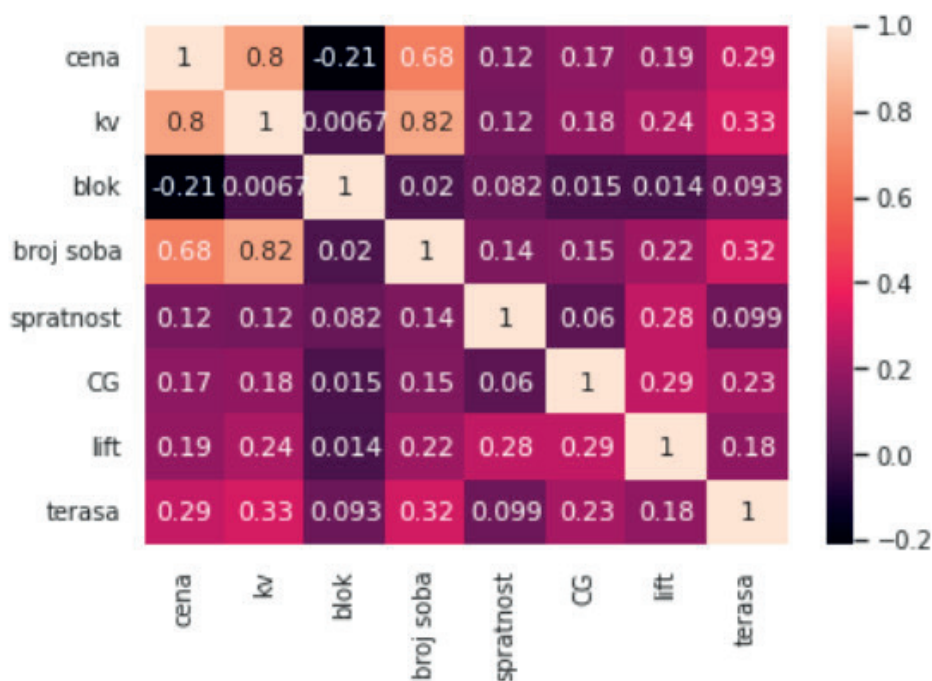


Slika 2. Scatter dijagram vrednosti atributa kvadrati i cena

Urađena je provera da li ima polja sa nepostojećim podacima (*non-null count*). Lokacija je uvrštena u obeležja tako da su stanovi razvrstani u zone tj. blokove na osnovu mape. Tekstualni podaci su pretvoreni u numeričke tako da su nazivi lokacije (tekst) zamenjene brojem bloka, na primer lokacija „Hotel Jugoslavija“ zamenjeni brojem bloka 11. Tekstualni podaci, da li ima lift ili nema, zamenjeni su numeričkim vrednostima 1 ili 0.

Na Slici 2 se vidi da postoji visoka korelacija između atributa kv i cena što je i prirodno, sa porastom broja kvadrata raste i cena stana.

Ulazni podaci su pokazali korelaciju, koja je numerički prikazana u tabeli *heatmap*. Ako postoji korelacije među obeležjima (atributima), i ako je visoka ne utiču dobro na tačnost predikcije za formirani model i preporuka je da se izbace kolone koje sadrže visoko korelisane attribute. Na Slici 3 se uočava korelacija između atributa kv i broj soba (0,82), što je očekivano, sa povećanjem kvadrata raste i broj soba. Ni jedan atribut nije izbačen iz tabele atributa, svi su zadržani, tako da je u ovom radu korišćeno 7 atributa (kv, blok, broj soba, spratnost, CG, lift, teresa), kao ulazni podaci u modelu. Kao izlazni podatak modela korišćena je cena.



**Slika3.** Heatmap-a prikaz korelacije izmedju različitih atributa

Od unetih podataka 70% podataka je uzeto za treniranje modela, a 30% za testiranje modela. Podaci su birani slučajnim uzorkom. Podaci su skalirani na opseg vrednosti [0,1]. Kreirani su modeli sa različitim vrstama regresora i upoređena tačnost modela preko metrika MAE, MSE, RMSE i  $R^2$  (korišćena je *Python* biblioteka *scikit-learn*) i rezultati su prikazani u Tabeli 2. Tačnost koja je dobijena za model koji se najbolje pokazao, Random Forest, iznosi 0.80.

**Tabela 2:** Poređenje tačnosti dobijenih modela regresije uz korišćenje različitih regresora (bez kolone novogradnja i svih podataka o stanovima koji spadaju u novogradnju)

	Model	MAE	MSE	RMSE	R2 Square	Cross Validation
0	Linear Regression	26975.612008	1.638361e+09	40476.671942	0.680941	0.625006
1	Polynomail Regression	28282.882386	1.622080e+09	40275.059260	0.684112	0.625006
2	Robust Regression	34880.964273	2.892578e+09	53782.694723	0.436692	0.562518
3	Ridge Regression	28922.909610	1.999190e+09	44712.300888	0.610673	0.625340
4	SVM Regressor	23247.018000	1.349620e+09	36737.180458	0.737171	0.000000
5	Random Forest Regressor	19157.162901	1.004423e+09	31692.631547	0.804396	0.000000
6	Stochastic Gradient Descent	26975.720906	1.638381e+09	40476.913884	0.680938	0.000000
7	Elastic Net Regression	26999.429644	1.647986e+09	40595.391710	0.679067	0.632290
8	Lasso Regression	29861.323214	1.920316e+09	43821.415430	0.626033	0.625010

Ulazna tabela za testiranje je dopunjena podacima o 60 stanova koji spadaju u novogradnju i dodat je atribut novogradnja (sa numeričkim vrednostima polja 1 - za da i 0 - za ne). Nad ovako formiranim podacima ponovljen je prethodno navedeni postupak modeliranja i rezultati predikcije su prikazani u Tabeli 3. Uočava se opadanje tačnosti predikcije modela. Tačnost koja je dobijena za model koji se najbolje pokazao za drugo testiranje (takođe model Random Forest) iznosi 0.65.

**Tabela 3:** Poređenje tačnosti dobijenih modela regresije uz korišćenje različitih regresora (ubačena kolona novogradnja i svi podaci o stanovima koji spadaju u novogradnju)

	Model	MAE	MSE	RMSE	R2 Square	Cross Validation
0	Linear Regression	36792.068876	4.332000e+09	65817.933062	0.591509	0.660827
1	Polynomial Regression	38149.398438	5.697132e+09	75479.347569	0.462782	0.660827
2	Robust Regression	43979.323768	7.562279e+09	86961.363174	0.286905	0.556518
3	Ridge Regression	36164.981467	4.658806e+09	68255.449031	0.560692	0.660591
4	Ridge Regression	36164.981467	4.658806e+09	68255.449031	0.560692	0.660591
5	SVM Regressor	32114.781743	6.038571e+09	77708.240581	0.430585	0.000000
6	Random Forest Regressor	28497.968461	3.622544e+09	60187.574439	0.658408	0.000000
7	Stochastic Gradient Descent	36791.784008	4.331974e+09	65817.733563	0.591511	0.000000
8	Elastic Net Regression	36556.610746	4.326896e+09	65779.147486	0.591990	0.571024

#### 4. ZAKLJUČAK

Cilj ovog istraživanja je bio da se metode mašinskog učenja primene na procenu cene stanova, na teritoriji opštine Novi Beograd za period maj mesec 2022. godinu. Rezultati eksperimenta su pokazali da se metoda mašinskog učenja, regresija, može primeniti na postupak predviđanja cene stana. Tačnost koja je dobijena za model koji se najbolje pokazao, za Random Forest regresor je 0.8. Osim izborom regresora, tačnost modela se može povećati na više načina: dodavanjem ili oduzimanjem pojedinih atributa, povećavanjem broja ulaznih podataka za treniranje i testiranje (zbog nepostojanja većeg broja podataka u radu su korišćeni podaci o samo 400 stanova, a poželjno bi bilo par hiljada da bi to bio reprezentativan statistički uzorak).

Uključivanje kolone novogradnja u tabelu atributa, da bi se dobili potpuniji i realniji podaci o tržištu stanova, je poželjno, ali u tom slučaju bi trebalo voditi računa da su podaci balansirani tj. da je približno isti broj stanova starogradnje i novogradnje uključeno u istraživanje.

#### LITERATURA

- [1] Ceh M, Kilibarda M, Liseć A, Bajat B. (2018). Estimating the Performance of Random Forest versus Multiple Regression for Predicting Prices of the Apartments, *International Journal of Geo-Information*, 7, 168.
- [2] Ilić A, Stojanović Z, Srdić S. (2018). Assistance of geographical information system in real estate appraisal, *Zbornik radova "SYM-OP-IS 2018"*, Beograd, 68-72.
- [3] <https://www.rgz.gov.rs/usluge/procena-i-vođenje-vrednosti-nepokretnosti>
- [4] <https://mtt.gov.rs/extfile/sr/33340/PRIRUCNIK.pdf>
- [5] <https://www.urbel.com/srp/planovi/biblioteka-planova>
- [6] <https://www.rgz.gov.rs/content/Datoteke/masovna-procena/2022/ProfGodisnji2021.pdf>
- [7] Raschka S, Mirjalili V. (2020). Python mašinsko učenje. *Kompjuter biblioteka*, Beograd, 315-351.



## MODEL UPOTREBE PRENOSNIH TABLET RAČUNARA U POSTUPKU PROVERE I DOPUNE SADRŽAJA TOPOGRAFSKIH KARATA

### MODEL OF USING PORTABLE TABLETS IN THE PROCESS OF CHECKING AND UPDATING THE CONTENTS OF TOPOGRAPHIC MAPS.

SINIŠA DROBNJAK<sup>1,2</sup>, MARKO STOJANOVIĆ<sup>1,2</sup>, NENAD GALJAK<sup>2</sup>, DEJAN ĐORĐEVIĆ<sup>1,2</sup>

<sup>1</sup> Vojnogeografski institut, [sinisa.drobnjak@vs.rs](mailto:sinisa.drobnjak@vs.rs), [stojanovicm80@yahoo.com](mailto:stojanovicm80@yahoo.com), [dejandjordjevic.vgi@gmail.com](mailto:dejandjordjevic.vgi@gmail.com)

<sup>2</sup> Univerzitet odbrane, Vojna akademija, Katedra prirodno matematičkih nauka, [galjaknenad@gmail.com](mailto:galjaknenad@gmail.com)

**Rezime:** Topografska karta u razmeri 1:25000 (TK25) je proizvod Vojnogeografskog instituta, zasnovan na GIS tehnologiji i premeru terena digitalnom fotogrametrijskom tehnologijom koja obuhvata aerofoto snimanje terena, digitalnu fotogrametrijsku restituciju i kartiranje sadržaja, proveru i dopunu kartiranog sadržaja na terenu, finalnu digitalnu kartografsku obradu i vizuelizaciju. U tom tehnološkom lancu izuzetno važno mesto pripada terenskoj proveru i dopuni sadržaja, bez koje je topografska karta gotovo neupotrebljiva. U ovom radu je opisan model upotrebe prenosnih tablet računara u postupku provere i dopune sadržaja TK25.

**Glavne reči:** Topografska karta, centralna baza podataka, terenska provera i dopuna, tablet računari, replika, sinhronizacija.

**Abstract:** Digital topographic map at scale 1:25000 (TM25) is a product of the Military Geographical Institute, which is based on GIS technology and surveying by digital photogrammetrical technology and which includes areal surveying, digital photogrammetrical restitution, mapping, data checking, and updating on the terrain, final cartographic processing, and visualization. This paper describes a model of using portable tablets in the process of checking and updating the contents of TM25.

**Keywords:** Topographic map, Central Geodatabase, terrain checking and updating, tablet computer, replica, synchronization

## 1. UVOD

Terenska provera i dopuna sadržaja je deo tehnološkog lanca izrade topografske karte 1:25000 (TK25), koji ima vitalni značaj za tačnost i ažurnost podataka koji se prikazuju, odnosno unose u Centralnu geoprostornu bazu podataka (CGBP) TK25.

Kao deo jedinstvenog tehnološkog procesa, uz određene specifičnosti usled rada van okruženja računarske mreže Vojnogeografskog instituta (VGI), terenska provera i dopuna povezuje i odlučujuće utiče na kvalitet i pozdanost rezultata rada na izradi TK25 dve veoma značajne tehnološke celine:

- Primarno prikupljanje podataka fotogrametrijskim kartiranjem, odnosno izrade „prve“ verzije sadržaja baze podataka TK25 i
- Kartografsko modelovanje prikupljenog sadržaja TK25, odnosno završnu kartografsku obradu baze podataka TK25.

Terenska provera i dopuna sadržaja TK25 nije samo specifična zbog mesta u tehnološkom lancu izrade TK25, nego je i svojevrsna prostorna i vremenska kontrola kvaliteta primarno prikupljenih podataka, zbog toga što predstavlja, uz aerofotogrametrijsko snimanje, fazu rada koji se odvija u realnim uslovima.

Kvalitet i pouzdanost informacija TK25 ne mogu se zamisliti bez verifikacije u realnim, terenskim uslovima. Ova aktivnost ne zadire u probleme matematičke tačnosti, kartografskog modelovanja sadržaja ili logičke strukture podataka, ali može doprineti njihovom unapređenju. Glavni cilj ove tehnološke faze, koja prethodi završnoj kartografskoj obradi podataka je kvalitativna i jednoznačna provera sadržaja baze podataka TK25, dobijenog na osnovu fotogrametrijske obrade aerofoto snimaka i ukoliko je neophodno, dopuna sadržaja detaljnima nastalim nakon završetka fotogrametrijske obrade.

U okruženju CGBP TK25, u kojem se svaka faza odvija prema unapred utvrđenom logičkom i fizičkom modelu podataka, terenska provera i dopuna sadržaja nema neke posebne specifičnosti. Rad na terenu se odvija primenom identičnih softverskih alata na prenosivom uređaju za prikupljanje podataka, opremljenim

modulom za GPS navigaciju. Svaki rukovodilac terenske ekipe mora da vodi računa o prikupljenim podacima, tako što redovno evidentira izmene u replici CGBP koja je kreirana pre polaska na zadatak. Sinhronizacija izmena saržaja TK25 u CGBP vrši se nakon povratka terenskih ekipa u VGI, a prema ukazanoj potrebi može se vršiti i u toku terenskih radova, ukoliko postoji mogućnost za dolazak rukovodioca terenske ekipe u VGI.

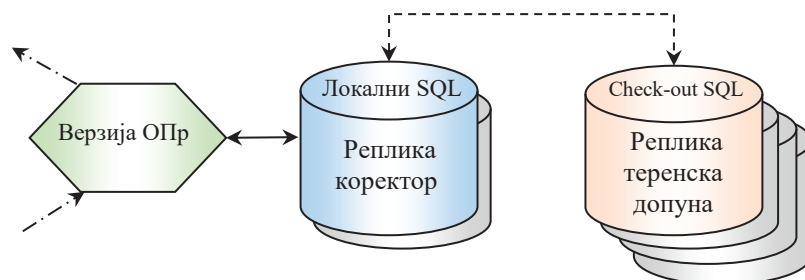
Veoma važan element sadržaja TK25 predstavljaju podaci o elektroenergetskoj mreži, šumskom pokrivaču, tačke geodetske osnove, fabrički i sportski kompleksi i drugi objekti javnog značaja za koje su nadležna javna preduzeća Republike Srbije, lokalni organi Republičkog geodetskog zavoda i privredni subjekti. Sa njima VGI uspostavlja i održava saradnju radi razmene podataka o elementima sadržaja CGBP TK25 koji imaju širi, nacionalni značaj.

Posebno osetljiv element sadržaja TK25, kome se i u fazi pripreme i prilikom rada na terenu posvećuje posebna pažnja su geografski nazivi. Radi se o elementu sadržaja koji je najteže prikupiti usled neažurne dokumentacije kojom raspolažu lokalni državni organi i usled depopulacije u pograničnim i planinskim krajevima, kao i promene nacionalne strukture stanovništva u pojedinim oblastima Republike Srbije.

## 2. TERENSKA PROVERA I DOPUNA SADRŽAJA BAZE PODATAKA TK25

Centralna geoprostorna baza podataka TK25 zasnovana je na hijerarhijsko-relacionom modelu podataka. Opšte tematske celine iskazane su logičkim modelom podataka kao tematske klase, u okviru kojih su organizovane tematske tabele [1]. U svakoj tematskoj tabeli jedna ili više kolona namenjena je za tekst, na osnovu kojeg se kreiraju anotacije, kao najkompleksniji tip podataka. Nazivi naseljenih mesta, predela i reljefnih oblika nisu obični tekstualni zapisi, već su deo jedinstvene baze podataka u okviru opšte tematske klase Nazivi [1].

Blok šema terenska provere i dopune sadržaja CGBP TK25 prikazana je na slici 1, i njom je definisana posebna verzija na nivou CGBP TK25 namenjena za unos sadržaja terenske provere i dopune, koja se naziva Verzija OPr [2].



Slika 1: Blok šema faze terenske provere i dopune sadržaja CGBP

Sadržaj novog izdanja TK25 ne razlikuje se u načelu od sadržaja prethodnog, ali je model podataka prilagođen radu sa vektorskim elementima sadržaja, koji su raščlanjeni na tematske celine (feature datasets) i povezani u jedinstvenu geokodiranu bazu podataka. Baza podataka DTK25 sastoji se od 5 tematskih celina (tabela 1), unutar kojih je svaki element sadržaja definisan kao tačka, linija, poligon ili tekst [3].

Tabela 1: Broj elemenata sadržaja CGBP TK25 razvrstanih po objektnim klasama

Tematska klasa	Tačkasti elementi	Linijski elementi	Poligoni
Nazivi	67	-	-
Gerip	138	679	94
Hidrografija	17	198	32
Reljef	4	18	7
Vegetacija	3	3	8
UKUPNO	229	898	141

Terenska provera i dopuna originala TK25, u uslovima trenutno dostignutog tehnološkog nivoa u VGI ne uključuje mogućnost direktnih izmena podataka u bazi dok je izvršilac na terenu. Taj proces se realizuje naknadnim unosom na osnovu prikupljenih .shp tačkastih, linijskih i površinskih datoteka, tekstualnih zapisa i fotografija objekata (mostova, nadvožnjaka i sl.).

### 3. FORMIRANJE ELABORATA TERENSKE PROVERE I DOPUNE TK25

Nakon prijema i pregleda kartiranog sadržaja CGBP TK25 koji je kartiran u odeljenju za fotogrametriju (OF), eventualno dorade na osnovu primedbi nadležnog starešine za terensku proveru i dopunu sadržaja TK25, pristupa se pripremama za terenske radove. Priprema za terensku dopunu obuhvata preuzimanje podataka iz CGBP, kompletiranje terenskog elaborata, kreiranje planiranje (procenu) vremena, prenos podataka u GPS prijemnik i ličnu pripremu izvršioca terenskih radova.

Terenski elaborat sačinjavaju [4]:

- original terenske dopune – replika CGBP na prenosivom računaru;
- karta pripreme;
- karta promene naziva sa oleatom i spiskom naziva za proveru;
- karta terenske dopune;
- oleata specijalnog sadržaja;
- radna sveska;
- ortofoto u digitalnom i analognom obliku;
- karta TK25 i TK50 prethodno izdanje u analognom i digitalnom obliku.

Uvidom u dokumenata terenskog elaborata procenjuje se potrebno vreme za terensku dopunu, a na osnovu obima i vrste izmena na terenu. Vreme za terensku dopunu procenjuje se u radnim danima i upisuje u radnu svesku terenskog elaborata.

Nakon izvršene procene vremena potrebnog za terensku dopunu, nadležni starešina za poslove terenske dopune u VGI određuje za svaku terensku ekipu radilište. Radilište, svake terenske ekipe, treba da čini geografsku celinu sa mogućnošću prilaza komunikacijama svim njegovim delovima.

Nakon prijema zadatka izvršioци terenske dopune (terenske ekipe) pripremaju se za terensku dopunu, koja obuhvata sledeće [4]:

- proučavanje Upustva za izradu TK25;
- kreiranje tekućeg projekta na prenosivom GIS uređaju za prikupljanje podataka;
- upoznavanje sa sadržajem svakog list karte svog radilišta;
- proučavanje ekonomskih i geografskih karakteristika radilišta;
- upoznavanje sa obimom i vrstom kartiranih izmena i dopuna;
- plotiranje ortofoto-a u razmeri karte, dok za urbana područja plotirati u krupnijoj razmeri (1:10000 ili 1:5000);
- plotiranje karte pripreme sa posebno naznačenim elementima sadržaja koje je neophodno proveriti na terenu;
- plotiranje karte terenske dopune sa geografskim nazivima, u koju se unose opšta zapažanja o svim proverenim i izmenjenim elementima sadržaja TK25;
- plotiranje karte terenske dopune bez geografskih nazivima, u koju se unose posebna zapažanja o provernim i izmenjenim posebno naglašenim elementima sadržaja (dalekovodi, gasovodi, ostali linijski i tačkasti elementi specijalnog sadržaja TK25, podaci o šumama i voćnjacima i slično);
- prenošenje sadržaja sa planova i karata krupnijih razmera u original terenske dopune koji se nije jednoznačno mogao identifikovati i kartirati sa raspoloživih podloga u prethodnoj fazi izrade TK25;
- upoređenje neslaganja geografskih naziva na različitim izdanjima TK25 i TK50 i unošenje uočenih neslaganja u listu promene naziva;
- zaduženje literature i prateće dokumentacije (reversi, pravilnici, stručna literatura);
- provera kompletnosti i ispravnosti tehničkih sredstava;
- provera kompletnosti elaborata;

#### 3.1. Original terenske dopune – replika CGBP na prenosnom računaru

Nakon završetka fotogrametrijskog kartiranja nadležni starešina organizacione jedinice koja izvodi terenske radove podnosi zahtev administratoru CGBP u VGI za izradu jedne ili više replika CGBP TK25 koje obuhvataju teritoriju na kojoj se vrši terenska provera i dopuna.

Ova replika, u verziji organizacione jedinice koja je nadležna za terenske geodetske radove, služi da se od nje kreiraju terenske replike na prenosivim računarima i odgovarajući tekući projekti na GIS uređajima za prikupljanje podataka. Na osnovu ove replike izrađuju se datoteke za štampanje karte pripreme (slika 2) i karte terenske dopune.

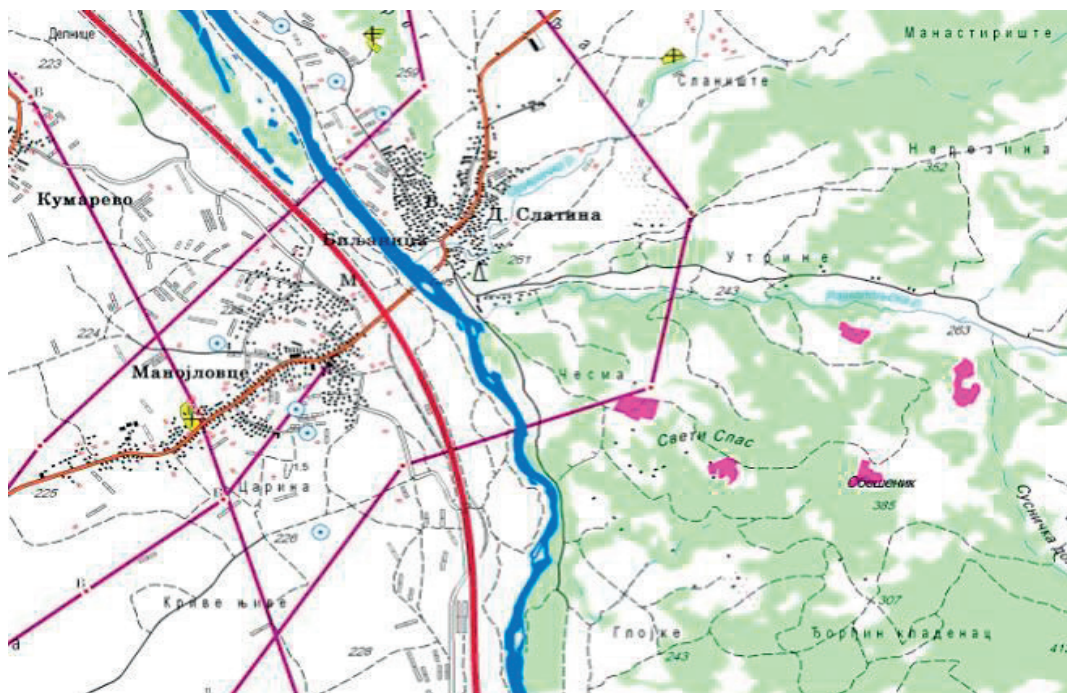
Osnovna uloga replike CGBP TK25 u verziji organizacione jedinice VGI koja izvodi terenske radove je zaštita podataka u centralnoj bazi na taj način što se izmene u replikama terenskih izvršilaca najpre sinhronizuju u ovu repliku, a tek nakon kontrole od strane nadležnog stručnog lica vrši se sinhronizacija u centralnu bazu [4].

### 3.2 Karta pripreme

Za potrebe izrade štampane verzije grafičkog sadržaja lista TK25 koji se dopunjuje na terenu, izrađuje se posebna grafička datoteka u kojoj se posebnim simbolima ističu objekti i delovi sadržaja TK25 na koje treba obratiti posebnu pažnju, a koji su na taj način lakše uočljivi licu koje izvodi terensku proveru [4].

Ova datoteka po pravilu ne sadrži temu geografski nazivi, ali može da sadrži u zavisnosti od gustine prikazanog sadržaja na listu TK25. Priprema datoteke iz koje se plotira karta terenske pripreme vrši se na sledeći način (slika 2):

- podese se transparentcija 50% za temu šume, putevi i stambeni objekti (kućice);
- izvrši se promena simbologije teme voćnjaci, tako da bude naglašena razlika u odnosu na šume, tako da budu prikazani ljubičastom bojom;
- podebljaju se svi linijski elemnti specijalnog sadržaja (dalekovodi, gasovodi i toplovodi);
- izvrši se promena simbologije teme objekti javnog značaja 1 i objekti u naselju 1 (izuzimajući stambene objekte), tako da elementi budu prikazani većim, crveno obojenim simbolom;
- elementima teme hidrografija 2 svetle boje dodeli se tamnija, čitljiva nijansa.



Slika 2: Karta pripreme sa naglašenim elementima sadržaja koji moraju biti provereni na terenu

## 4. PRIPREMA GPS TRIMBLE T10 UREĐAJA ZA TERENSKU PROVERU I DOPUNU TK25

Trimble je objedinio snagu desktop računara i praktičnost velikog ekrana u izuzetnom tablet računaru prilagođenom terenskim aplikacijama. Trimble® T10 je snažan i robusan uređaj za geodetski premer i prikupljanje podataka za GIS, koji može da izvršava i zahtevne desktop aplikacije, za efikasnu obradu podataka na terenu, sve u jednom uređaju. Projektovan da podrži celodnevni rad u složenim terenskim i vremenskim uslovima, Trimble T10 omogućuje primenu najrazličitijih aplikacija za premer, prikupljanje i analizu podataka na terenu, uz prikaz na izuzetnom i velikom ekranu, oslonjen na operativni sistem Microsoft Windows® 10.

Snažan računarski sistem Trimble T10 odgovara na sve izazove bez ikakvih teškoća. Brza obrada digitalnih karata, satelitskih snimaka, fotografija, oblaka tačaka i drugih podataka o prostoru, odlike su T10. Uštedite vreme i novac povećanjem produktivnosti na terenu, bržim završetkom posla, sa više detalja, povećavajući na taj način vrednost vaših podataka.

Bilo da se radi o geodetskom premeru, prikupljanju podataka za GIS, inspekciji na terenu ili drugim zadacima, veliki ekran od 25,6 cm (10,1"), čitljiv na dnevnom svetlu i direktnom suncu, omogućuje udoban

rad i pregled vaših podataka. Visoka rezolucija ekrana obezbeđuje lak rad sa digitalnim kartama, pregled i analize, olakšavajući donošenje odluka na terenu (slika 3).



**Slika 3:** Terenska provera i dopuna sadržaja TK25

Aplikacije koje se koriste za prikupljanje podataka je Arc Pad (izabrani softver u VGI). Nakon pokretanja softvera automatski se povezuje GPS prijemnik, koji potom počinje da traži satelite.

Uređaj je pripremljen za terensku dopunu ukoliko je [5]:

- kreiran novi projekat,
- podešen koordinatni sistem,
- podešena preciznost očitavanja koordinata,
- učitane odgovarajuće grafičke datoteke (digitalne podloge),
- kreirane tabele za unošenje opisnih atributa na osnovu kojih će za svaki snimljeni detalj u post-processing obradi biti izabran odgovarajući znak iz kartografskog ključa i ostali podaci od značaja za ažuriranje Centralne baze podataka,
- proverena tehnička ispravnost (baterije, oprema, prijem satelita) i
- etaloniran i označen propisnom nalepnicom.

Za prikaz podataka koristi se georeferencirana karta kao pozadinski layer. Objekti na karti se prikazuju predviđenim simbolima. Klikom na simbol objekta, dobijaju se informacije o koordinatama objekta, visini u izabranom sistemu visina (nadmorska, elipsoidna) i vrednostima nekih atributa [6].

Prilikom snimanja detalja na terenu, osim o atributskim karakteristikama objekta, mora se voditi računa i o topološkim pravilima, što najviše dolazi do izražaja kod linijskih elemenata.

## 5. PRIKUPLJANJE PODATAKA NA TERENU

Terenska dopuna se obavlja po planu sačinjenom u periodu pripreme. Od zadatog plana može se odstupiti ako se ustanovi da se na dotičnoj teritoriji izvode radovi koji će uticati na promene u sadržaju karte, a koji će do kraja terenske dopune će biti u fazi koja daje mogućnost unošenja tih promena.

Kriterijumi za izmene i dopune sadržaja definisani su uputstvom za izradu TK25. Jedan od polaznih činilaca koji utiču na izbor elemenata je razmer karte, a to je odnos dužina i površina prikazanih na karti 1:25 000 i u prirodi: 1mm na karti = 25m u prirodi; 1mm<sup>2</sup> na karti = 625 m<sup>2</sup> u prirodi.

Uputstvom za izradu TK25 jasno definiše koji se elementi sadržaja prikazuju u razmeru, a koji uslovnim znakom, što najviše zavisi od površine prikazanog prirodnog ili izgrađenog objekta (areala).

Izvršioци radova prikupljaju podatke shodno stvarnom stanju na terenu. Prikupljeni podaci se unose na list terenske dopune, kartu terenske dopune, ortofoto, oleatu promene naziva, spisak provere naziva, radnu svesku.

Obrada podataka prikupljenih na terenu vrši se svakodnevno posle terenske dopune. Svi podaci prikupljeni na terenu moraju biti uneseni u odgovarajuću repliku CGBP TK25. Svaki dan po povratku sa terena prikupljeni podaci iz GPS uređaja prebacuju se u na PC računar, a zatim se učitavaju i u projekat. Nakon identifikacije i izbora odgovarajućeg znaka iz jedinstvenog šifarnika, u bazu podataka se unose tekstualni i broječno-slovni podaci (slika 4).



**Slika 4:** Rad u terenskim uslovima sa Trimble T10

Nakon završetka terenske provere i dopune sadržaja TK25 nadležno lice iz VGI kontroliše da li su svi podaci pravilno prikupljeni na terenu i da li su uneti u bazu podataka. Takođe kontroliše da li su terenski izvršioци postupili u skladu sa preporukama stručnog lica koje je vršilo terensku pripremu, da li su provereni svi nazivi sa spisaka naziva predviđenih za proveru na terenu i da li su svi novoprikupljeni i provereni nazivi ispisani na oleati promene naziva. Ako se ustanovi da su svi podaci pravilno prikupljeni i uneseni u bazu podataka, promene u projektu koje su unesene u odnosu na sadržaj projekta pre terenskih radova sinhronizuju se u CGBP TK25. Time prestaje direktan rad terenskih izvršilaca nad replikom CGBP TK25, ali ne i tumačenje određenih rešenja koja su primenjena na terenu koja mogu zahtevati korektori ili druga stručna lica u fazi završne kartografske obrade.

## 6. ZAKLJUČAK

Baza podataka TK25 ima višefunkcionalnu namenu u procesu prikupljanja, upravljanja, analize i prezentacije podataka koji predstavljaju sadržaj TK25 i ostalih TK razmernog niza. Provera i dopuna sadržaja baze podataka TK25 ima odlučujući značaj u pogledu obezbeđenja njenog kvaliteta i funkcionalnosti i predstavlja jedan od najznačajnijih koraka u tehnološkom lancu izrade TK25. Zbog toga terensku proveru i dopunu sadržaja treba posmatrati kao proces pomoću kojeg je moguće posmatrati određene fenomene u realnim uslovima i na osnovu povratnih informacija neprekidno poboljšavati ne samo strukturu i model podataka u CGBP TK25, nego i osmišljavati funkcionalniji i ekonomičniji pristup prikupljanju, prikazu, analizi i prezentaciji podataka koji čine geografski informacioni sistem.

## LITERATURA

- [1] Zeiler, M. (1999). *Modeling of our world*, Environmental Systems Research Institute, Inc, Redlands, California.
- [2] Drobnyak, S., Radojčić, S., Božić, B. (2014) Primena ISO 19157 standarda u tehnološkom procesu izrade digitalnih topografskih karata, *Tehnika*, vol.66, br.4, str. 551-556.
- [3] Marković, V. (2010). *Logički model podataka digitalne topografske karte u razmeri 1:25000*, Vojnogeografski institut, Beograd.
- [4] Vojnogeografski institut. (2016). *Uputstvo za terensku proveru i dopunu sadržaja TK*, Vojnogeografski institut, Beograd.
- [5] Srdić, Z. (2010). *Geoexplorer CE, osnovno uputstvo za rukovanje i održavanje*, Vojnoizdavački zavod, Beograd.
- [6] Vojnogeografski institut. (2016). *Uputstvo za izradu digitalne topografske karte u razmeri 1:25000*, Vojnogeografski institut, Beograd.

## VIŠEKRITERIJUMSKI MODEL OPTIMIZACIJE PROCENE UGROŽENOSTI OD ŠUMSKIH POŽARA PRIMENOM GIS-A

### MULTICRITERIA DECISION MODEL OF FOREST RISK ASSESSMENT BY GIS APPLICATION

LJUBOMIR GIGOVIĆ<sup>1</sup>, DARKO LUKIĆ<sup>1</sup>, NENAD GALJAK<sup>1</sup>, DRAGOLJUB SEKULOVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet odbrane u Beogradu, Vojna akademija, gigoviclj@gmail.com

<sup>2</sup> Univerzitet „Union-Nikola Tesla“ u Beogradu, Fakultet za IT i inženjerstvo, sekulovicdr@yahoo.co.uk

**Rezime:** U radu je izvršena klasifikacija i kartiranje rizika od pojave šumskih požara sa ciljem predviđanja i smanjenja njihove učestalosti i ekološke štete koje izazivaju. Predloženi model se zasniva na kombinovanoj primeni Geografskih informacionih sistema (GIS) i višekriterijumskog odlučivanja (MCDA) korišćenjem fazi logike i analitičkog hijerarhijskog procesa (AHP), sa ciljem procene i mapiranja rizika od požara na studiji oblasti Nacionalnog parka „Đerdap“. Postupak je razvijen uz pomoć 8 kriterijuma. Primenom Fuzzy logike izvršena je standardizacija kriterijuma, dok su AHP-om izračunati težinski koeficijenti u odnosu na njihovu osetljivost i sposobnost da iniciraju nastanak požara. Konačna mapa rizika od šumskih požara klasifikovana je u 4 kategorija, od veoma niskog do veoma visokog rizika. Predložena metoda i rezultati ovog rada mogu se koristiti za politiku održivog razvoja na svim nivoima državne uprave.

**Ključne reči:** Rizik, Šumski požari, Geografski informacioni sistem (GIS), Fuzzy logika, AHP, Đerdap

**Abstract:** The paper describes the process of mapping forest fire risk in whose prediction and prevention. The proposed model is based on the combined application of Geographic Information Systems (GIS) and multi-criteria decision-making (MCDA) using fuzzy logic and analytic hierarchy process (AHP), with a view to evaluating and mapping the risk of fires in the study area of the National Park "Djerdap". The procedure was developed based on the 8 criteria. Fuzzy logic is applied in the standardization of criteria, while the AHP to a calculated weight coefficients with respect to their sensitivity and the ability to initiate the occurrence of fire. The final map of the risk of forest fires is classified into 4 categories, from very low to very high risk. The proposed method and the results of this work can be used for a policy of sustainable development at all levels of government.

**Keywords:** Risk, Forest fires, Geographic information system (GIS), Fuzzy logic, AHP, Djerdap

## 1. UVOD

Nedavna istraživanja JRC Technical Reports pokazuju da bi se, usled klimatskih promena povećanja temperature i smanjenja vlažnosti, mogla udvostručiti područja pogođena šumskim požarima u Evropi [1]. Ovaj podatak zabrinjava i upućuje na ozbiljne analize ovog fenomena i preventivno modeliranje rizika od šumskih požara. Rad ima za cilj izradu karata zona rizika od pojave šumskih požara. Zone rizika od pojave šumskih požara su lokacije gde je verovatno da će se zapaliti vatra i odakle se lako može proširiti i na druge susedne oblasti. Njihovom identifikacijom ukazuje se na područja povećanog rizika od nastanka i razvoja požara, što predstavlja osnovu za planove hitnih intervencija. Time se stvaraju povoljni uslovi za minimiziranje broja požara i za uklanjanje uslova za njihovo nastajanje. Geografski informacioni sistemi (GIS), u kombinaciji sa drugim oblicima tehnologije, kao što su daljinska detekcija i kompjutersko modeliranje, se sve više koriste u svim aspektima upravljanja rizika od požara. Multikriterijumska analiza (MCDA) u kombinaciji sa GIS-om je najčešće korišćeni način u procesu kartografskog modeliranja zona rizika od požara. U ovom istraživanju predstavljen je GIS-MCDA-FAHP model izrade karte rizika šumskih požara sa ciljem identifikacije i klasifikacije osetljivijih zona na pojavu požara za područje Nacionalnog parka Đerdap. Izrada mape rizika od šumskih požara može značajno umanjiti rizik od požara i poboljšati preventivne aktivnosti [2].

## 2. STUDIJA OBLASTI

Nacionalni park „Đerdap“ nalazi se u istočnom delu karpatske Srbije. Smešten je između 21°36' i 22°35' istočne geografske dužine i 44°43' i 44°16' severne geografske širine. Prostire se na prostoru Đerdapske klisure od Golubačkog grada do Karataša na dužini od oko 100 km i u širini od 2-8 km. Severna granica parka prati reku Dunav i jasno je određena, dok se približna južna granica proteže razvođem Dunava i Peka, kao i najvišim delovima Liškovca, Velikog grebena i Miroča. Visina ovog prostora se kreće od 50-803 mnv.

Nacionalni park „Đerdap“ se svojim jedinstvenim prirodnim odlikama i specifičnostima i stanjem biodiverziteta i zaštićenim vrstama biljaka i životinja, svrstava među najatraktivnije nacionalne parkove na Balkanu. Deo je EMERALD mreže područja (Emerald Network of Areas of Special Conservation Interest – AsCI) identifikovano/ustanovljeno pod imenom Nacionalni park Đerdap. Emerald područja su značajna sa stanovišta primene Konvencije o očuvanju evropske divlje flore i faune i prirodnih staništa (Bernska konvencija) u Srbiji. Geoprostor Nacionalnog parka je spada u prostore bogatije šumom (63,3%) i predstavlja zonu potencijalnog rizika od šumskih požara, pre svega tokom perioda jul-avgust kad su zabeležene najveće temperature i najmanja količina izlučenih padavina.

### 3. REZULTAT

Metodološki model u ovom radu zasnovan je na prostornoj GIS-MCDA strukturi. Sa metodološke tačke gledišta, predloženi GIS-AHP MCDA model definisanja kriznih zona za požare obuhvata sledeće korake (Slika 1):

1. Definisane cilja/problema i arhitekture modela
  - identifikacija glavnog cilja
  - definisanje mrežne strukture modela
2. Identifikacija kriterijuma osetljivosti na šumske požare
3. Prikupljanje podataka i izgradnja GIS prostorne baze kriterijuma
4. GIS-MCDA evaluacija
  - Individualno vrednovanje kriterijuma
  - Fuzzy standardizacija kriterijuma
  - Formiranje matrice odlučivanja i relativna procena težine kriterijuma (AHP)
  - Rezultati agregacije (Suitability map) (WLC)
5. GIS vizuelizacija konačnog rešenja i preporuke

#### 3.1. Identifikacija kriterijuma

Pravilan izbor kriterijuma od kojih zavisi nastanak šumskih požara predstavlja važnu fazu u prevenciji njihovog nastanka. U istraživanju je učestvovalo šest eksperata sa bogatim iskustvom u oblasti upravljanja rizicima i životne sredine. Tokom intervjuisanja eksperata prikupljeni podaci su obrađeni i izvršena je agregacija njihovih mišljenja. U Tabeli 1 dati su kriterijumi koji su od ključnog značaja za pojavu i ocenu rizika od šumskih požara u Nacionalnom parku „Đerdap“.

**Tabela 1:** Opis kriterijuma uzroka šumskih požara

Kriterijum	Opis kriterijuma
C1 Land use	Korišćenje zemljišta je faktor koji u najvećoj meri utiče na pojavu i širenje požara, jer definiše vegetaciju i dostupnu količinu goriva za požar [3].
C2 Ekspozicija	Osunčani deo terena je najpogodniji za nastanak i širenje požara, jer prima više sunčeve svetlosti pa ima i veću temperaturu, a samim tim i manju vlažnost [3].
C3 Nagib	Nagib terena utiče na ponašanje vatre. Strme padine povećavaju brzinu širenja vatre [4]. Povećanje nagiba za 10% može udvostručiti brzinu širenja vatre.
C4 Visina	Nadmorska visina je faktor koji utiče na ponašanje vetra, a samim tim ima veliku ulogu u širenju požara. Takođe, nadmorska visina utiče na strukturu vegetacije, količinu dostupnog goriva, vlažnost vazduha i temperaturu [4].
C5 Temperatura	Požari nastaju pri svakoj temperaturi, ali da njihova učestalost u zavisnosti od porasta temperature [5].
C6 Padavine	Padavine su činilac koji odlučuje o mogućnosti nastanka i širenja požara, javljaju se u obliku vlažnosti i vlažnosti gorivog materijala. Ako je gorivni materijal suv, bez vlage, gori brže i brže se postižu veće temperature.



C7 Rastojanje od puta	Šume u blizini puteva imaju veći rizik od pojave šumskog požara, 95% šumskih požara na izazvano je ljudskom nepažnjom [6].
C8 Rastojanje od naselja	Utvrđeno da je čovek glavni uzrok požara, tako da se udaljavanjem od njegovog prebivališta broj požara smanjuje.

### 3.2. Prikupljanje podataka i unos u GIS

Podaci korišćeni u ovoj studiji su sastavljeni iz različitih izvora. Svi procesi GIS modelovanja podataka su su izvedeni korišćenjem integrisanih alata ArcGIS 10.6 ESRI softvera. Karte prostorne distribucije temperature i padavina dobijene su korišćenjem Kriging interpolacione metode na osnovu georeferencirane mreže od 12 meteoroloških stanica sa unetim klimatskim vrednostima za referentni period od 35 godina.

### 3.3. GIS-MCDA

S obzirom da su podaci prikupljeni na različite načine i različitog su formata, prvi korak MCDA jeste da svi skupovi podataka budu standardizovani i u jedinicama koje se mogu uporediti. U radu, svi kriterijumi su standardizovan korišćenjem fazi skupova.

**Tabela 2.** Fuzzy standardizacija kriterijuma

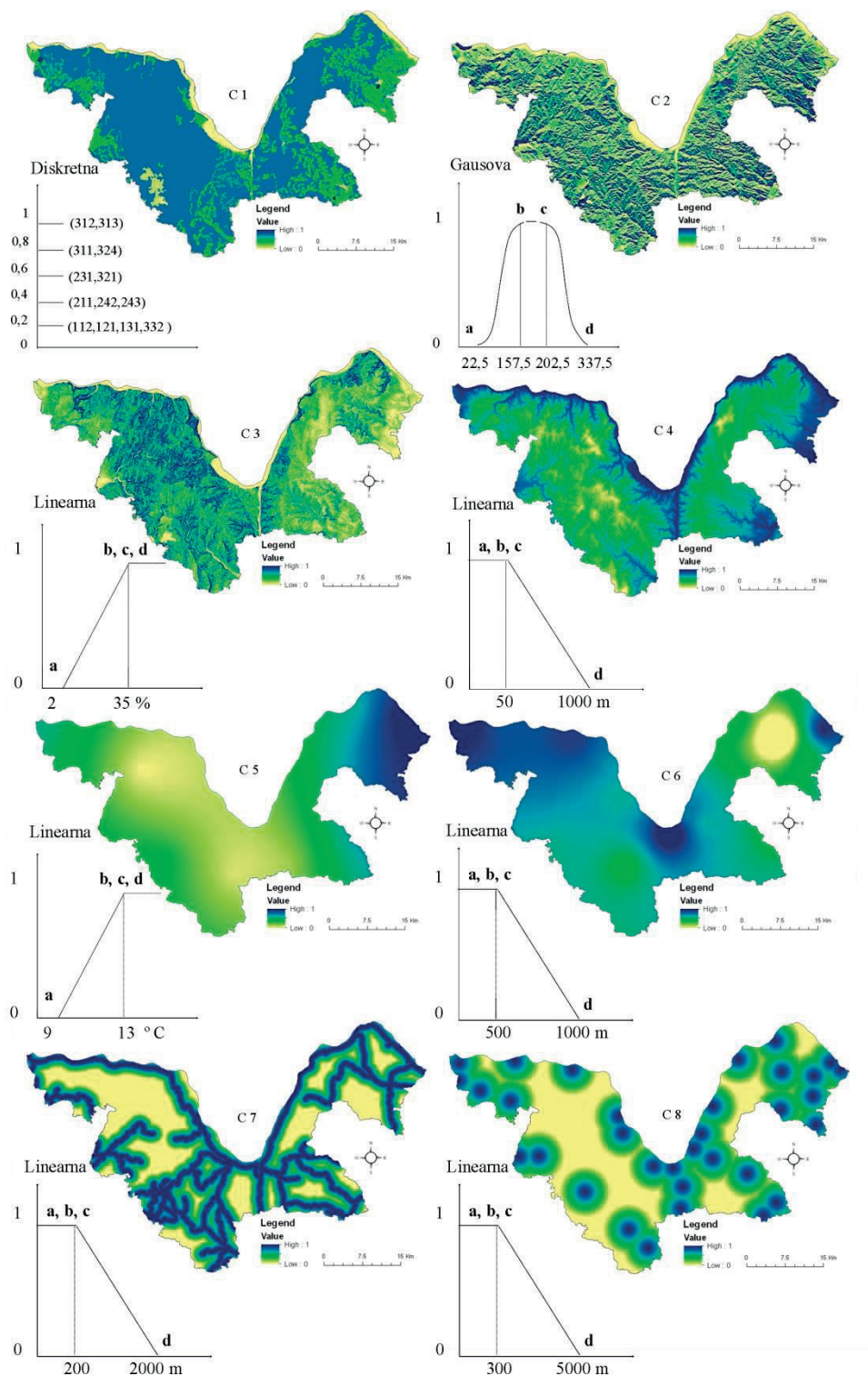
Kriterijum	Fuzzy funkcije	Kontrolne tačke	Opis fuzzy članstva
C1 Korišćenje zemljišta*	Diskretna Kategorični podaci	(511,512,331) jednako 0, (112,121,131,332) jednako 0.2, (211,242,243) jednako 0.4, (231,321) jednako 0.6, (311,324) jednako 0.8 i (312,313) jednako 1	
C2 Ekspozicija	Gausova Simetrična	a= 22,5° c= 202,5° b= 157,5° d= 337,5°	Sever (337,5-22,5°) jednako 0, 22,5-157,5° između 0-1, jug (157,5-202,5°) jednako 1, 202,5-337,5° između 1-0
C3 Nagib	Linearna Monotono rastuća	a= 2 % b= 35 %	0-2% jednako 0, 2- 35% između 0-1, preko 35% jednako to 1
C4 Visina	Linearna Monotono opadajuća	c= 50m d= 1000m	0-50 m jednako 1, 50- 1000m između 1-0, preko 1000m jednako 0
C5 Temperatura	Linearna Monotono rastuća	a= 9 C° b= 13 C°	ispod 9 C° jednako 0, 9-13 C° između 0-1, preko 13 C° jednako 1
C6 Padavine	Linearna Monotono opadajuća	c= 500m d= 1000m	0-500 m jednako 1, 500- 1000m između 1-0, preko 1000m jednako 0
C7 Rastojanje od puta	Linearna Monotono opadajuća	c= 200m d= 2000m	0-200 m jednako 1, 200- 2000m između 1-0, preko 2000m jednako 0
C8 Rastojanje od naselja	Linearna Monotono opadajuća	c= 300m d= 5000m	0-300 m jednako 1, 100- 5000m između 1-0, preko 5000m jednako 0

\* Treći nivo CORINE nomenklature

a = članstvo raste iznad 0; b = članstvo postaje 1; c = članstvo padne ispod 1; d = članstvo postaje 0

S obzirom da ulazni podaci mogu imati diskretne ili kontinuirane vrednosti, korišćene su metode diskretne i kontinuirane fuzzy standardizacije. Diskretna standardizacija, u kojoj eksperti neposredno dodeljuju vrednosti atributa na definisanoj skali pripadnosti, je korišćena za fuzzy standardizaciju kategoričnog kriterijuma - Vegetacija. U standardizaciji ostalih skupova podataka, kod kojih se vrednosti atributa postepeno menjaju sa jedne lokacije na drugu, primenjena je kontinuirana standardizacija. Ovi skupovi podataka su standardizovani primenom fuzzy koncepta na kontinuiranoj skali u zavisnosti od izabrane funkcije pripadnosti. Za vrednovanje podobnost atributa korišćena je kontinuirana skala u opsegu od 0 do 1 bajta, gde je 0 najmanje rizična, a 1 najrizičnija vrednost atributa u odnosu na mogućnost izbivanja

šumskog požara. Standardizovani kriterijumi za ocenu sa fuzzy funkcijama i oblikom članstva prikazani su u Tabeli 2 i Slika 1.



Slika 1: Standardizovani kriterijumi sa oblikom fuzzy članstva

### 3.4. Određivanje težine kriterijuma (AHP)

Pri rešavanju realnih problema, kriterijumi nemaju isti stepen značajnosti, pa je potrebno da donosici odluke definišu značajnost pojedinih kriterijuma izračunavajući odgovarajuće težinske koeficijente (težine) ili pondere za kriterijume. Za izračunavanje normalizovanih težina kriterijuma iskorišćena je višekriterijumska tehnika u vidu AHP [9]. Metodološki posmatrano, AHP je višekriterijumska tehnika koja se zasniva na međusobnom poređenju elemenata na datom hijerarhijskom nivou u odnosu na elemente na

višem nivou. Na vrhu je cilj, ispod su kriterijumi (podkriterijumi, ako postoje) i na dnu su alternative. AHP zahteva da se prvo međusobno porede kriterijumi i izračunaju njihove relativne težine u odnosu na cilj. Alternative se zatim porede u parovima u odnosu na svaki kriterijum i analognim postupkom određuju se njihove relativne težine u odnosu na kriterijume.

Posmatrajući definisani cilj za svaki par kriterijuma su unešene vrednosti značaja jednog u odnosu na drugi. Na taj način polja po dijagonali matrice iznose 1. Nakon unošenja vrednosti iz Satijeve skale [9] u matricu poređenja, izračunavaju se težinske vrednosti kriterijuma ( $w_i$ ). Matrica poređenja na nivou klastera prikazana je u Tabeli 3.

**Tabela 3.** Matrica poređenja i težine kriterijuma

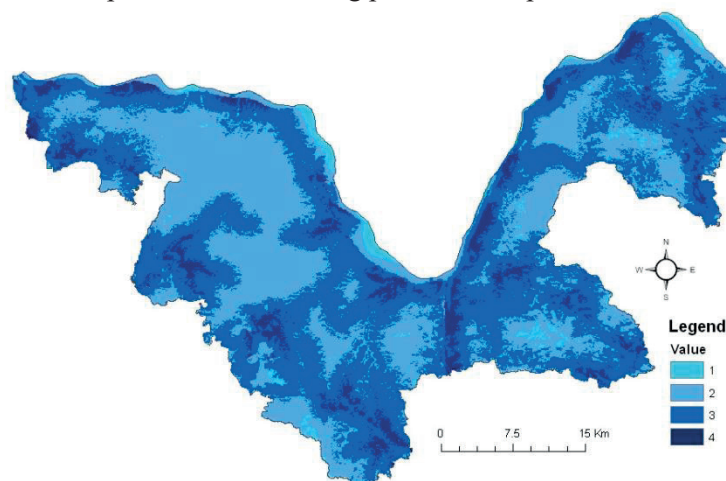
Kriterijum	C1	C2	C3	C4	C5	C6	C7	C8	$w_i$
C1 Korišćenje zemljišta	1	3	4	5	6	5	2	2	0.291
C2 Ekspozicija	1/3	1	2	1/2	2	2	1/3	1/3	0.082
C3 Nagib	1/4	1/2	1	2	1/2	1/2	1/4	1/5	0.050
C4 Visina	1/5	2	1/2	1	1/2	1/2	1/4	1/4	0.054
C5 Temperatura	1/6	1/2	2	2	1	1/2	1/3	1/4	0.059
C6 Padavine	1/5	1/2	2	2	2	1	1/3	1/3	0.073
C7 Rastojanje od puta	1/2	3	4	4	3	3	1	2	0.205
C8 Rastojanje od naselja	1/2	3	5	4	4	3	1/2	1	0.185

### 3.5. Agregacija konačne mape

U realizaciji konačne koristi se Ponderisana linearna kombinacija (Weighted Linear Combination, WLC) koja je integrisana u Spatial analyst tools, prema formuli:

$$S = \sum w_i x_i \quad (1)$$

Gde je S indeks osetljivosti od požara,  $w_i$  je normalizovana težina kriterijuma  $i$ , i  $x_i$  je fuzzy vrednost podobnosti rizika od požara kriterijuma  $i$ . Na taj način, množe se težine kriterijuma, dobijenih kao rezultat AHP-a, sa fuzzy skorom ćelija svakog kriterijuma  $i$  kao rezultat se generiše konačna mapa rizika od požara. Na bazi usvojenih kriterijuma, dobijena je konačna mapa, koja je predstavljena u istom vrednosnom opsegu kao i kriterijumi od 0 do 1. Veće vrednosti ćelija karakterišu prostor koji je više rizičan sa stanovišta pojave požara. U sledecem koraku izvršena je defazifikacija na 4 klasa pogodnosti pojave šumskog požara od slabog do veoma osetljivog područja (Slika 2). Tamnijom nijansom boje je obeležen prostor koji je najkritičniji za pojavu požara na prostoru Nacionalnog parka „Đerdap“.



**Slika 2.** Defazifikovana mapa rizika od požara

Analizom rezultata, ukupna površina najugroženijeg prostora (indeks osetljivosti 4) od šumskog požara u NP „Đerdap“ iznosi 115,1 km<sup>2</sup>, što je 7,7% ukupne teritorije. Uglavnom to su delovi koji se nalaze u blizini komunikacija i naselja, na strmim stranama i pretežno pokriveni četinarskom i mešovitom šumom (Tabela 4)

**Tabela 4.** Prostorna diferencijacija NP „Đerdap“ sa stanovišta rizika od požara

	Indeks osetljivosti	Area (km <sup>2</sup> )	%
1	Nizak	21,7	1,4
2	Umeren	596,9	39,7
3	Visok	770,4	51,2
4	Veoma visok	115,1	7,7

#### 4. ZAKLJUČAK

Učestalost pojave šumskih požara, uspeh gašenja, i veličina pričinjenih šteta u mnogom zavisi od detaljno proučenih i razrađenih preventivnih mera. Zaštita šuma od požara koncipirana je preko procene rizika od požara. U radu je prikazan GIS-AHP MCDA model s ciljem evaluacije zona rizika od pojave požara na području np „Đerdap“. Predviđanje šumskih požara izvršeno je na osnovu 8 kriterijuma. Standardizacija kriterijuma je primenom fuzzy logike. U postupku određivanja odnosa između kriterijuma i klastera korišćen je AHP višekriterijumski metod. Matrice poređenja zasnovane su na iskustvu eksperta, literaturi i dosadašnjoj praksi. Konačna mapa rizika je dobijena primenom WLC metode. Dobijeni rezultai nam ukazuju da 7,7% površine NP „Đerdap“ pripada zoni veoma visokog, 51,2% zoni visokog, 39,7% zoni umerenog i 1,4 % zoni niskog rizika od pojave požara. Ove informacije predstavljaju osnovu za prevenciju, upravljanje i minimizaciju štete izazvanih šumskim požarima. Primenjenim modelom proširuje se teorijski okvir znanja iz oblasti upravljanja rizicima. Postojeću metodologiju moguće je primeniti na područja sa sličnim geografskim karakteristikama. Takođe, prikazani model je moguće unaprediti sa novim i modifikovanim kriterijumima koji u dosadašnjim modelima nisu razmatrani, a koji su od značaja za ovu problematiku.

#### LITERATURA

- [1] JRC Technical Reports: Forest Fire in Europe, Middle East and North Africa, [http://forest.jrc.ec.europa.eu/media/cms\\_page\\_media/9/FireReport2012\\_Final\\_2pdf](http://forest.jrc.ec.europa.eu/media/cms_page_media/9/FireReport2012_Final_2pdf) (21.09.2015)
- [2] L. Gigović, H.R.Pourghasemi, S. Drobnjak, S.Bai: Testing a New Ensemble Model Based on SVM and Random Forest in Forest Fire Susceptibility Assessment and Its Mapping in Serbia's Tara National Park. *Forests* 2019, 10, 408. (2021).
- [3] E. Chuvieco, R. G. Congalton: Application of Remote Sensing and geographic Information Systems to Forest Fire Hazard Mapping. *Remote sensing of Environment*, (1989).
- [4] S. J. Pyne, P. L. Andrewa, R. D. Laven: *Introduction to Wildland Fire*. John Wileys and Sons Inc., 221-227 (1996).
- [5] S. Milanović, N. Marković, D. Pamučar, Lj. Gigović: Forest Fire Probability Mapping in Eastern Serbia: Logistic Regression versus Random Forest Method, *Forests*, 12(1), 5, (2021).
- [6] C. Vasilakos, K. Kalabokidis, J. Hatzopoulos, G. Kallos, Y. Matsinos: Integrating new methods and tools in fire danger rating. *International Journal of Wildland Fire*, 16 (3), 306 (2018).
- [7] European Environment Agency: Corina Land Cover. <https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012>
- [8] [http://www.hidmet.gov.rs/latin/meteorologija/klimatologija\\_godisnjaci.php](http://www.hidmet.gov.rs/latin/meteorologija/klimatologija_godisnjaci.php)
- [9] T.L. Saaty: *The Analytic Hierarchy Process*. McGraw-Hill, New York, 287 (1980).

## OPTIMIZACIJA PROFESIONALNIH MALIH BESPILOTNIH LETELICA ZA POTREBE GEODEZIJE I GIS-A

### OPTIMIZATION OF PROFESSIONAL SMALL UNMANNED AERIAL VEHICLES FOR THE NEEDS OF GEODESY AND GIS

RADOJE BANKOVIĆ<sup>1</sup>, BOBAN MILOJKOVIĆ<sup>2</sup>, DRAGOLJUB SEKULOVIĆ<sup>3</sup>, MILOŠ BASARIĆ<sup>1</sup>, IVAN GARIĆ<sup>1</sup>

<sup>1</sup> Vojnogeografski institut „General Stevan Bošković“, Beograd, Srbija, radojebankovic@yahoo.com, mbasarićbp@gmail.com, ivangarić@yahoo.com

<sup>2</sup>Kriminalističko-policijski univerzitet, Beograd, Srbija, boban.milojkovic@kpu.edu.rs

<sup>3</sup> Fakultet za poslovne studije i pravo, Univerzitet „Union – Nikola Tesla“, Beograd, Srbija, dragoljub.sekulovic@fpp.edu.rs

**Rezime:** Razvoj savremenih naučno-tehnoloških otkrića u aero industriji doprineo je ubrzanom povećanju masovnosti primene malih bespilotnih letelica u mnogim sferama ljudske delatnosti, od vojnih, preko komercijalnih, do inženjerijskih. S tim u vezi, u radu je pored pojmovnog razgraničenja, konceptualnih rešenja, opisa produkata snimanja, oblasti primene i načina upotredne profesionalnih malih bespilotnih letelica, data njihova optimizacija za potrebe geodezije i GIS-a primenom metode iskustvene generalizacije. Na osnovu taktičko-tehničkih performansi i iskustva u njihovom korišćenju, izrađeni su težinski kriterijumi i potkriterijumi optimizacije i izvršeno je rangiranje letelica koje su dostupne u okviru nacionalnog kapaciteta – privatnog sektora za geoinformacione tehnologije i geomatiku. Dobijeni rezultati ukazuju da je letelica eBee Geo SenseFly po „redosledu pogodnosti“ najprimerenija za potrebe geodezije i GIS-a u kojima se traži veća funkcionalnost opreme misije, zahtevne letne performasne, pouzdanost i bezbednost letenja u složenim uslovima, veća mobilnost tehničkih rešenja i svestrane tehničke podrške ovlašćenih distributera.

**Ključne reči:** Optimizacija, bespilotne letelice, geodezija, GIS.

**Abstract:** The development of modern scientific and technological discoveries in the aerospace industry has contributed to the rapid increase in the mass use of small unmanned aerial vehicles in many spheres of human activity, from military, through commercial, to engineering. In this regard, in addition to conceptual delimitation, conceptual solutions, descriptions of imaging products, areas of application and methods of use of professional small drones, their optimization for the needs of geodesy and GIS using the method of experiential generalization. Based on tactical-technical performance and experience in their use, weighting criteria and optimization sub-criteria were developed and the ranking of aircraft available within the national capacity - the private sector for geoinformation technologies and geomatics - was performed. The obtained results indicate that the eBee Geo SenseFly spacecraft is the most suitable for surveying and GIS in the „order of convenience“, which requires greater functionality of mission equipment, demanding flight performance, reliability and safety of flying in complex conditions, greater mobility of technical solutions and versatile technical support from authorized distributors.

**Keywords:** Optimization, unmanned aerial vehicle, geodesy, GIS.

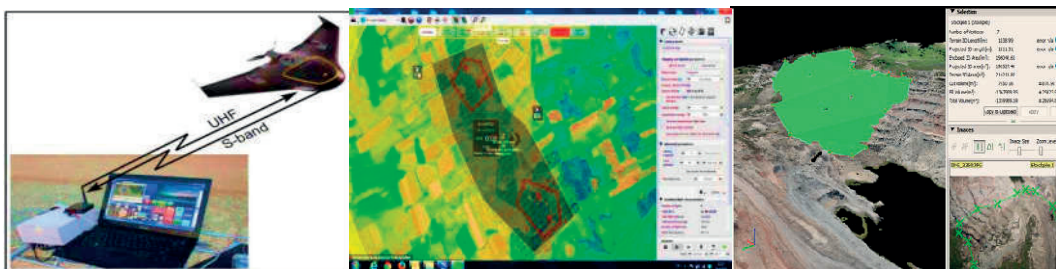
## 1. UVOD

Danas se u aktuelnoj naučnoj i tehničkoj praksi, za najrazličitije potrebe savremenog čoveka, koriste bespilotne letelice. Veći broj tih letelica namenjen je za potrebe sistema odbrane i bezbednosti, a nešto manji za potrebe privrede, nauke, sporta, zabave i dr. Na nastanak bespilotnih letelica uticale su povećane vojne potrebe sa početka 20. veka i sve donedavno bile su rezervisane isključivo za vojne primene u „neprijateljskom“ okruženju, u rejonima visokog rizika za letelice sa pilotskom posadom zbog prisustva snažnih artiljerijsko-raketnih jedinica, ali i za izviđanje, korekturu artiljerijske vatre, lasersko ozračivanje ciljeva radi napada vođenim projektilom i kao oružje izbora u borbi protiv ekstremista. Pri tome, letelice moraju ispuniti opravdati stroge zahteve kao što su raznovrsna ubojna sredstva i optika, autonomija leta,

minimalna elektromagnetska emisija tokom misije, otpornost na ometanje i nepovoljne meteo uslove i dr., što sve utiče na njihovu visoku cenu projektovanja i izrade [3].

## 2. PROFESIONALNE MALE BESPILOTNE LETELICE

Jedan deo bespilotnih letelica koristi se u režimu profesionalne civilne primene, gde se izdvajaju male bespilotne letelice (UAV - *Unmanned aerial vehicle*) sa posebnim performansama za inženjerske potrebe u oblasti geodezije i GIS-a. Takve letelice su okosnica bespilotnih aerofotogrametrijskih sistema – *Unmanned aerial photogrammetric systems (UAFS)* [1]. Radi se o sistemu zasnovanom na podsistemima koga čine profesionalna mala bespilotna letelica (PMBL) za autonomno aerofotogrametrijsko snimanje površi Zemlje pomoću visoko kvalitetnih, lakih i kalibriranih digitalnih senzora – kamera, zatim, kontrolni računar, radio veza za komunikaciju s letelicom, softvera za planiranje leta, rad na terenu i naknadnu obradu podataka u birou, bazna stanica za diferencijalno GNSS pozicioniranje u **RTK/PPK** režimu, oprema za nošenje i transport, inteligentni punjač baterije i ostali pribor (Slika 1).



Slika 1: Sistem profesionalne male bespilotne letelice za inženjerske primene (izvor: arhiva autora)

Najsavremenije profesionalne male bespilotne letelice koje imaju napredna tehnička rešenja sa ugrađenim multi frekventnim RTK GNSS visoke preciznosti, moćne letne performanse i opremu misije koju čine senzori za daljinsku detekciju (spektralne, multispektralne i termalne digitalne kamere sa visokom geoprostornom, radiometrijskom i spektralnom rezolucijom) predstavljaju veoma efektivan i efikasan način za masovno prikupljanje geoprostornih podataka za geodetski premer i GIS projekte, kao i za interaktivna 2D/3D merenja (koordinate, dužine, visine, visinske razlike, podužni i poprečni profili, površine i zapremine, ukop/iskop i drugo). Naknadnom obradom prikupljenih snimaka nastaje niz digitalnih geotopografskih materijala, kao što su 3D oblak tačaka, digitalni model površi terena, digitalni model terena, digitalni ortofoto, pojedinačne fotografije koji predstavljaju topografsku osnovu za najrazličitije inženjerske primene (Slika 2).



Slika 2: Produkti snimanja profesionalnom malom bespilotnom letelicom za potrebe geodezije i GIS-a (izvor: arhiva autora)

Osnovna prednost date geoinformacione tehnologije za daljinsku detekciju se ogleda se u obezbeđenju povećane produktivnost koja je nedostižna za terestičku fotogrametriju, uz ekonomičnost, rentabilnost i jednostavnost koju aerofotogrametrija ne može da postigne. Naime, savremene metode daljinskog sondiranja Zemljine površi stalno se usavršavaju. Različiti rezultati aerofotogrametrijskih i kosmičkih snimaka sve više se koriste u različitim oblastima ljudske delatnosti. Da bi se napravio georeferencirani ortofotoplan područja, više nije potrebno koristiti skupe metode dobijanja fotografija iz vazduha (satelitsko snimanje, manji avioni). To je postalo moguće zahvaljujući brzom razvoju industrije ultra-malih bespilotnih letelica. U jednom letu, takva letelica je sposobna da obuhvati od nekoliko desetina do nekoliko stotina hektara geoprostora sa velikim brojem snimaka, koji će se naknadno koristiti za

pravljenje georeferenciranog ortofotoplana područja [9]. Takođe, navedena tehnologija je nezamenljiva i kod snimanja teško prohodnog i ugroženog (opasnog) područja [4].

Zato se adekvatnom primenom PMBL optimalno i svrsishodno dolazi do raznovrsnog seta produkata koji se mogu koristiti za dokumentovanje zatečenog stanja i promena kroz vreme (infrastrukturnih objekata, gradilišta, klizišta, kopova i rudarskih basena, kamenoloma, odlagališta industrijskog i komunalnog otpada), zatim za idejno/detaljno projektovanje u oblasti elektroenergetskih, telekomunikacionih, putnih i železničkih infrastruktura, zatim u preciznoj poljoprivredi i šumarstvu, izradi i ažuriranju geoprostornih informacionih sistema, vršenju inspeksijskog nadzora i komunalne delatnosti, kao i mnogim srodnim inženjersko – tehničkim delatnostima (analiza poplavnih talasa, analiza šteta nakon prirodnih i tehničkih katastrofa, analiza stanja bezbednosti, ekološke analize itd.). PMBL najbolji je primer za vizuelizaciju promene[5].

Dosadašnja iskustva iz primene PMBL za geodeziju, GIS projekte i najrazličitije inženjerske namene, upućuju na izvesna ograničenja u vezi projektovane tačnosti i veličine pokrivenosti koju propisuje proizvođač i otkazima podsistema u toku leta. S tim u vezi, u toku snimanja, operator bi trebalo da konstantno upravlja dometom radio linka, meteorološkim uslovima i rastojanjem između plana leta i visokokvalitetne baze GNSS stanice koju bi, po mogućstvu, trebalo postaviti u sredini zone snimanja. Pri planiranju leta potrebno je proceniti vertikalnu rasčlanjenost reljefa i visinske razlike (naročite preko 500 m), zatim postojanje i visinu vegetacionog pokrivača i izgrađene objekte koji utiču na sigurnost i tačnost u radu zbog višestruke refleksije GNSS signala. Za centimetarsku tačnost potrebno je koristiti visokokvalitetnu letelicu sa opremom misije koja podržava kinematiku u realnom vremenu (RTK) i kinematiku na naknadnom obradnom signalu (PPK), odnosno, ugrađeni GNSS modul mora bar da obrađuje i L1 i L2 frekvencije. Za potvrdu potpune tačnosti dobijenih podataka potrebno je uspostaviti nekoliko oslonih tačaka visoke preciznosti koja uključuje korišćenje uobičajenih brendova Trimble ili Leica, a kod nepristupačnog i opasnog terena, potrebno je postaviti barem jednu oslonu tačku visoke preciznosti na ivici zone snimanja. Takođe, potrebno je izvršiti procenu lokacije PMBL u vreme svakog snimanja slike zasnovano na kinematici za naknadnu obradu (PPK) [6].

Međutim, pored teorijskog definisanja, klasifikacije, sagledavanja mogućnosti, ograničenja, smernica i mera predostrožnosti prilikom korišćenja PMBL, pred geostručnjake se često postavlja pitanje analize izazova dizajna hardvera i softvera PMBL sa različitim funkcionalnim mogućnostima u inženjerskim primenama, tj. dolazanje do optimalnog rešenja za njihovu nabavku. Upravo, rešenje mogućih dilema bio je cilj istraživanja u ovom radu.

### 3. MATERIJAL I METODE

Tržište malih profesionalnih bespilotnih letelica za geodeziju, GIS i druge inženjerske namene je u permanentnom usponu, ali nije razvijeno kao tržište komercijalno-civilnih i vojnih bespilotnih letelica. S tim u vezi u istraživanju je korišćen namerni uzorak koga su činile 3 najzastupljenija modela PMBL u nacionalnom okviru i to: WingtraOne GEN II (u varijanti leteće krilo – avion), Sensefly eBee Geo (u varijanti leteće krilo – avion) i DJI Phantom 4 RTK (u varijanti multikopter).

U postupku dolaženja do optimalnog modela PMBL za potrebe geodezije i GIS-a u nacionalnom okviru najtačnije je korišćene metode analitičkih hijerarhijskih procesa, kao vrsta egzaktno naučne metode operacionih istraživanja koja se sastoji iz četiri faze: 1) strukturiranje problema; 2) prikupljanje podataka; 3) ocenjivanje relativnih težina, i 4) određivanje rešenja problema [7], [8]. Međutim, da bi se izvršila optimizacija modela PMBL i došlo do „redosleda pogodnosti“ PMBL za potrebe geodezije i GIS-a kada postoje izvesni ograničavajući kriterijumi (npr. mali uzorak zbog malog broja dostupnih modela na nacionalnom tržištu, softverska ograničenja za ekspertizu i drugo), mogu se primeniti i prigodne metode optimizacije kao što je metod iskustvene generalizacije [2].




Naime, za procenu relativnih težina kriterijuma optimizacije, kao i za određivanje relativnih težina pojedinih alternativa po konkretnom kriterijumu, korišćeno je iskustvo autora i njihovih kolega u primeni PMBL. S tim u vezi, utvrđeni su i rangirani po stepenu težine sledeći kriterijumi i potkriterijumi na osnovu kojih će se taj redosled formirati: 1) Univerzalnost opreme za snimanje (bolje su rangirane PMBL koje mogu da vrši snimanje sa opremom misije (kamere) koja podržava RTK i PPK, koja se može opciono nadograditi i kojom se dobijaju geoprostorni podaci čijom se naknadnom obradom nastaju digitalni ortofoto snimci, termalni i multispektralni snimci); 2) Ekonomičnost (bolje su rangirane PMBL koje imaju nižu cenu koštanja i manji broj poslužioca, što je u skladu sa restriktivnim finansijskim sredstvima kojima se raspolaže); 3) Klimorobusnost opreme misije (bolje su rangirane PMBL koje mogu da lete pri niskim i visokim temperaturama, i po umerenoj kiši i snegu); 4) Letne mogućnosti (bolje su rangirane PMBL koje imaju bolje letne karakteristike koje se odnose na duže vreme, veću brzinu i visinu leta, mogućnost izbegavanja

prepreka, kao i na veću dužinu operativnog dometa koja je pokrivena radio linkom i siguran povratak u slučaju nedovoljnog napajanja, što je potrebno da bi se dobili geoprostorni podaci metrički tačni i u realnom vremenu sa što veće površine snimanja); 5) Mobilnost tehničkog rešenja (bolje su rangirane PMBL koje mogu da nose težu opremu misije, koje su lakše i manjih dimenzija, što ih čini jednostavnijim za transport i montiranje za snimanje i pakovanje), i 6) Organizacioni problemi (bolje su rangirane PMBL za koje je moguća obuka, tehnička podrška i servis posle određenog broja časova leta, kao i opcija iznajmljivanja opreme i usluga planiranja leta i naknadne obrade od strane ovlašćenog distributera).

#### 4. REZULTATI I DISKUSIJA

Na osnovu iskustvene generalizacije i izvršene analize performansi iz tehničkih specifikacija za optimalne uslove leta 3 najzastupljenija modela PMBL u nacionalnom okviru, modelovana je *tabela 1* sa prikazom iskustvenih težinskih kriterijuma i potkriterijuma kako bi se došlo do „redosleda pogodnosti“ PMBL za potrebe geodezije i GIS-a, a samim tim i do preporuke za njihovu moguću nabavku za te ali i za druge inženjerske namene.

**Tabela 1:** Ispunjenost težinskih kriterijuma i potkriterijuma za optimizaciju profesionalnih malih bespilotnih letelica za geodeziju i GIS(\* - prvi nivo ipunjenosti)

TIP/MODEL PROFESIONALNE MALE BESPILOTNE LETELICE RTK & PPK				
		Avion/WingtraOne GEN II	Avion/Sensefly eBee Geo	Multikopter/DJI Phantom 4 RTK
Iskustveni težinski kriterijumi optimizacije	ISPUNJENOST POTKRITERIJUMA ZA OPTIMIZACIJU			
Univerzalnost opreme za snimanje	3D snimanje terena	Da	Da	Da
	3D planiranje leta	Da	Da	Da
	Multispektralno snimanje	Da	Da	Da
	Termovizija	Da	Da	Da
	Multifrekventni RTK GNSS visoke preciznosti	Da	Da	Da
Opciona oprema	Pametni punjač baterija	SenseFly Corridor adapter, Radio tragač u slučaju gubljenja* Pametni punjač baterija	Da, D-RTK 2 GNSS Mobilna stanica, Pametni punjač baterija	
Ekonomičnost	Cena u Evrima - bez PDV	22.000 sa RGB kamerom od 20 MP; 29.000 sa kamerom od 42 MP	16.500 sa RGB kamerom od 20 MP	5.500 sa RGB kamerom od 20 MP*
	Broj poslužioaca (optimalno 1 za teren, 1 za biro, opciono 1)	Da, 1+1/1	Da, 1+1/1	Da, 1+1/1



<b>Klimorobusnost</b>	Maksimalna otpornost na vetar	Do 12 m/s	Do 12,8 m/s*	Do 10 m/s
	Temperaturni opseg korišćenja	-10 do +40 °C	-15° do +40°C*	0°C do +40°C
	Vlažnost (kiša, magla, sneg)	Ne	Otpornost na slabu kišu*	Ne
<b>Letne mogućnosti</b>	Brzina leta	57,6 km/h	40-110 km/h*	58 km/h
	Vreme leta	59 min*	59 min*	30 min
	Tip poletanja/sletanja	Vertikalno/vertikalno	Iz ruke/automatsko linearno sletanje	Vertikalno/vertikalno
	Maksimalna pokrivenost jednim letom	310 ha*	220 ha	100 ha
	Izbegavanje prepreka	Da	Da -LiDAR senzor za uočavanje prepreka (do 120 m)*	Da - Infracrveno opažanje prepreka do 30 m
	Rezolucija na tlu (GSD)	1 cm*	1,5 cm	3 cm
	Daljina leta pokrivena radio linkom	3 – 10 km*	3 – 8 km	3 -7 km
	Automatski povratak kada je baterija prazna	Da	Da	Da
<b>Mobilnost tehničkog rešenja</b>	Težina letelice sa opremom misije	4,5 kg	1,3 kg*	1,391 kg
	Težina letelice pri transportu	18,6 kg (putna torba)	5,9 kg (ranac)*	8,19 kg (putna torba)
	Vreme pripreme letelice za snimanje/pakovanje	4 min	3 min*	4 min
<b>Organizacioni problemi</b>	Mogućnost obuke za upravljanje letelicom	Da	Da	Da
	Tehnička podrška i servis na 100 sati leta	Da	Da	Da
	Dostupnost rezervnih delova	Da	Da	Da
	Opcija iznajmljivanja, usluga planiranja leta i naknadne obrade	Ne	Da*	Ne

Na osnovu procene relativnih težina kriterijuma optimizacije, kao i na osnovu određenih relativnih težina pojedinih alternativa po konkretnom kriterijumu izvršeno je sledeće rangiranje:

1) *Univerzalnost opreme za snimanje*. Sve tri analizirane PMBL poseduju funkcionalna hardversko-softverska rešenja, tj.mogu da vrši 3D planiranje i snimanje sa opremom misije (kamerama) koja podržava RTK i PPK metod pozicioniranja, odnosno, mogu svojim softverskim rešenjima da naknado obrade prikupljene geoprostorne podatke čime nastaju digitalni ortofoto snimci, termalni i multispektralni snimci projektovane tačnosti za geodeziju, GIS projekte i druge inženjerske namene. Što se tiče potkriterijuma opcione nadogradnje, sve tri letelice imaju pametne punjače baterije, s tim da letelica eBee Geo SenseFly ima mogućnost namenske dogradnje opreme za snimanje koridora - Corridor adapter i Radio tragač u slučaju gubljenja, čime se izdvaja od druga dva modela i u kriterijumu univerzalnosti opreme misije zauzela bi prvo mesto.

2) *Ekonomičnost*. Sve tri letelice može opsluživati minimalan broj poslužioca. Letelica WingtraOne GEN II ima najveću cenu koštanja, ali ima i najmoćniju kameru od 42 MP, dok letelica DJI Phantom 4 RTK Multikopter ima najnižu cenu koštanja sa kamerom od 20 MP koja zadovoljava projektovanu tačnost, pa bi po kriterijumu ekonomičnosti zauzela prvo mesto.

3) *Klimorobusnost opreme misije*. Letelica eBee Geo SenseFly se izdvaja po klimorobusnim performansama, tj.može da leti po najvećem temperaturnom opsegu (pri niskim i visokim temperaturama), i jedina može da leti pri umerenoj kiši, pa bi po kriterijumu klimorobusnosti zauzela prvo mesto.

4) *Letne mogućnosti*. Sve tri letelice imaju izuzetne letne karakteristike, i ujednačene su po pitanju tehničkih rešenja o načinu poletanja i sletanja, izbegavanju prepreka i autonomnog povratka kada je baterija prazna. Ipak, po podkriterijumu maksimalne pokrivenosti jednim letom, rezolucije na tlu (GSD) i daljine leta

koja je pokrivena radio linkom, izdvaja se letelica WingtraOne GEN II, pa bi po kriterijumu letnih performansi zauzela prvo mesto.

5) *Mobilnost tehničkog rešenja*. Sve tri letelice imaju izuzetna tehnička rešenja po pitanju mobilnosti tehničkih rešenja. Ipak, letelica eBee Geo SenseFly se izdvaja po ovom kriterijumu jer ima najmanju težinu sa opremom misije, najlakša je za transport i ima najkraće vreme pripreme letelice za snimanje i pakovanje, pa bi po kriterijumu mobilnosti tehničkog rešenja zauzela prvo mesto.

6) *Organizacioni problemi*. Sve tri letelice prodaju ovlašćeni distributeri koji u cenu letelice uračunavaju cenu obuke za upravljanje letelicom, imaju obezbeđenu tehničku podršku i servis posle određenog broja časova leta, s tim da samo ovlašćeni distributer letelice eBee Geo SenseFly, nudi mogućnost iznajmljivanja kompletne opreme za aerofotogrametrijsko snimanje (letelica, GNSS baza/rover, usluga planiranja leta i naknadne obrade), pa bi po kriterijumu organizacionih problema zauzela prvo mesto.

Na osnovu iznetog, najveći broj težinskih kriterijuma i potkriterijuma zadovoljava letelica eBee Geo SenseFly, tj. po „redosledu pogodnosti“ bila bi najprimerenija za potrebe geodezije i GIS-a, ali i za druge inženjerijske potrebe, i ako letelica WingtraOne GEN II ima veću tačnost i veću pokrivenost, a letelica DJI Phantom 4 RTK ima manju cenu koštanja. Sagledavajući rezultate optimizacije malih bespilotnih letelica za potrebe policije [3], letelica kompanije SenseFly je takođe bile prva po „redosledu pogodnosti“.

## 5. ZAKLJUČAK

Bespilotne letelice svakim danom postaju sve veća tehnička podrška odlučivanja u najrazličitijim sferama ljudske delatnosti, počev od stvaralačke, pa do destruktivne. Profesionalne male bespilotne letelice za prikupljanje geoprostornih podataka sa opremu misije koju čine senzori za daljinsku detekciju (aerofotogrametrijske, termalne i multispektralne kamere) brzo su se razvijali u poslednjih desetak godina.

Za efektivnu i efikasnu primenu PMBL za potrebe geodezije i GIS-a, kao i drugih inženjerijskih namena, potrebno je detaljno poznavati njihove mogućnosti opreme misije, cenu koštanja, robusnost, letne i druge performanse, tj. pre same nabavke, potrebno je izvršiti njihovu optimizaciju po „redosledu pogodnosti“ korišćenjem metoda operacionih istraživanja kao što su višekriterijumska optimizacija, ili kada okolnosti to nalažu, korišćenjem prigodnih metoda kao što je metoda iskustvene generalizacije.

Dobijeni rezultati u radu primenom metoda iskustvene generalizacije ukazuju da je po analiziranim težinskim kriterijumima i potkriterijumima, letelica **eBee Geo SenseFly** po „redosledu pogodnosti“ na prvom mestu jer ima 4 od 6 razmatranih težinskih kriterijuma, Time je ona najprimerenija za potrebe geodezije i GIS-a u kojima se traži veća funkcionalnost opreme misije, zahtevne letne performasne, robusnost i bezbednost letenja u složenim uslovima, veća mobilnost tehničkih rešenja i svestrane tehničke podrške ovlašćenih distributera. Takođe, potrebno je naglasiti da letelica **eBee Geo SenseFly** ima zadovoljavajuću horizontalnu i vertikalnu tačnost, zavidnu veličinu pokrivenosti, ali i znatno veću cenu u odnosu na najjeftinija rešenja koja imaju izvesna ograničenja u primeni, i znatno manju cenu u odnosu na najsuperiornije letelice u tretiranoj klasi.

## LITERATURA

- [1] Adžemović, Lj., & Milenović, V. (2014). Bespilotni aerofotogrametrijski sistemi, *Geodetska služba*, 43 (117), str. 102-108.
- [2] Milojević, S., (2010). Optimizacija posebnih oblika nastave na Kriminalističko-policijskoj akademiji, *Bezbednost*, (52/3), str. 172-237.
- [3] Milojković, B., (2015). Optimizacija modela malih bespilotnih letelica za potrebe policije, *Bezbednost*, 57(3), str. 5-27.
- [4] Casado, R. M., Irvine, T., Johnson, S., Palma, M., & Leinster, P. (2018). The Use of Unmanned Aerial Vehicles to Estimate Direct Tangible Losses to Residential Properties from Flood Events: A Case Study of Cockermonth Following the Desmond Storm. *Remote Sensing*, 10(10), 1548.
- [5] Singhal, G., Bansod, B., & Mathew, L. (2018). Unmanned Aerial Vehicle classification, Applications and challenges: A Review. *Preprint, November*, p.1-19.
- [6] Taddia, Y., Stecchi, F., & Pellegrinelli, A. (2019). USING DJI PHANTOM 4 RTK DRONE FOR TOPOGRAPHIC MAPPING OF COASTAL AREAS, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W13, p. 625-630.

- [7] Tuba, Z., Vidnyánszky, Z., Bottyán, Z., Wantuch, F.,&Hadobács, K. (2013). Application of Analytic Hierarchy Process in fuzzy logic-based meteorological support system of unmanned aerial vehicles, *Academic and Applied Research in Public Management Science*, 12 (2), стр. 221-228.
- [8] Vasiljević, D.,&Vasiljević, J. (2015). Optimization of the elements of unmanned aircraft systems, In Prociding, *International scientific conference of IT and Business Related Research – Synthesis 2015*, Singidunum University in Belgrade, p. 213-217.
- [9] Захлебин, С. А. (2021). Методика построения ортофотопланов местности с помощью беспилотного квадрокоптер, *Журнал „Доклады ТУСУР”/Архив /Выпуск журнала, № 3, Том 24, стр.44-49.*



## RAD SA PROSTORNIM PODACIMA

### WORK WITH SPATIAL DATA

SASA BAKRAČ<sup>1,2</sup>, BORIS VAKANJAC<sup>1</sup>, DEJAN ĐORĐEVIĆ<sup>1,2</sup>, IVAN POTIĆ<sup>1</sup>, SINIŠA DROBNJAK<sup>1,2</sup>,

<sup>1</sup> Vojnogeografski institut-"General Stevan Bošković", Beograd, [sasa.bakrac@vs.rs](mailto:sasa.bakrac@vs.rs); [borivac@gmail.com](mailto:borivac@gmail.com); [dejandjordjevic.vgi@gmail.com](mailto:dejandjordjevic.vgi@gmail.com); [ipotic@gmail.com](mailto:ipotic@gmail.com); [sdrobnjak81@gmail.com](mailto:sdrobnjak81@gmail.com)

<sup>2</sup> Univerzitet odbrane, Vojna akademija, Beograd

**Rezime:** U mnogim naučnim disciplinama se koristi terenski rad kao način prikupljanja podataka. Neke od njih su geografija, geodezija, geologija, poljoprivreda, biologija, meteorologija i druge. Često mladi stručnjaci, tačnije tek završeni studenti, iako su imali predmete gde se učilo kartiranje ili neki sličan predmet nemaju dovoljno praktičnog iskustva u izvođenju procedura vezanih za izradu karte ili "rad sa prostornim podacima". Slično može da se dogodi i iskusnim stručnjacima iz nekih oblasti, ali koji se nisu bavili terenskim radom. Termin "rad sa prostornim podacima" smo izabrali ili uveli iz razloga što se metodologija nekadašnjeg kartiranja dosta promenila. Danas skoro da ne postoji prostor na planeti Zemlji koji na neki način nije posećen i u različitoj meri proučen. Generalno proces podrazumeva tri faze. Prva bi bila prikupljanje postojećih podataka vezanih za određen prostor i problematiku, sa formiranjem preliminarne baze podataka od dostupnog postojećeg materijala. Druga faza je izlazak na teren gde bi se "in situ" vršilo proučavanje objekata od interesa, a koji se nalaze na nekom prostoru, pri čemu se beleže koordinate observacionih tačaka, rade dokumentacione fotografije i vrši opis observacija (ili observacionih tačaka) kao i eventualno uzorkovanje različitog materijala. Na kraju u trećoj fazi bi se na osnovu postojećih podataka i rezultata terenskog rada formirala finalna baza podataka sa tekstualnim opisom i generisanom kartom (kartama) područja i objekata od interesa. U ovom radu su dati glavni principi i metode pomenutih aktivnosti.

**Ključne reči:** prostorni podaci, baze podataka, obrada podataka

**Abstract:** In many scientific disciplines, fieldwork is used as a way of collecting data. Some of them are geography, geodesy, geology, agriculture, meteorology and others. Often young professionals, more precisely newly graduated students, although they have had subjects where mapping was taught, or some similar subject, do not have enough practical experience in performing procedures related to mapping or "working with spatial data". A similar thing can happen with experienced experts in some fields, but who have not dealt with fieldwork. We chose to introduce the term "work with spatial data" because the methodology of the then mapping has changed a lot. Today, there is almost no space on Planet Earth that has not been visited and studied in various ways. Generally, the process involves three phases. The first would be the collection of existing data related to a particular area and issues, with the formation of a preliminary data base of available existing material. The second phase is going out to the field where "in situ" studies of objects of interest would be performed, which are located in a certain area, recording the coordinates of observation points, making documentary photographs, and describing observations (or observation points) and possibly sampling of a different material. Finally in the third phase, based on the existing data and results of fieldwork, a final database would be formed with a textual description and map (maps) of areas and objects of interest will be generated. This paper presents the main principles and methods of the mentioned activities.

**Keywords:** spatial, data, databases, data processing.

## 1. UVOD

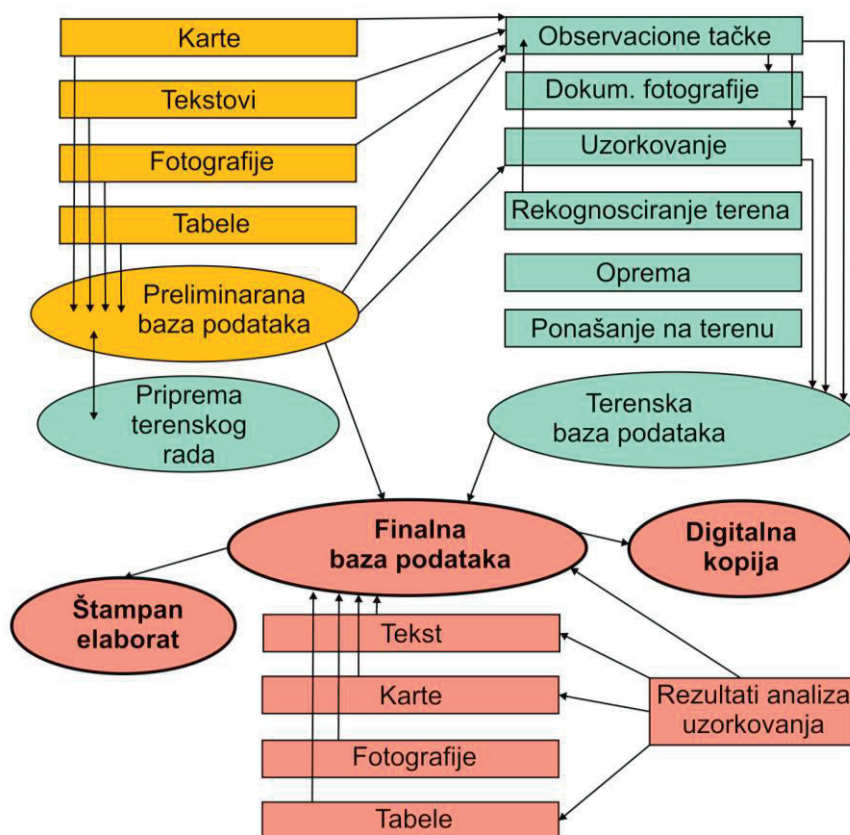
Terenski podatak je nezamenjiv izvor velikog broja informacija u raznim naukama koje se bave prostorom na planeti Zemlji, a od skorije i prostorom Zemljinog satelita Meseca i planete Mars, gde su sa Meseca donešeni uzorci [1], a na Marsu se tlo ispituje tzv. "roverima" koji su opremljeni malim automatskim laboratorijama za ispitivanje, stenskog, hemijskog sastava tla i drugo [2].

Terenski podatak je određen nizom informacija, od kojih neke imaju, a neke ne moraju nužno imati smisla. Takođe postoji veliki broj nauka koje su vezane za prostor što je navedeno u rezimeu.

Na teren se ne može ići "ad hoc", npr. okvirno se odredi prostor od interesa, pa se ode tamo, uzimaju uzorci, rade dokumentacione fotografije i taj materijal se na neki način organizuje. Iz iskustva znamo da ako se fotografije odmah posle terenskog dana ne stave u već pred pripremljenu bazu podataka tj. fotografija posle recimo tri dana neažurnog pristupa dokumentacionim fotografijama nastaje niz problema koje nije lako rešiti (kad je fotografija urađena, koji objekat itd.). Tako je sa skoro svim aktivnostima vezanim za terenski rad.

U ovom radu ćemo dati neke okvirne smernice koje bi bile od velike pomoći posebno mladim stručnjacima koji treba da donesu i odrade terenske/prostorne podatke. Nekad i neko bi ove postupke nazvao kartiranjem. Može se reći da je to i tačan i ne tačan termin. Tačan je iz razloga što rezultat pomenutog rada jeste (i mora da bude) karta, međutim tokom vremena i sa upotrebom komputera, tačnije GIS programa situacija se menja. Prostorni podatak koji je bio ili je prikazan na karti danas može da nosi niz atributa što ranije nije bilo moguće. Tačno je da su tekstualni tumači pratili većinu karata [3], ali u njima nije bilo moguće objasniti i prikazati svaku tačku, liniju, poliliniiju i poligon kao što je to danas. Takođe treba napomenuti da je moguće uneti u sistem veliki broj atributa vezanih za observacionu tačku ili neki objekat, što je prednost GIS programa i računara, ali ne moraju svi "mogući" atributi imati u konkretnom slučaju smisla.

Rad sa prostornim podacima ima generalno tri faze koje ćemo zbog ograničenog broja strana pokušati da koliko je to moguće objasnimo. Faze su 1. Faza - Kabinetski pred terenski rad; 2. Faza - Terenski rad; 3. Faza - Formiranje finalne prostorne baze podataka i pisanje izveštaja (Slika 1).



Slika 1. Dijagram rada sa prostornim podacima

## 2. I FAZA KABINETSKI PRED TERENSKI RAD, FORMIRANJE PRELIMINARNE BAZE PODATAKA

Kao što je gore napisano planeta Zemlja je posećena i opisana i skoro da ne postoji deo, a da nije na neki način ispitan. Čak i "Marijana trenč" na dubini od oko 10500 m, u okeanu, su u posebnoj vrsti podmornice posetili ljudi 1960 godine [4]. U ovom radu se nećemo baviti okeanima i morima iako su i ovi objekti dobro obrađeni i zanimljivi. Opisaćemo rad i procedure koje se rade na kopnu. Drugim rečima za svaki deo površine kopna postoje podaci i neka vrsta karte kao i fotografije sa tla, aviona/helikoptera ili satelita. Prva faza - Kabinetski pred terenski rad, bi se prvenstveno sastojala od sledećih aktivnosti:

a.) Prikupljanje postojećeg materijala za prostor od interesa. Potrebno je u institucijama koje se bave određenom problematikom npr. geologijom, biologijom, geodezijom meteorologijom, poljoprivredom..., ili arhivima, potražiti postojeće materijale. To mogu biti karte, pisani izveštaji, fotografije, knjige, elaborati i sl.

b.) Obrada prikupljenog materijala. Ovo podrazumeva pregled i selekciju materijala na: karte, skice, tekstualne zapise, fotografije i tabele. Nekada to mogu biti i rukom pisani dnevници koji mogu imati veliku vrednost u smislu observacija autora u određenim oblastima.

- Istorijske (stare) karte je potrebno geo-referencirati. Uglavnom će biti skenirane u JPEG ili TIFF formatu. Nekada su karte u lošem stanju pa ih je, koliko je moguće, potrebno obraditi u nekom od programa za obradu slika. Treba imati u vidu da neke, veće, skenirane karte u rezoluciji 200-300 DPI mogu zauzimati veliku memoriju [5], pa i o tome treba voditi računa. Ako je moguće, tj. ako se mogu prepoznati referentni objekti bilo bi dobro geo-referencirati i skice. Sve geo-referencirane i ne referencirane karte stavljamo u posebne foldere, a njih u jedan folder sa nazivom objekta koje ispituje npr. "Zlatibor\_bolesti\_biljaka\_istorijske\_karte".

- Tekstovi mogu biti u različitim programima (Word, PDF) ili kao slike (JPEG). Nekada je to knjiga ili radna sveska prethodnika koji je obrađivao sličnu problematiku pre možda 70 godina (obično se ovaj materijal skenira). Koliko je moguće pokušaćemo da uniformišemo podatke u nekom od programa za pisanje npr. Word\_u. Naravno to ne znači prepisivanje celih izveštaja, nekada je to nemoguće, a i često nema smisla. Međutim sav dostupan pisani materijal treba pročitati i složiti po tematici, autoru ili formatu (skenirano, Word itd.).

- Sa starim fotografijama može biti sledećih problema: malog su formata pa ih u nekom od programa za crtanje treba uveličati (ako je moguće) i izoštriti. Mogu biti crno-bele ili sepija monohromatske, ne treba ih naknadno bojiti i kao takve - monohromatske mogu dati važne informacije. Fotografije takođe stavljamo u foldere npr. "Zlatibor\_bolesti\_biljaka\_istorijske\_fotografije", takođe u ovom folderu se fotografije mogu pamtiti po autorima, godinama ili ako ne znamo, onda "crno\_bele", "sepija", "u\_boji", "Do\_II\_svetskog\_rata", "pedesete" ili "sedamdesete" i sl.

- U tabelama su često date različite vrednosti (hemijskih analiza), koordinate ili neko drugo objašnjenje pozicije objekata odakle su analize uzete (npr. TE\_Kolubara\_deponija\_pepela\_kaseta\_A). Problem je što, institucije ili biblioteke, tabele imaju u štampanoj formi dok je za početak najlakše raditi u Excel\_u. Često se neke tabele mogu preuzeti i sa interneta, date kao časopis ili naučni rad. Problem je što nama treba "živa" Excel tabela koju možemo da koristimo u nekom od GIS programa [6] i u tom slučaju treba da preradimo tabelu. Takođe kao prilog nekim starim kartama su date tabele, a tačke uzorkovanja su nacrtane na karti, tada je potrebno prebaciti pozicije "starih" tačaka koje su verovatno u Gauss-Kruger\_u u UTM, ako se radi o prostorima bivše SFRJ.

- Priprema terenskog rada podrazumeva observaciju dostupnih karata i formiranje mogućih ruta i pozicija observacionih tačaka. Pomenuto naravno nije apsolutno, po izlasku na teren situacija može biti manje ili više različita od plana, o čemu se donosi odluka na samom terenu. Istorijske podatke je potrebno proučiti i uporediti sa programskim zadatkom. Ovim postupcima se štedi vreme i novac jer nekad nije potrebno ponavljati seriju ispitivanja nego je samo proveriti.

## 3. II FAZA TERENSKI RAD

Pre izlaska na teren potrebno je pripremiti opremu, pozicije observacionih tačaka kao i putanje koje vode do njih. Posle prve pripremne, opisane faze sledi izlazak na teren. Generalno je potrebno: vozilo, GPS, saradnik/ca, kompas, sveska i olovka, marker, fotoaparatus, lupa, ruksak, topografska karta, prikladna odeća, rukavice, alatka za pristup rizičnim mestima, pribor za uzimanje sitnih uzoraka, multi funkcionalni nožić ili slično, plastične kese i kartonske pločice za uzorke, malo joda, zavoj i flaster. Takođe treba znati šta je observaciona tačka.

▪ Observacione tačke su deo prostora koji se detaljno opisuje i zbog čega se u suštini ide na teren. Tačka je geometrijski pojam i opisuje se kao presek dve prave, u smislu terenskog rada situacija je drugačija. Mesto observacije nije nužno tačka, može biti linija ili poligon. U smislu poligona to može biti deo gde je u šumi iz nekog razloga drveće obolelo, na nekom prostoru različitih dimenzija, ili gde su polutanti doveli do izumiranja živog sveta. Prvenstveno observaciona tačka je određena koordinatama i nadmorskom visinom. Koji koordinatni sistem će se koristiti zavisi od slučaja do slučaja. Danas na prostoru Republike Srbije se uglavnom koristi UTM34 N (T). Može postojati problem sa nadmorskom visinom jer ručni GPS prijemnici u tom smislu nisu apsolutno tačni, ali se taj problem može kasnije rešiti u kabinetu. U terenski dnevnik se upisuju podaci koji se dobijaju posmatranjem objekata od interesa na observacionoj tački. Prva stavka treba da bude ime projekta, slede: ime obrađivača, datum, broj observacione tačke koji bi trebao da ima veze sa datumom, opis tj. observacija - šta je na tom mestu viđeno (tip zemljišta, biljke, stene, stanje šume, tačnije ono zbog čega se terenske aktivnosti izvode u konkretnom slučaju). Kod davanja opisa i fotografisanja treba imati mere, sa većim iskustvom istraživač više vidi i nekad se potroši puno vremena na nekoj observacionoj tački. Ponekad je to neophodno, ali opis do 20 rečenica i do 5 fotografija je uglavnom dovoljno da se prostor i problematika od interesa na observacionoj tački kvalitetno prikažu. Retko, ali se događa da se na nekoj tački ili poligonu potroši ceo terenski dan. Potrebno je po povratku sa terena istog dana prekucati tekst iz terenskog dnevnika i staviti u za to pripremljen folder, a isto treba uraditi i sa fotografijama kao i sa koordinatama koje bi bilo dobro prekucati u npr. Excel\_u. Folderi određene observacione tačke sa tekstom i fotografijama treba da imaju isto ime npr. "Ljubovija\_Caparica\_reka\_deponija\_2". Izuzetno rade se i skice kada je potrebno naglasiti neku pojavu i kasnije u kabinetu u nekom od programa za crtanje prikazati je na karti ili fotografiji. Jedan od razloga ažuriranja podataka terenskog rada na dnevnom nivou je što je moguće da se oprema, fotoaparati ili terenski dnevnik mogu izgubiti ili oštetiti (npr. mogu upasti u potok).

Na mestu observacione tačke se rade i dokumentacione fotografije, najmanje dve - uglavnom 4 do 5 i bilo bi poželjno dati neku informaciju o položaju - smeru u kom su fotografije urađene, bilo bi najbolje da se da azimut, ali i termini kao što su "u pravcu severozapada" i slično mogu biti zadovoljavajući. Obavezno treba staviti u terenski dnevnik broj fotografije ako postoji na uređaju kojim se slika. Često je potrebno fotografisati objekte sa nekim razmernikom. To može da bude metalna novčanica poznatih dimenzija ili mali lenjir. Razmernik se ne stavlja u centar fotografije nego sa strane. Treba imati u vidu da je moguće uraditi i tzv. "makro" fotografije različitih objekata - insekata, obolelog lišća, svežih preloma stena što znatno olakšava kasniji rad pri analizi i pisanju izveštaja u kabinetu.

▪ Uzorkovanje se vrši u zavisnosti šta se radi u konkretnom projektu. Geolozi će uzorkovati komade stena iz razloga proučavanja tipova stena i eventualnih promena na njima. Geohemičari i ekolozi mogu uzorkovati zemljište iz razloga ispitivanja zagađenja, a agronomi zbog ispitivanja kvaliteta zemljišta. Biolozi će uzorkovati živi svet iz različitih razloga. Kao i za observacionu tačku i za uzorkovanje postoje pravila. Uglavnom je potreban alat koji zavisi od struke do struke kojim se uzorci od interesa uzimaju (od ašova, plastičnih ili čeličnih kašika, do skalpela, pinceta i drugog), kao i oprema u kojoj će se uzorci transportovati. Najčešće se uzorci sakupljaju i transportuju u čvrstim plastičnim kesama različitih veličina. Ovo naravno ne znači da se uzorci vode tako uzimaju. Analiza uzoraka vode tzv. "Obima V", podrazumeva više staklenih i plastičnih sudova gde se u nekim nalaze hemikalije koje stabilizuju određene procese koji se mogu dogoditi u vodi od trenutka uzorkovanja do ispitivanja u laboratoriji. Takođe i za uzorkovanje npr. živih insekata postoji posebna transportna oprema i procedure. Uzorke i mesto uzorkovanja je potrebno opisati i fotografisati, takođe ako mesto uzorkovanja nije u okviru observacione tačke treba dati i koordinate. Često se kod uzorkovanja rastresitog materijala koriste kartonske pločice koje mogu biti plastificirane. Na njima se upisuje niz podataka koji se upisuju i u terenski dnevnik. Takođe često se koriste flasteri različitih širina koji se lepe na kese sa uzorcima i na kojima je moguće hemijskom olovkom ili markerom upisati broj uzoraka i druge osnovne podatke.

▪ Rekognosciranje terena se preporučuje kada se prvi put dolazi na neki nepoznat prostor. Trase lokalnih puteva mogu biti promenjene, prelazi preko manjih vodotoka uništeni vremenom, nekorišćenjem ili bujicom. Iz pomenutih i niza drugih razloga potrebno je proći teren, videti kuda je moguće, a kuda nije, proći i izvršiti izmene planiranih ruta i observacionih tačaka u skladu sa situacijom.

▪ Oprema, ovde bi pojasnili nekoliko stvari. Kao vozilo preporučili bi Ladu Nivu iz više razloga. Može se nabaviti polovna u dobrom stanju i relativno je jeftina u odnosu na druga znatno skuplja terenska vozila. Takođe moguće ju je relativno lako, u slučaju da dođe do kvara, ako ne popraviti, onda osposobiti da se može stići do prve automehaničarske radionice gde bi se kvar otklonio. Današnji mobilni telefoni imaju ugrađene GPS prijemnike, kao i dobre multi-funkcionalne kamere tako da jedan uređaj može da zameni posebni GPS i fotoaparati. Topografska karta je neophodna iz razloga orijentacije na terenu. Nije nužno da to bude određeni list topografske karte 1:25000 ili 1:50000, danas na internetu postoji više dostupnih



topografskih karata različitih proizvođača [7] koje se mogu uveličati do npr. 1:10000 ili 1:5000, često su već geo-referencirane, mogu se odštampati ili nositi u elektronskom obliku u tabletu, takođe postoji i veliki broj satelitskih snimaka koji mogu biti od velike pomoći za orijentaciju. Kompas je neophodan da bi imali uvid u poziciju severa (time i drugih strana sveta) i orijentisali naš položaj na terenu i kartu (treba imati u vidu da položaj magnetnog severa delimično varira, ali ipak ne toliko da kompas nije potreban). Rukavice i alatka (čekić, cepin, manja šipka i sl.) za pristup rizičnim pozicijama su neophodni jer nekad treba uzeti uzorak koji se ne vidi dobro od rastinja ili kamenja, a da je na toj poziciji raspadnuta uginula životinja, zmija ili insekt koji može da ubodom nanese povredu. Iz tog razloga uzorci se nikada ne uzimaju golim rukama, a rizična i slabo vidljiva mesta uvek treba proveriti pomenutim alatima.

- Ponašanje na terenu treba da je u skladu sa osnovnim principima zaštite na radu. Nije preporučljiva praksa da na teren ide jedna, sama osoba. Na terenskim radovima su moguće povrede, nekada teške, a dešavaju se i incidenti sa smrtnim ishodom. Kolege na terenu uvek treba da su jedan drugom u vidokrugu, a ne da se čuju, tj. dovikuju i sl. Moguće je da se zvuk odbije od stene ili većeg drveta i da se stekne pogrešan utisak o položaju kolege u datom trenutku. Takođe ne treba po svaku cenu rizikovati povredu ili život na osulinama i strminama da bi se uzeo uzorak ili uradila fotografija. Prvenstveno treba imati u vidu da se terenske aktivnosti vrše dok ima dnevnog svetla. Poželjno je otići i vratiti se sa terena za vreme dana i izbegavati vraćanje sa terena noću kada su ljudi posle terenskog dana umorni. Na observacionoj tački ne bi trebalo ostajati duže od 45 minuta do 1 sata, naravno uvek postoje izuzetni slučajevi kada je potrebno zbog neke pojave provesti na određenom prostoru i ceo dan. Neophodno je napraviti pauzu, sa lakšim obrokom i uneti dovoljno tečnosti. Nekada je potrebno i pored velikih vrućina obući laganiju košulju sa dugačkim rukavima i eventualno po izlasku na teren se poprskati nekim insekticidom predviđenim za takve situacije (autan). Obadi, a posebno stršljeni mogu izazvati povrede, a kod alergičnih osoba i probleme sa disanjem. Takođe koža može reagovati i sa biljkama. Moguće su i ogrebotine od kojih prikladna terenska odeća štiti. Posebno je važno imati, terensku, obuću sa debljim đonom i kramponima da bi se sprečilo eventualno proklizavanje. Takođe moguće su i vremenske nepogode, tako da bi trebalo pratiti razvoj vremenske situacije na prostoru koji se istražuje. Jedna od meteoroloških opasnih situacija, pored udara munje je i naglo povećanje količine vode u vodotocima, pri velikim padavinama, posebno potocima gde je nagib terena nešto izraženiji. Malo opreme prve pomoći nije teško nositi, jer su moguće ozbiljnije ogrebotine, ubodne rane i sl. Naravno kod ozbiljnijih povreda potrebno je kontaktirati stručne službe, ili ako je moguće, povređenog transportovati terenskim vozilom do medicinske ustanove. Preporuka je i da se sa terena ne vraća kući isti dan posle terenskog rada, nego da se prespava i da se ekipa sutra dan sveža i odmorna vrati sa terena.

#### 4. III FORMIRANJE FINALNE PROSTORNE BAZE PODATAKA I PISANJE IZVEŠTAJA

Generalno baza podataka je na neki način osmišljena još u prvoj fazi rada sa prostornim podacima (folderi sa kartama, fotografijama, tekstovima i tabelama). Napominjemo da postoje specijalizovane prostorne baze podataka u raznim institucijama, ali njihovo objašnjenje i elaboracija nije svrha ovog rada. U ovoj "trećoj" fazi treba da se organizuju prethodni i terenskim radom prikupljeni podaci i uporede sa starim. Često dosta vremena je potrebno, u ovoj fazi, za upoređivanje podataka. Moguće je da će neki podaci biti slični, a neki različiti.

Jedan od prvih zadataka je upoređivanje tekstova rezultata starih istraživanja sa novim podacima prikupljenih sa observacionih tačaka. Tokom ovog procesa potrebno je prikazati pojave koje su ispitivane i dati objašnjenje zašto su novi podaci slični ili različiti od starih podataka i zaključaka.

Odmah po pregledu tekstualnih podataka, a i tokom tog procesa radi se na izradi karata koje su rezultat terenskog rada. Pozicije observacionih tačaka, putanje, poligoni koji predstavljaju različite objekte, eventualne distribucije polutanata i drugo od interesa se stavlja na neku (npr. topografsku podlogu), nove karte se štampaju, ne nužno na papiru, može kao JPEG, TIFF ili PDF i potom se vrše observacije i upoređenja sa prethodnim podacima i donose zaključci. Ovde se mogu generisati detaljne karte pojava od interesa koje mogu biti na satelitskoj, topografskoj ili nekoj drugoj osnovi. Nekada prikaz na topografskoj ili satelitskoj osnovi ima mnogo detalja koji odvlače pažnju od kartiranog objekta od interesa pa je tada potrebno prikazati nove podatke na praznoj podlozi koja je ograničena istražnim prostorom, opštinom ili nekom drugom granicom (zaštićeno područje). U zavisnosti od situacije i potrebe, a i zbog praktičnosti neke karte mogu biti štampane na formatima A4 i A3 ili B5 koji su zgodni za korićenje u okviru tekstualnog dela, ako se na tim formatima vide objekti od interesa, ako ne, štampaju se u potrebnoj razmeri i formatima i pakuju u za to pripremljene džepove na zadnjim koricama izveštaja.

Rezultati analiza (hemijske, tehnološke...) uzorkovanja se uvek prikazuju tabelarno i u tekstu, a kada je to moguće i prikladno i na kartama.

Fotografije, kojih kod nekih istraživanja sa puno observacionih tačaka može biti mnogo, se selektuju i vrši se odabir onih koje će ići u izveštaj. Sve fotografije i one koje nisu ušle u izveštaj se čuvaju kao grafičko tehnička dokumentacija.

Finalni proizvod može biti jedna knjiga (izveštaj ili elaborat) ili više knjiga sa prepisom observacionih tačaka, prikazom svih fotografija i drugim priložima (radne verzije karata, tabele i drugo).

Sve pomenuto se nalazi i u digitalnoj formi u folderima, jednostavnoj bazi podataka i može se uvek nadograđivati. Obavezno se u digitalnoj formi pravi rezervna kopija svog materijala.

## 5. ZAKLJUČAK

Cilj ovog rada je bio da se prikaže postupak tokom izvođenja terenskog rada stručnjacima različitih oblasti, koji iz raznih razloga treba da rade na terenu i prikupe prostorne podatke vezane za neku problematiku. Polazna pretpostavka je da ovi stručnjaci nemaju dovoljnih iskustava u terenskom radu. Takođe, kao početna premisa rada je činjenica da svaka struka ima svoja pravila, ali u terenskim uslovima postoje neki opšti principi istraživanja. Tako je predmet ovog rada zasnovan na osnovama opštih principa terenskog rada, koje smo ovde pokušali da u kratkim crtama prikažemo. U suštini, rad predstavlja model koji bi svako ko se prvi put susreće sa terenskim radom, tj. radom na nekom prostoru trebalo da pročita. Naglašava se da je problematika mnogo šira i zahtevnija. S toga, smatramo da je potrebno više proučavanja i rada po opisanim stavkama. Nadamo se da ćemo u razumnom vremenskom roku dati i širi prikaz.

## LITERATURA

- [1] Nacionalna geografija Srbija, 2021: Novo otkriće, nova misterija, Analiza uzoraka stene sa Meseca otvara nova pitanja. <https://nationalgeographic.rs/nauka/vesti-i-zanimljivosti/a27677/Analiza-stene-sa-meseca.html>
- [2] Euronews Serbia, 2021: NASA ostvarila istorijski uspeh potvrđeno da je rover Perzeverans sakupio uzorke tla sa Marsa. <https://www.euronews.rs/magazin/tehnologija/15815/nasa-ostvarila-istorijski-uspeh-potvrđeno-da-je-rover-perzerverans-sakupio-uzorke-tla-sa-marsa/vest>
- [3] Rakić, M., Dimitrijević M, Marković M.,1976. Tumač za list Kruševac K34-19, Redakcioni odbor, Milorad Dimitrijević Stevan Karamata Boris Sikošek Dobra Veselinović, Osnovna geološka karta 1:100 000, Savezni geološki zavod, Beograd, s 1-54
- [4] Charlotte Jee.2019. The deepest ever dive to the bottom of the Mariana Trench, found litter there. <https://www.technologyreview.com/2019/05/14/135393/the-deepest-ever-dive-to-the-bottom-of-the-mariana-trench-found-litter-there/>
- [5] Geospatial World, 2009. GIS and scanning technology. <https://www.geospatialworld.net/article/gis-and-scanning-technology/>
- [6] <https://desktop.arcgis.com/en/arcmap/latest/manage-data/tables/understanding-how-to-use-microsoft-excel-files-in-arcgis.htm>
- [7] <https://sasplanet.geojamal.com/>

## DALJINSKO PRAĆENJE STANJA ENERGETSKIH POSTRIJENJA

### REMOTE MONITORING OF ENERGY PLANT STATE

NENAD GALJAK<sup>1</sup>, MIROSLAV VUJASINOVIĆ<sup>2</sup>, MIODRAG REGODIĆ<sup>2</sup>, MILOŠ BLAGOJEVIĆ<sup>2</sup>

<sup>1</sup>Univerzitet odbrane, Vojna akademija Beograd, galjaknenad@gmail.com

<sup>2</sup>Univerzitet u Banjoj Luci, Arhitektonsko-građevinsko-geodetski fakultet, miroslav.vujasinovic@aggf.unibl.org, miodrag.regodic@aggf.unibl.org, milos.blagojevic@student.aggf.unibl.org

**Rezime:** Termalno snimanje predstavlja nekontaktni metod, kojim se u vrlo kratkom vremenu registruje emitovanje toplote, odnosno infracrveno zračenje. Registrovanje tih zračenja našlo je široku primjenu kod praćenje različitih pojava u mnogim oblastima ljudskog djelovanja, kao što su elektroenergetika, mašinstvo, građevinarstvo, arhitektura, geodezija. Termalni senzori koriste jednu ili više unutrašnjih temperaturnih referenci za poređenje sa otkrivenom radijacijom, tako da mogu biti povezane apsolutnom radijalnom temperaturom. U radu su predstavljeni suština termalnog snimanja, kao i primjena termovizije prilikom dijagnostike toplotnih stanja energetskih transformatora, odnosno, termovizijsko snimanje i analiza stanja elektroenergetskih postrojenja metodom infracrvene termografije.

**Ključne reči:** Termalno snimanje, Senzor, Termogram, Elektroenergetsko postrojenje.

**Abstract:** Thermal observation is non-contact method for register of heat emission in short time., ie infrared radiation. Registration of these radiation has found application in the monitoring of various phenomena in areas of human activity, such as power engineering, mechanical engineering, construction, architecture and geodesy. Thermal sensors use one or more inside temperature reference to compare with detected radiation, so they can be refer to absolute radial temperature. In paper presents the essence of thermal observation as well as the application of thermal observation in the diagnosis of thermal condition of power transformers, thermal observation and analysis of the power plant state by infrared termography.

**Keywords:** Thermal observation, Sensor, Thermogram, Power plant

## 1. UVOD

Daljinska detekcija omogućava globalno sagledavanje Zemlje, a samim tim i promjena koje se dešavaju na Zemlji. Te promjene mogu biti izazvane prirodnim ili antropogenim djelovanjem. Snimci načinjeni daljinskim snimanjem pomogli su čovjeku da bolje koristi životnu sredinu i prirodne resurse, a sa razvojem novih tehnologija očekuje se i proširenje primjene satelitskih snimaka za dobrobit čovječanstva [2].

Mnogi multispektralni sistemi djeluju na području termalnog infracrvenog, kao i vidljivog i infracrvenog dijela spektra elektromagnetne energije. Termalni senzori koriste foto detektore osjetljive na direktne kontakte fotona na svojoj površini, da bi otkrili emitovano termalno zračenje. Detektori se hlade na temperaturi blizu apsolutne nule, kako bi se ograničili na vlastitu termalnu emisiju.

Termalni senzori u suštini mjere temperaturu površine i termičko svojstvo cilja. Termalne kamere su uglavnom tipa poprečnog skenera koji detektuju zračenje emitovano samo u termalnom dijelu spektra.

Termografski senzori imaju veliku primjenu. Oblasti u kojima se termalno snimanje primjenjuje su: energetika, elektronika, mašinstvo, građevinarstvo, geodezija, pa čak i određena polja medicine.

Podaci termalnih senzora se često kombinuju sa podacima iz drugih područja spektra elektromagnetne energije, kako bi se što bolje analiziralo stanje posmatranog objekta. Važna primjena termalnog snimanja je u šumarstvu i ogleda se u ranoj detekciji šumskih požara, naročito u zaštićenim područjima.

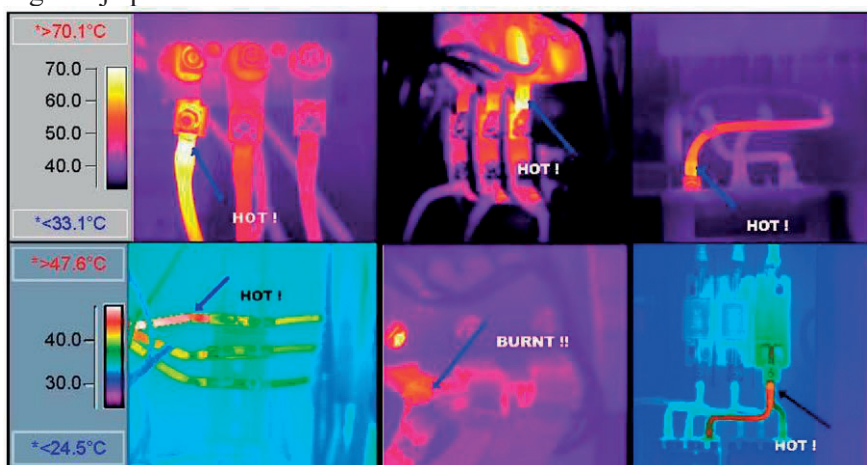
Kod industrijskih postrojenja, da bi ostvarili ciljeve proizvodnje, pogoni često rade bez zastoja. Kako bi se izbjegli skupi kvarovi, koji najčešće nastaju usljed pregrijavanja pojedinih dijelova kao i nepredviđeno izgubljeni vrijeme, kao preventiva se koriste termalna snimanja.

Kada se u velikim postrojenjima ne bi koristile preventivne mjere, često osim gubitaka u proizvodnji mogle bi izazvati se i veće nesreće, poput požara i ljudskih žrtava.

Termalni senzori se koriste pri kontroli opreme niskog, srednjeg i visokog napona, te distributivnih, napojnih i energetskih transformatora. Pojava zagrijavanja u mnogim slučajevima ukazuje na novonastali kvar. Velika prednost termografskih senzora je ispitivanje i mjerenje dok su električni sistemi pod naponom.

U oblasti mašinstva i energetike, termografski podaci predstavljaju neprocjenjiv izvor informacija. Kada se mehaničke komponente istroše, one prestaju da djeluju i temperatura počne da se povećava.

Istrošeni ležaj će se uočiti na infracrvenom zapisu, pa se njegovom zamjenom postiže normalno operativno stanje određenog uređaja. Pored ove primjene osigurava se i precizna analiza pri održavanju kotlovnica i upravljanje vatrostalnim gubicima i slično. Takođe se mogu otkriti i začepjenja cijevi. Primjena termografije u energetici je prikazana na slici 1.

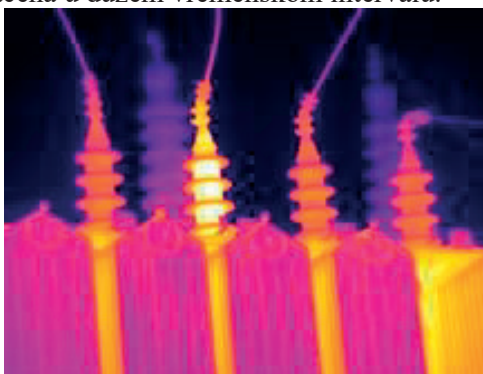


Slika 1: Primjena termografije u energetici [5]

## 2. PRIMJENA TERMOGRAFIJE U DIJAGNOSTICI STANJA ENERGETSKIH POSTROJENJA

Termalna snimanja su neophodna kod otkrivanja neregularnih toplotnih stanja na energetskim transformatorima, kroz kontrolu priključnih veza transformatora, transformatorskog suda, sklopova i prateće opreme na transformatoru.

Kvarovi u unutrašnjosti transformatora, odnosno na strujnim vezama od priključaka ka namotajima nisu tako česti, ali kada se dese zahtijevaju posebnu pažnju, naročito kod transformatorskih jedinica većih snaga. Pravovremeno otkrivanje ove vrste neispravnosti daje mogućnost planiranja isključenja za najpovoljniji trenutak po elektroenergetski sistem. U međuvremenu se opterećenje transformatora smanjuje do nivoa koji dozvoljava sistem i pooštreno prati stanje transformatora kako bi se utvrdio trend razvoja kvara. Na slici 2 je predstavljen slučaj pregrijavanja tijela provodnog izolatora transformatora, kod kog je termovizijom pronađena neispravnost, koja je praćena u dužem vremenskom intervalu.



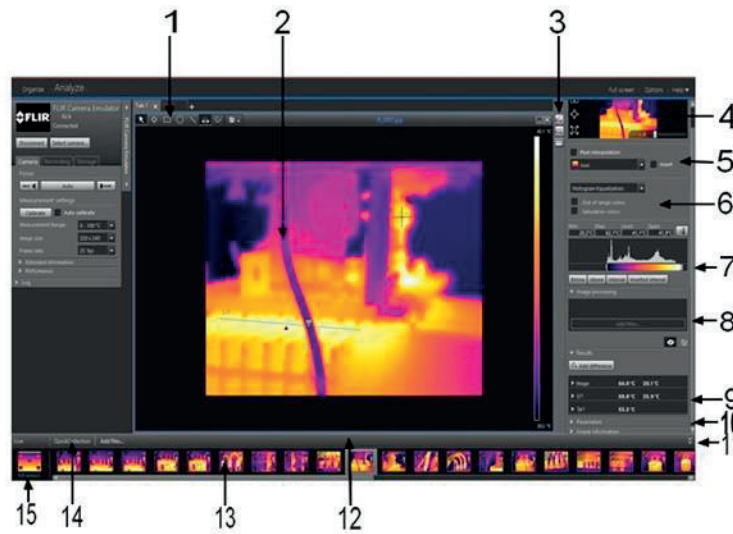
Slika 2: Pregrijavanje provodnih izolatora transformatora, otkriveno termalnim snimanjem [4]

Sistem za hlađenje i čišćenje ulja, koji se kontroliše termografskom metodom, čine hladnjaci i uljne pumpe. Ocjena efikasnosti rada sistema za hlađenje sa prinudnom cirkulacijom ulja se vrši na osnovu temperaturnih promjena po visini hladnjaka. Ocjenjuje se karakter zagrijavanja razmjenjivača toplote, koji zavisi od čistoće spoljašnjih i unutrašnjih površina cijevi, kao i od brzine obrtanja ventilatora za hlađenje i od uljnih pumpi. Zagrijavanje uljnih pumpi se upoređuje sa tipičnim zagrijavanjem pumpi analognog tipa [9]. Na slici 3 prikazani su termalni snimak, fotografski snimak i temperaturni profil uljne pumpe.



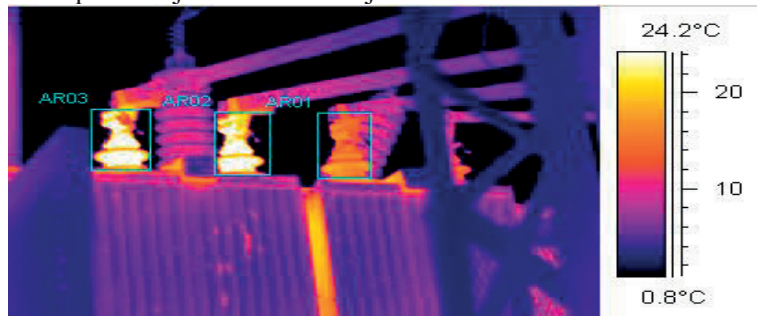
Slika 3: Termalni snimak, fotografski snimak i temperaturni profil uljne pumpe [4]

Za analizu termograma najčešće se koristi softver FLIR R&D Software 3.3. (slika 4), koji ima mogućnost upravljanja režimom snimanja termograma, postavljanje uslova u pogledu temperature (min, max.) pojedinih zona na termogramu, upozorenje korisnika o prekoračenju postavljenih uslova, generisanju izvještaja za odabrani termogram [6].



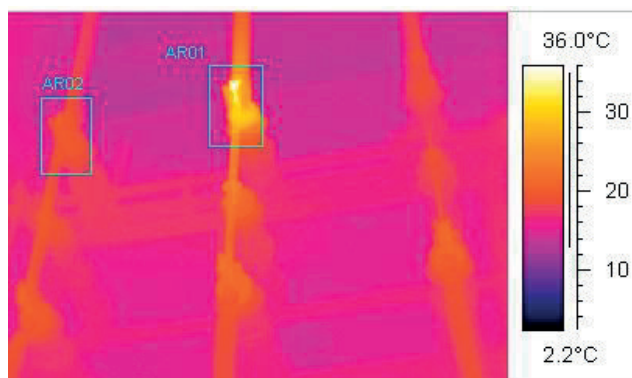
Slika 4: Radno okruženje softvera FLIR R&D Software 3.3. [3]

Pregledom sledećeg termograma, primjetno je da postoji zagrijavanje na transformatoru, što je prikazano na termogramu transformatora T3 110/35/10 kV na slici 5, koje može biti prouzrokovano velikim prelaznim otporom kontakata, velikim opterećenjem ili nesimetrijom faza.



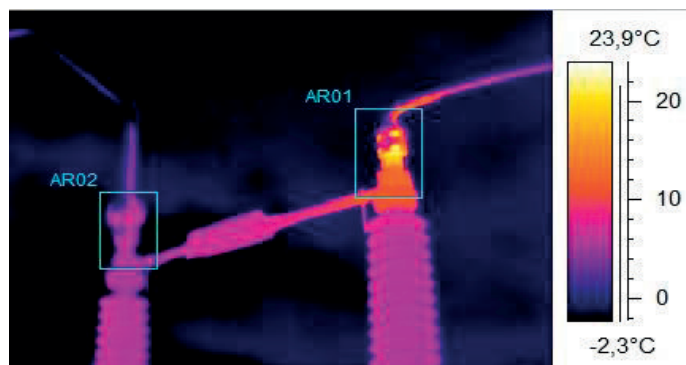
Slika 5: Termogram transformatora T3 110/35/10 kV [3]

U svrhu preventivnog održavanja i izbjegavanja neželjenih prekida u isporuci električne energije potrebno je interventno otklanjanje uzroka zagrijavanja i sanacija mjesta zagrijavanja. Pregledom termograma uočeno je povećano zagrijavanje priključka izlaznog rastavljača prema dalekovodu. Termogram izlaznog rastavljača je predstavljen na slici 6.



Slika 6: Termogram izlaznog rastavljača [3]

Obradom i analizom termograma linijskog rastavljača (slika 7), utvrđuje se zagrijavanje na priključku rastavljača prema sabirnicama, pa se preporučuje saniranje mjesta zagrijavanja na priključku sabirničkog rastavljača.



Slika 7: Termogram linijskog rastavljača [3]

Na slici 8 prikazan je termogram pola prekidača prema sabirničkom rastavljaču u dalekovodnom polju.



Slika 8: Termogram pola prekidača [3]

Na osnovu navedenih primjera može se zaključiti da se zagrijavanja se najčešće dešavaju na priključcima rastavljača i prekidača. Primjetno je da, ispravni spojevi, kontakti prekidača i rastavljača imaju temperaturu približnu temperaturi okoline, dok „loši“ spojevi i kontakti imaju temperature koje znatno odstupaju od temperature okoline, što direktno ukazuje na stanje tačaka koje je potrebno sanirati. Najčešći uzrok ovoj pojavi su oksidacija i loši spojevi gdje zbog prelazne otpornosti kontakta dolazi do povećanog zagrijavanja.

Da bi se umanjila ova pojava potrebno je redovno vršiti preventivnu termovizijsku kontrolu, tako da se na vrijeme mogu uočiti tačke u kojima je došlo do povećanja temperature, koja može dovesti do oštećenja elemenata u postrojenju, pa i do ispadanja dijelova ili cijelog postrojenja i sistema, što za posljedice može imati veliku materijalnu i finansijsku štetu, kako za proizvođača tako i za potrošača. Uzroci povećanog zagrijavanja ne moraju biti samo loši spojevi.

Povećana temperatura na nekom od elemenata postrojenja, npr. na provodnim izolatorima na transformatoru, ukazuje da je posmatrani dio u lošem stanju i potrebno je taj dio zamjeniti u što kraćem roku.

### 3. ZAKLJUČAK

Dosadašnji rezultati daljinskih snimanja i istraživanja Zemlje u različitim oblastima: geologiji, hidrologiji, šumarstvu, poljoprivredi, kartiranju zemljišta i vegetacije, inženjerskom projektovanju, fotogrametriji, kartiranju prirodnih i vještački izazvanih katastrofa, prostornom planiranju, ekologiji, arheologiji, okeanografiji, vojnim i drugima potrebama, predstavljaju snažan podsticaj razvoju novih senzora i platformi za snimanja i istraživanja iz svemira [1].

Termalno snimanje predstavlja beskontaktni metod snimanja kojim se u kratkom vremenu registruje emitovanje toplote, odnosno infracrveno zračenje. Kao takva, ova tehnika je idealna za identifikovanje loših mjesta, odnosno mjesta u energetskim postrojenjima u kojima postoji određena nepravilnost ili kvar, kao i da bi se dobila procjena gubitka toplote.

Rezultat termografskog snimanja predstavlja termogram, koji u sivim tonovima ili nekom kodu boja, daje sliku temperaturne raspodjele na površini posmatranog objekta. Pomoću infracrvene termografije moguće je vidjeti gdje nastaju gubitci toplotne energije kod zgrada, elektropogona, toplovodnih ili paravodnih cijevnih instalacija i elemenata (podzemnih i nadzemnih), kao i vršiti kontrolu stanja požara.

Efikasnim praćenjem temperaturnih režima energetskih transformatora stvara se podloga za upravljanje ovim režimima. Rezultat toga može biti smanjenje trajanja zastoja i troškova popravki, odnosno produženje vijeka korištenja transformatora.

Primjenom infracrvenog termičkog zračenja dolazi se do informacija o stanju opreme u eksploataciji, a praćenjem porasta temperature u odnosu na temperaturu okoline i trenutnog oprećenja na karakterističnim mjestima, može se odrediti stepen opasnosti od povećanog zagrijavanja, te se blagovremenom preventivnom intervencijom znatno može povećati pouzdanost i raspoloživost određenog postrojenja.

Sagledavši sve metode i karakteristike termografskog snimanja, sa sigurnošću se može tvrditi da ova tehnika predstavlja najpouzdaniji vid otkrivanja nepravilnosti i grešaka u radu elektroenergetskih postrojenja.

### LITERATURA

- [1] Regodić, M. (2008). Važniji satelitski programi sistematskog praćenja Zemlje, *Vojnotehnički glasnik* 4/2008, 70-88.
- [2] Regodić, M., & Željana Benić, Ž. (2014). Daljinska opažanja prirodnih pojava, *Vojnotehnički glasnik*, vol. 62, br. 3, 101-120.
- [3] Šaraba, P., Popović, B., & Krsmanović, D. (2015). Preventivno termovizijsko snimanje elektroenergetskih postrojenja metodom IR termografije, *INFOTEH*, Vol. 14, Jahorina
- [4] Čičkarić LJ. (2009). Primena termografije u dijagnostici toplotnih stanja energetskih transformatora, *Elektrotehnički institut „Nikola Tesla”*, Beograd
- [5] <http://www.lceng.com/services/infrared-thermography-services-malaysia>





## SATELITSKO SNIMANJE PROGRAMOM COPERNICUS

### SATELLITE OBSERVATION WITH COPERNICUS PROGRAM

NENAD GALJAK<sup>1</sup>, MIROSLAV VUJASINOVIĆ<sup>2</sup>, MIODRAG REGODIĆ<sup>2</sup>, SLAĐANA STANIŠIĆ<sup>2</sup>

<sup>1</sup>Univerzitet odbrane, Vojna akademija Beograd, galjaknenad@gmail.com

<sup>2</sup>Univerzitet u Banjoj Luci, Arhitektonsko-građevinsko-geodetski fakultet, miroslav.vujasinovic@aggf.unibl.org, miodrag.regodic@aggf.unibl.org, sladjana.stanistic@student.aggf.unibl.org

**Rezime:** Evropski satelitski program za opažanje Zemlje Copernicus, jedan je od najvećih programa Evropske komisije, čiji je zadatak davanje poboljšanih informacija o Zemlji. Copernicus servisi prate stanje u oblastima zaštite životne sredine, upravljanja urbanim područjima, regionalnog i lokalnog planiranja, poljoprivrede, šumarstva, ribarstva, zdravstva, transporta, promjene klime, održivog razvoja i drugim. Primarni Copernicus servisi daju potrebne informacije za donošenje odluka u slučaju vanrednih situacija, kao što su prirodne katastrofe i humanitarne krize. U radu su detaljno opisani servisi i senzori satelitskog programa Copernicus, te primjeri primjene, interpretacije i analize satelitskih snimaka ovog programa u mnogim oblastima ljudskog djelovanja.

**Ključne reči:** Satelitsko snimanje, Copernicus, Satelitski snimci, Životna sredina.

**Abstract:** The European satellite program for Earth observation Copernicus is one of the largest programs of the European Commission, whose task is provide improved information about the Earth. Copernicus service monitors of environmental protection areas, urban management, regional and local planning, agriculture, forestry, fisheries, health, transport, climate change, sustainable development and others. Copernicus services provide the necessary information for decision-making in case of emergencies, such as natural disasters and humanitarian crises. The paper describes in detail the services and sensors of the Copernicus satellite program, as well as examples of application, interpretation and analysis of satellite images of this program in many areas of human activity.

**Keywords:** Satellite observation, Copernicus, Satellite images, Environment

#### 1. UVOD

Evropski Copernicus program za opažanje Zemlje jedan je od najvećih programa Evropske komisije čiji je cilj davanje poboljšanih informacija o Zemlji. Copernicus servisi pokrivaju šest tematskih područja: zemlju, more, atmosferu, promjenu klime, upravljanje hitnim intervencijama i sigurnost.

Proizvodi Copernicus servisa našli su široku primjenu u oblastima zaštite životne sredine, upravljanja urbanim područjima, prostornog planiranja, poljoprivrede, šumarstva, ribarstva, zdravstva, transporta, promjene klime, održivog razvoja i drugim. Primarni Copernicus servisi pružaju potrebne informacije za donošenje odluka o životnoj sredini i donošenju odluka u slučaju vanrednih situacija, kao što su prirodne katastrofe i humanitarne krize [3].

Svim Sentinel misijama i satelitima, na kojima se zasniva Copernicus program, upravlja Evropska svemirska agencija. Za održavanje infrastrukture svemirske komponente programa je osim ESA-e zadužen je i EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites). Djelovanje vezano za Sentinele povjereno je ESA-i i EUMETSAT-u, zbog specifičnih vještina koje te dvije organizacije posjeduju, a sateliti Sentinel misije (SM) su vlasništvo Evropske Unije.

## 2. COPERNICUS SERVISI

Komponenta Copernicus programa koja koristi podatke dobijene Sentinel satelitskim misijama i nadopunjuje ih in-situ mjerenjima su Copernicus servisi. Njihova namjena je davanje usluga za upravljanje i zaštitu životne sredine, upravljanje prirodnim resursima.

Copernicus servisi pružaju informacije za šest glavnih područja: praćenje kopna, praćenje mora i oceana i praćenje atmosfere, hitne intervencije, sigurnost i klimatske promjene. Ovu komponentu nadgleda Uprava Evropske komisije. U rad Copernicus servisa uključene su i neke druge evropske agencije, poput European Environment Agency (EEA), European Agency for the Management of Operational Cooperation at the External Borders of the Member States of the European Union (FRONTEX), European Satellite Centre [3]. Na slici 1 prikazani su servisi programa Copernicus.

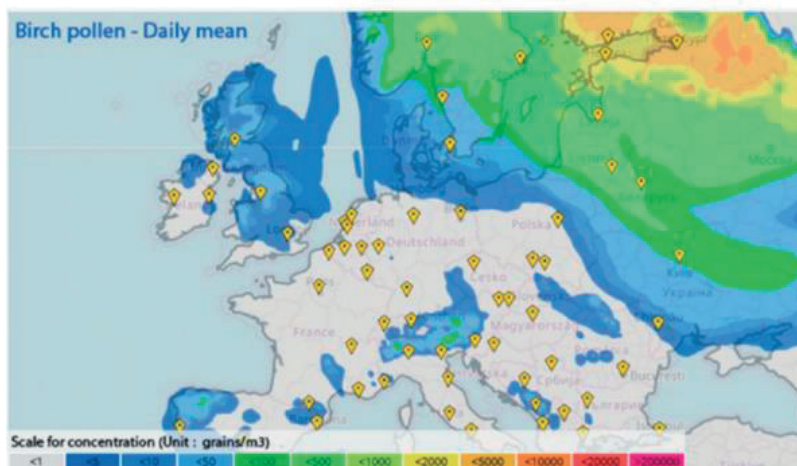


Slika 1: Servisi programa Copernicus [4]

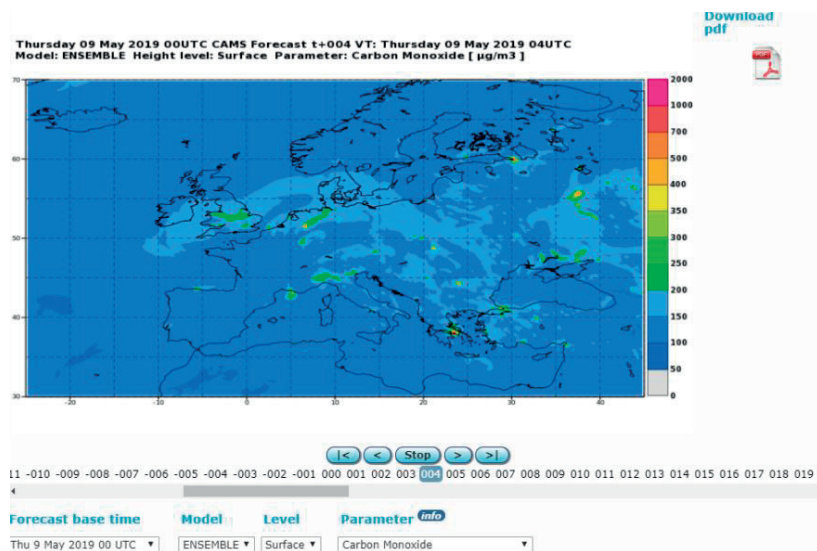
**Servis za praćenje atmosfere (CAMS)** radi za područja kvalitet vazduha i sastav atmosfere, stanje ozona i UV zračenje, emisija i dioba CO<sub>2</sub>, solarno zračenje, klimatski faktor.

Servis pruža analize u realnom vremenu i četvero-dnevne prognoze, kao i reanalizu evropskog kvaliteta vazduha, čime se omogućuje stalna procjena vazduha.

Na slici 2 je prikazana je koncentracija polena na području Evrope, a na slici 3 stanje CO nad Evropom, što su proizvodi servisa za praćenje atmosfere.



Slika 2: Koncentracija polena na području Evrope [3]

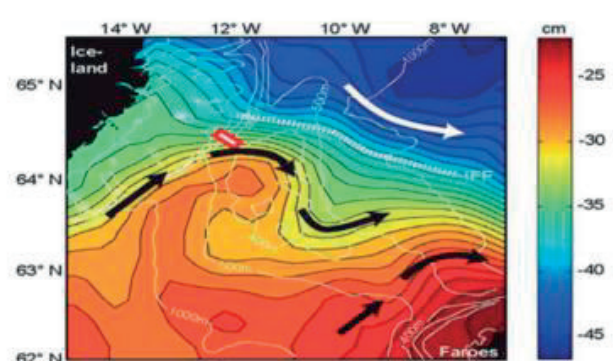


Slika 3: Stanje CO nad Evropom [4]

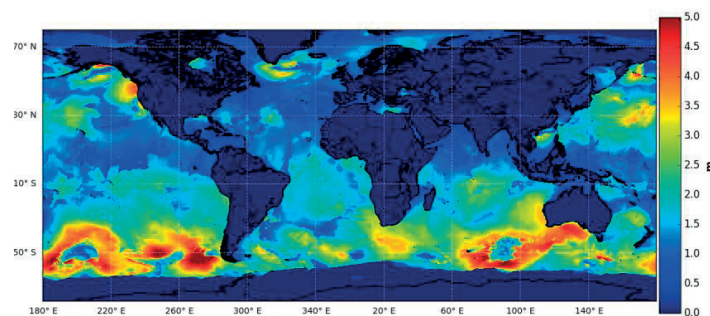
**Servis za praćenje mora (CMMS)** pruža informacije o fizičkom stanju, varijabilnosti i dinamici ometa i morskih ekosistema za globalni okean i evropski regionalni okean.

Mnogi podaci koje servis isporučuje (npr. temperatura, slanost, nivo mora, struje, vjetar i debljina leda), takođe imaju ključnu ulogu u domenu vremenskih, klimatskih i sezonskih predviđanja.

Na slici 4 prikazan je jedan od konačnih rezultata servisa, kretanje morskih struja, za potrebe praćenja promjena nivoa mora na nekom području. Bijele linije povezuju mjesta s istim srednjim nivoom mora čiji je raspon prikazan određenom bojom. Strelice pokazuju smjer kretanja morskih struja. Analiza visine talasa na osnovu Copernicus snimaka prikazana je na slici 5, što je takođe rezultat servisa za praćenje mora.



Slika 4: Kretanje morskih struja [3]  
EU Copernicus Marine Service Global Waves analysis and forecast



Slika 5: Analiza visine talasa na osnovu Copernicus snimaka [4]

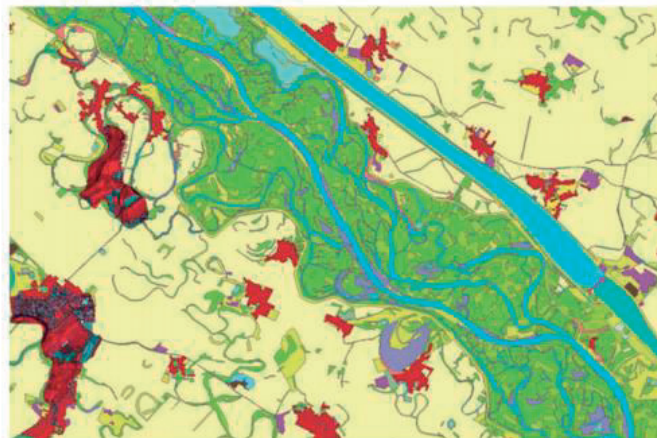
Na slici 6 prikazan je tajfun Hagibis koji je često ugrožava japansko ostrvo Honšu. Japan je u stalnoj opasnosti od dejstva tajfuna i time izložen velikoj šteti od jakog vjetra i bujica. Ovaj ogromni tajfun, koji se upoređuje sa uraganom kategorije 5, snimljen je Copernicus Sentinel-3 misijom 10. oktobra 2019. godine. Središnji dio, tzv. „oko oluje“ ima prečnik oko 60 km [5].



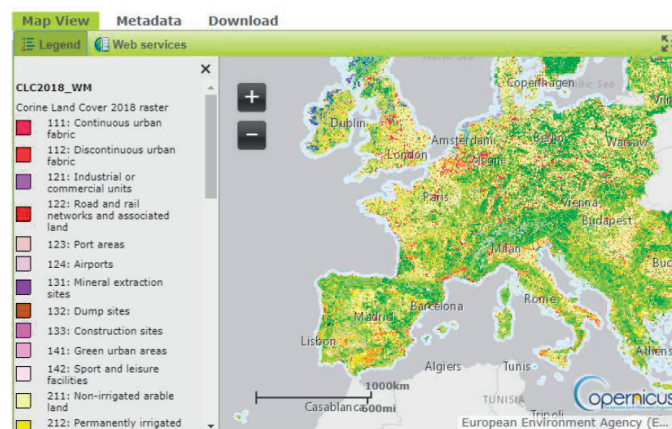
Slika 6: Tajfun Hagibis, snimljen Copernicus Sentinel-3 satelitom [5]

**Servis za praćenje zemljišta (CLMS)** pruža informacije koje se odnose na korištenje zemljišta i promjeni korištenja zemljišta. Servis podržava aplikacije na različitim područjima za potrebe prostornog planiranja, upravljanja šumama, upravljanja vodama, poljoprivrede.

Na slici 7 je prikazana karta zemljišta servisa za praćenje zemljišta. Zelenom bojom je prikazana vegetacija, plavom voda, a crvenom naseljena mjesta. Stanje zemljišta na području Evrope krajem 2018. godine prikazano je na slici 8, što je rezultat servisa za praćenje zemljišta.

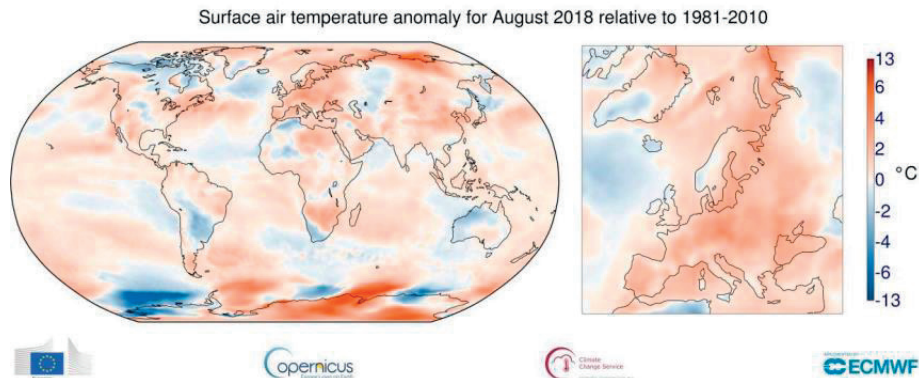


Slika 7: Karta zemljišta servisa za praćenje zemljišta [3]



Slika 8: Stanje zemljišta na području Evrope krajem 2018. godine [4]

**Servis za praćenje klimatskih promjena (CZS)** zamišljen je kao reakcija na ekološke i društvene izazove povezane sa klimatskim promjenama. Usluga prati nekoliko klimatskih pokazatelja (povećanje temperature, porast nivoa mora, topljenje ledenog sloja, zagrijavanje okeana) i klimatskih indeksa (na temelju zapisa o temperaturi, padavinama i sušnim događajima). Na slici 9 prikazana je anomalija temperature vazduha za avgust 2018. godine što je rezultat servisa za praćenje klimatskih promjena.

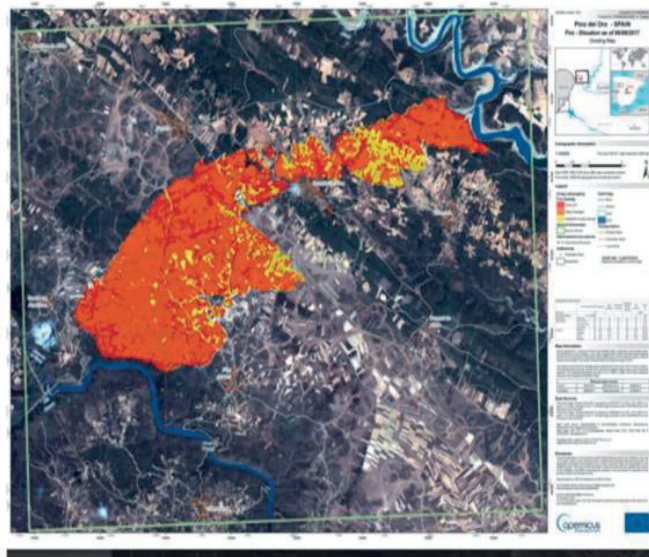


Slika 9: Anomalija temperature vazduha [4]

Na osnovu satelitskih snimaka sinoptičari dobijaju i pouzdane informacije o vremenskim uslovima, analizirajući osobine oblaka, a koje se odnose na stanje temperature, temperature oblačnog sloja, stepen oblačnosti, profile sadržaja vlage, parametre vjetra, koje osnovnu primjenu nalaze u u prognozi vremena [2].

**Servis za hitne slučajeve (CEMS)** pruža informacije svim akterima uključenim u upravljanje prirodnim katastrofama, ljudskim potrebama i hitnim slučajevima, sa preciznim geo-prostornim podacima izvedenim iz satelitskih Sentinel misija i dopunjenim in-situ mjerenjima.

Evropski informacioni sistem šumskih požara (EFFIS) pruža u realnom vremenu informaciju o šumskim požarima i šumskim požarnim režimima u evropskim, bliskoistočnim i sjevernoafričkim regijama. Na slici 10 je prikazana digitalna karta rizika od pojave požara, crvenom bojom su prikazana područja koja imaju najveću mogućnost pojave požara [3].



Slika 10: Digitalna karta rizika od pojave požara [2]

**Servis za sigurnost (CSS)** ima za cilj da podrži politiku EU pružanjem informacija, kao odgovor na sigurnosne izazove.

U području nadzora državne granice glavni ciljevi rada servisa su: smanjenje broja ilegalnih imigranata koji ulaze u EU neotkriveni, smanjivanje broja poginulih ilegalnih imigranata, spašavanje više života na moru, povećanje unutrašnje sigurnosti Evropske unije i doprinos prevenciji prekograničnog kriminala. U području pomorskog nadzora opšti ciljevi rada servisa su: osiguranje sigurnosti pomorske plovidbe, podrška kontroli ribarstva, borba protiv zagađivanja mora i provođenje zakona [2].

### 3. ZAKLJUČAK

Uvođenje modernih tehnologija za praćenje, analizu i upravljanje prirodnim stanjima i bogatstvima, kao i stalno praćenje stanja na površini Zemlje, zahtijeva pravovremene i objektivne informacije, često stanju u realnom vremenu i stvarnim mogućnostima kontrole promjena koje su neprekidne. Upravo zbog mogućnosti široke primjene podataka dobijenih posredstvom vještačkih Zemljinih satelita, osiguran je široki spektar proizvoda satelitskih misija [1].

Copernicus je satelitski program Evropske unije, čija je zadatak sistematsko snimanje područja zemlje, mora, atmosfere, promjena klime, upravljanja hitnim intervencijama i sigurnost. Kao rezultat aktivnosti Copernicus programa dobija se cijeli niz geoprostornih proizvoda i usluga.

Proizvodi programa se koriste za analizu i poboljšanje sigurnosti, održivog razvoja, načina upravljanja prirodnim resursima, zaštite životne okoline, praćenja klimatskih promjena i drugo. Satelitski snimci u kombinaciji sa in-situ mjerenjima imaju važnu ulogu u praćenju, interpretaciji i kartiranju područja pogođenih nekom prirodnom nepogodom, kao i u fazi otlanjanja posljedica djelovnja prirodnih nepogoda.

Copernicus je jedan od najboljih online servisa, po pitanju rezolucije snimaka i količine dostupnih podataka. Stalni razvoj servisa i povećanje kvaliteta proizvoda, obezbjeđuju servisu vodeću ulogu u oblasti praćenja prirodnih nepogoda i ostalih atmosferskih pojava.

### LITERATURA

- [1] Regodić, M. (2008). Važniji satelitski programi sistematskog praćenja Zemlje, *Vojnotehnički glasnik* 4/2008, 70-88.
- [2] Regodić, M., & Željana Benić, Ž. (2014). Daljinska opažanja prirodnih pojava, *Vojnotehnički glasnik*, vol. 62, br. 3, 101-120.
- [3] Jakopović, L., Kolarek, M., Rezo, M., & Kranjčić, N. (2018). Satelitska misija Copernicus – mogućnosti nadziranja prirodnih nepogoda, *Geotehnički fakultet*, Varaždin
- [4] Sabljak, D. (2019). Šta snimaju Sentinel sateliti, Zagreb
- [5] <https://www.copernicus.eu/en/media/images/typhoon-hagibis>

## UTICAJ VOJNIH INTERNET STVARI NA GEOPROSTORNU ANALIZU BOJIŠTA

### THE INFLUENCE OF MILITARY INTERNET OF THINGS ON BATTLEFIELD GEOSPATIAL ANALYSIS

IVANVULIĆ<sup>1</sup>, RADOMIR PRODANOVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet odbrane – Vojna akademija, ivan.vulic@mod.gov.rs

<sup>2</sup> Vojska Srbije – Centar za primenjenu matematiku i elektroniku, radomir.prodanovic@vs.rs

---

**Rezime:** U radu će biti analiziran uticaj vojnih internet stvari na geoprostornu analizu bojišta. Biće definisan pojam vojnih internet stvari, njihov značaj na savremenim bojištima, kao i njihov uticaj na geoprostornu analizu bojišta u geografskim informacionim sistemima.

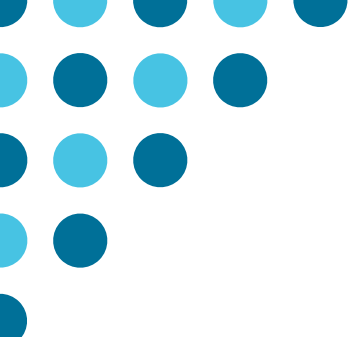
**Ključne reči:** Vojne internet stvari, geografski informacioni sistemi, geoprostorna analiza.

**Abstract:** The paper will analyse the influence of military internet of things on the battlefield geospatial analysis. The notion of military internet of things, their significance on modern battlefields, as well as their influence on battlefield geospatial analysis in geographic information systems will be defined.

**Keywords:** Military internet of things, geographic information systems, geospatial analysis







# GRAFOVI I MREŽE

## GRAPHS AND NETWORKS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business



## **Analiza simetrije u ekonomskom sistemu pre, tokom i posle ekonomske krize primenom teorije grafova** **Symmetry analysis of economic system before, during, and after economic crisis using graph theory**

VOJIN STEVIĆ<sup>1</sup>, MARIJA RAŠAJSKI<sup>2</sup>, MARIJA MITROVIĆ DANKULOV<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu-Elektrotehnički fakultet, Bulevar Kralja Aleksandra 73, 11120 Beograd, Srbija; voj.in.stevic@gmail.com

<sup>2</sup> Univerzitet u Beogradu-Elektrotehnički fakultet, Bulevar Kralja Aleksandra 73, 11120 Beograd, Srbija; rasajski@etf.bg.ac.rs

<sup>3</sup> Institut za fiziku u Beogradu, Univerzitet u Beogradu, Pregrevica 118, 11080 Beograd, Srbija; mitrovic@ipb.ac.rs

**Rezime:** *Ekonomske krize utiču na sve aspekte naših života. Razumevanje kako one nastaju i otkrivanje ranih indikatora njihovog nastanka su od velike važnosti. Dinamika i funkcija bilo kog ekonomskog sistema je usko vezana za strukturu mreže interakcija između njegovih gradivnih elemenata. U ovom radu, koristimo teoriju grafova i grupe automorfizama da analiziramo promenu simetrije u ekonomskom sistemu usled njegovog prolaska kroz krizu. Mreža interakcija između firmi u finansijskom sektoru Sjedinjenih Američkih Država ima visok stepen simetrije pre, tokom i posle krize. Simetrija mreže tokom krizne godine je manja nego simetrija mreža pre i posle krize. Naši rezultati doprinose boljem razumevanju evolucije ekonomskih sistema.*

**Ključne reči:** *teorija grafova, automorfizmi, kompleksne mreže, vremenske serije, ekonomski kompleksni sistemi*

**Abstract:** *Economic crises influence every aspect of our lives. Understanding how they emerge and detecting their early signs is of utmost importance. The dynamics and function of any economic system are closely related to the structure of the network interactions between its constituents. In this work, we use graph theory and a group of automorphisms to analyze the change of symmetry in the economic system due to the economic crisis. We show that the network of financial USA companies has a high level of symmetry before, during, and after the 2008 economic crisis. We show that the symmetry of the network is lower during the economic crisis than before and after. Our results contribute to a better understanding of the evolution of economic systems.*

**Keywords:** *graph theory, automorphism, complex networks, time series; complex economic systems*

### **1. UVOD**

Ekonomski sistem se sastoji od velikog broja interagujućih heterogenih elemenata, čije kolektivno ponašanje ne može da se predvidi na osnovu pojedinačnog ponašanja ovih elemenata. Stoga, interakcije u ekonomskim sistemima imaju vrlo važnu ulogu [1, 2]. Razumevanje dinamike i funkcije ekonomskih sistema je usko povezano sa razumevanjem strukture i dinamike mreže interakcija. Teorija grafova i teorija kompleksnih mreža pružaju vrlo bitne metode i alate za izučavanje evolucije ekonomskih sistema [3].

Ekonomski sistemi nisu statični. Oni konstantno evoluiraju i prolaze kroz faze rasta, krize i recesije [4]. Sve ove faze imaju ogroman uticaj na pojedince i ljudsko društvo [4]. Stoga je neophodno da se razume kako se evolucija mreže interakcija menja tokom ovih različitih faza i da li se na osnovu promene strukture mreže interakcija mogu predvideti trenuci kada sistem prelazi iz jedne u drugu fazu.

Jedna od važnih strukturnih osobina kompleksnih mreža je njihova simetričnost u odnosu na permutacije koje ne menjaju matricu povezanosti. U radu [5] je dat pregled različitih metoda za ispitivanje simetrije u grafovima među kojima se navodi i grupa automorfizama grafa kao jedan od metoda. Zastupljenost simetrije u realnim sistemima, određivanje uzroka simetrije i poređenje uzroka među realnim sistemima koristeći metod grupe automorfizama grafa je prikazan od radu [6]. U ovom radu je pokazan određen stepen univerzalnosti velike zastupljenosti simetrije u realnim sistemima kroz potvrđivanje u različitim realnim sistemima kao što je genetska struktura ljudske T ćelije, internet mreža i elektroenergetska mreža. U radu [7] je grupa automorfizama korišćena za ispitivanje simetrije u međunarodnoj trgovinskoj mreži.

U ovom radu su korišćene metode koje se oslanjaju na algebru, kombinatoriku i teoriju grafova kako bi se analizirala simetričnost mreže interakcija u ekonomskom sistemu. Simetričnost predstavlja nepromenljivost sistema u odnosu na određenu transformaciju koja u ovom slučaju znači permutaciju čvorova ali tako da se susednost među čvorovima ne promeni. Za ispitivanje simetričnosti koristili smo grupe automorfizama grafa, pri čemu je veličina grupe automorfizama pokazatelj stepena simetrije u ekonomskom sistemu.

U ovom radu posmatrali smo ekonomski sistem sačinjen od američkih kompanija koje se bave finansijskim poslovima kao što je bankarstvo, osiguranje, investicioni fondovi. Struktura ekonomskog sistema je posmatrana kao graf dobijen na osnovu korelacija između vremenskih serija cena akcija ovih kompanija na berzi. Posmatrane su mreže za tri različite godine koje odgovaraju periodu pre krize, 2005, godina krize, 2008, i godina posle krize kada je krenuo oporavak ekonomskog sistema, 2015. Za svaku od ove tri mreže odredili smo veličine grupe automorfizama. Rezultati ovog rada potvrđuju veliku zastupljenost simetrije u posmatranom ekonomskom sistemu. Pored toga, pokazali smo da je ekonomski sistem u kriznom periodu okarakterisan smanjenom simetričnošću u odnosu na period pre i posla krize.

Ovaj rad je organizovan na sledeći način. U sekciji 2 prikazan je pregled definicija i korišćene metodologije. U sekciji 3 su prikazani rezultati analize, dok su zaključci rada prikazani u sekciji 4.

## 2. METODOLOGIJA

### 2.1. Definicija grafa

Prema definiciji D. Cvetkovića i S. Simića iz [8], ukoliko je  $v$  neprazan skup i  $e$  binarna relacija u  $v$ , uređen par  $G = (V, E)$  se naziva graf. Čvorovi grafa su elementi skupa  $V$ , dok su veze elementi skupa  $E$ . Graf je neorjentisan ili simetričan ukoliko je  $E$  simetrična relacija, a to znači da je  $xEy = yEx$ . U ovom radu ćemo raditi sa neorjentisanim grafovima i zvaćemo ih prosto grafovima imajući u vidu da se podrazumeva neorjentisani graf. Broj elemenata skupa  $V$ , odnosno broj čvorova ćemo obeležavati sa  $N$ , a broj elemenata skupa  $E$ , odnosno broj veza sa  $L$ . Za graf ćemo koristiti oznaku  $G$ .

Sused čvora  $i$  je čvor spojen jednom vezom sa čvorom  $i$ . Broj suseda se označava kao stepen čvora  $i$ . Podgraf  $g(v, e)$  grafa  $G$  se definiše kao graf koji sadrži podskup čvorova grafa  $G$  dok veze između tih čvorova odgovaraju vezama u grafu  $G$  i podskup su skupa  $E$ , odnosno  $v \subset V$  i  $e \subset E$ . Za podgraf  $g$  grafa  $G$  se još kaže da je to podgraf koji je indukovano skupom čvorova  $n$ , gde su sve veze između čvorova  $n$  u grafu  $G$ , ne uključujući veze sa čvorovima van tog skupa, sadržane u podgrafu  $g$ .

Put koji povezuje čvor  $i$  sa čvorom  $j$  je niz susednih veza između ta dva čvora, tako da prva veza u nizu počinje u čvoru  $i$ , dok se poslednja veza u nizu završava u čvoru  $j$ . Susedne veze su one koje imaju zajednički čvor. Ukoliko se kroz svaki čvor u putu ne prolazi više od jednom, taj put se naziva elementarni put. Najkraći put između dva čvora u grafu je onaj koji sadrži najmanji broj veza.

Za graf se kaže da je povezan ukoliko se svaka dva čvora mogu povezati putem, u suprotnom se kaže da je nepovezan. Nepovezan graf se sastoji od komponenti povezanosti, odvojenih delova grafa, koje predstavljaju podgrafove u okviru kojih se svaka dva čvora mogu spojiti putem.

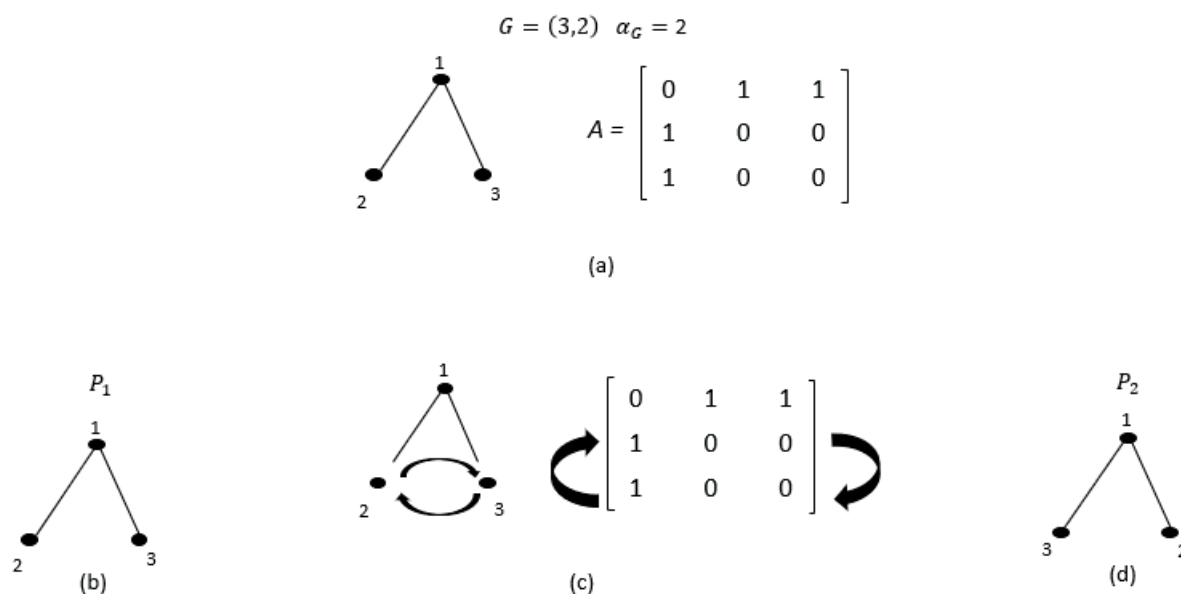
Za predstavljanje grafa se koristi matrica susedstva,  $A$ , koja jednoznačno određuje graf. Matrica susedstva je kvadratna matrica gde su oznake vrsta i kolona jednake oznakama čvorova. Elementi matrice susedstva  $a_{i,j}$  mogu biti jednaki 0 ili 1 u zavisnosti da li postoji veza između čvorova  $i$  i  $j$ .

### 2.2. Izomorfizam grafova. Automorfizam i grupe automorfizama

**Definicija 1.** Dva grafa su izomorfna ako postoji uzajamno jednoznačno preslikavanje skupa njihovih čvorova koje održava susednosti čvorova.

Dakle, za proizvoljna dva grafa  $G_1 = (V_1, E_1)$  i  $G_2 = (V_2, E_2)$  kažemo da su izomorfni ako i samo ako postoji bijektivno preslikavanje  $\varphi$  skupa  $V_1$  na skup  $V_2$  za koje važi

$$(\forall a, b \in N_1)((a, b) \in L_1 \Leftrightarrow (\varphi(a), \varphi(b)) \in L_2)$$



**Slika 1** Primer grupe automorfizama jednostavnog grafa sa tri čvorova

Ukoliko posmatramo izomorfizam grafa  $G$  na samog sebe, odnosno izvršimo permutaciju čvorova tako da zadržimo osobine navedne u definicijama više, dobijamo graf koji se naziva automorfizam grafa  $G$ . Dakle, uočavamo da automorfizam predstavlja permutaciju čvorova u grafu  $G$ . Skup svih automorfizama predstavlja grupu koja se naziva grupa automorfizama grafa  $G$  i označavamo je sa  $Aut(G)$ . Osobina koju ispitujemo u ovom radu jeste veličina grupe automorfizama odnosno broj svih permutacija grafova tako da se zadržava određena osobina a to je matrica susedstva. Veličinu grupe označavamo sa  $\alpha_G$ . U sledećem pasusu je dat razlog zašto zadržavamo matricu susedstva a ne samo strukturu grafa.

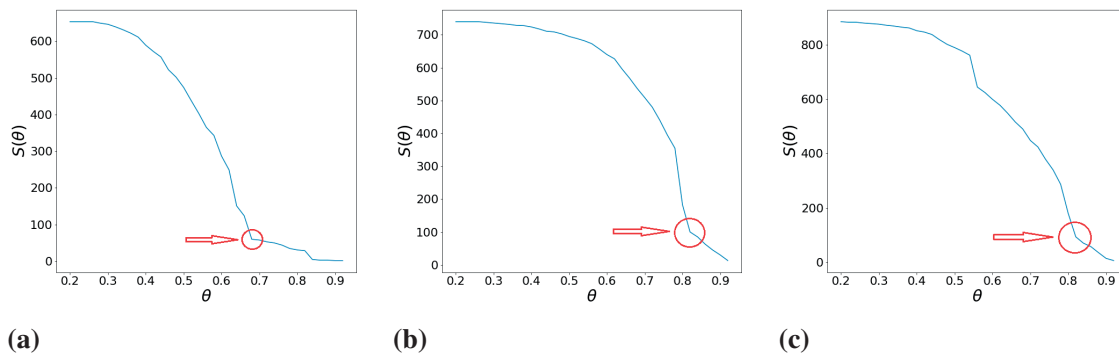
Posmatrajući automorfizme pri kojima se samo struktura grafa održava, a matrice susedstva mogu biti različite, je isto što i promena oznaka čvorova. Broj takvih izomorfizama je  $N!$ , gde je  $N$  broj čvorova grafa i isti je za svaki graf sa  $N$  čvorova. Dakle, veličina grupe u tom slučaju zavisi samo od broja čvorova i ne zavisi od strukture grafa. Međutim, u ovom radu se ispituje prisutnost simetrije u realnim sistemima. Ta simetrija je posmatrana kroz broj permutacija čvorova odnosno veličinu grupe automorfizama tako da se matrica susedstva ne menja. U ovom slučaju veličina grupe automorfizama zavisi od strukture grafa, ne samo od broja čvorova, i razlikuje se za svaki graf. Ono što je zanimljivo jeste da je simetrija, definisana na ovaj način, zastupljena u realnim sistemima. U ovom radu se takođe ispituje prisutnost simetrije u ekonomskom sistemu kao realnom sistemu.

### 2.3. Grupe automorfizama grafova sa malim brojem čvorova

Automorfizme jednog grafa dobijamo tako što fiksiramo oznake čvorova, a zatim menjamo mesta čvorovima u grafu tako da se ne promeni matrica susedstva. Ovo je ekvivalentno zameni mesta vrstama ili kolonama u matrici susedstva koji odgovaraju tim čvorovima tako da je krajnji rezultat nepromenjena matrica susedstva. Na jednostavnom primeru na slici 1 grafa od tri čvorova ćemo prikazati grupu automorfizama.

Na slici 1a je prikazan originalni graf  $G$  i njegova matrica susedstva. Veličina grupe automorfizama  $Aut(G)$  je jednaka 2 tj.  $\alpha_G = 2$ . Dakle, grupa sadrži dve permutacije koje imaju istu matricu susedstva kao originalni graf. Na slici 1b je prikazana jedna takva permutacija koja predstavlja identičko preslikavanje. Druga permutacija je slikovito prikazana na slici 1c. Dakle, vidimo da automorfizam odnosno permutaciju možemo posmatrati kao da otkačinjemo čvorove iz svoje pozicije i kačimo ih na mesto čvorova sa kojim vršimo zamenu. Takođe se na istoj slici vidi da se zamenom vrsta koje označavaju čvorove u matrici susedstva ne menja matrica susedstva. Na slici 1d je prikazana rezultujuća druga permutacija u oznaci  $P_2$  kao element grupe automorfizama  $Aut(G)$ . Dakle, naša grupa je  $Aut(G) = \{P_1, P_2\}$ .

Primetimo da postoje čvorovi koji ne učestvuju u permutaciji odnosno ne menjaju svoju poziciju ni u



**Slika 2** Zavisnost broja čvorova u najvećoj komponenti povezanosti u grafu,  $S$ , od praga,  $\theta$ . Slike (a), (b) i (c) prikazuju zavisnost za godine 2005, 2008 i 2015, respektivno.

jednom automorfizmu u okviru grupe. Takav čvor je čvor sa oznakom 1. Prema radu [7] razlikuje se lokalna i globalna simetrija prema broju čvorova koji učestvuju u permutaciji odnosno automorfizmu. Naime, ukoliko je automorfizam posledica zamene mesta čvorova kao u primeru na slici 1, onda se radi o lokalnom automorfizmu jer samo dva čvorova učestvuju u permutaciji. Međutim, postoje permutacije za koje je neophodno permutovati istovremeno sve čvorove kako bi se dobio automorfizam. Takva vrsta automorfizama se naziva globalna simetrija.

### 3. REZULTATI

#### 3.1. Opis mreža

Podaci koji su korišćeni u radu predstavljaju vremenske serije dnevnih cena deonica kompanija posmatranih u periodu od godinu dana. Odabrane su tri godine: 2005, 2008 i 2015. Preuzeti su sa sajta yahoo/finance i javno su dostupni. Kompanije su iz finansijskog sektora koje su prisutne na njujorškoj berzi i ima ih ukupno 653 za godinu 2005, 740 za godinu 2008 i 884 za godinu 2015. Poslove obavljaju u Severnoj Americi, a obuhvataju sledeće kategorije kompanija: investicione banke, banke koje posluju sa stanovništvom, fondovi koji investiraju u hartije od vrednosti, fondovi koji investiraju u nepokretnosti, lokalni investicioni fondovi. Dnevne cene predstavljaju cene koje su ostvarene na kraju radnog vremena berze i koje su korigovane tako da se ukloni efekat razvodnjavanja vrednosti zbog prodaje novih deonica, kao i zbog isplate dividendi.

S obzirom da preuzeti podaci sadrže trend, uklonjen je trend u svakoj vremenskoj seriji kako bi se dobile vremenske serije koje predstavljaju fluktuaciju kao posledicu pojedinačnog ponašanja kompanije, a ne posledicu zajedničkog faktora prisutnog kod svih kompanija. U cilju uklanjanja trenda korišćen je metod opisan u [9]. Na osnovu detrendovanih vremenskih serija, izračunati su korelacioni koeficijenti kao u radu [10]. Na ovaj način je dobijena korelaciona matrica  $\hat{\rho}_{X,Y}$ . Kako bismo dobili matricu susedstva koja jednoznačno određuje graf, primenjujemo metod praga na osnovu koga veza između dva čvorova postoji ukoliko je absolutna vrednost ocenjenog korelacionog koeficijenta između ta dva čvorova veća od zadatog praga. Element  $A_{i,j}$  matrice susedstva dobijamo na sledeći način:

$$A_{i,j} = \begin{cases} 1 & \text{ako važi } |\hat{\rho}_{i,j}| > \theta \\ 0 & \text{ako važi } |\hat{\rho}_{i,j}| \leq \theta \end{cases} \quad (1)$$

gde je  $\theta$  vrednost praga. Vrednost praga  $\theta$  određujemo tako što posmatramo zavisnost broja čvorova u najvećoj komponenti povezanosti u grafu,  $S$ , od vrednosti praga. Najveća komponenta povezanosti je komponenta povezanosti koja sadrži najveći broj čvorova [11]. Zavisnost je prikazana na slici 2, gde je na slici 2a prikazana zavisnost koja odgovara godini 2005, na slici 2b godini 2008 i na slici 2c godini 2015.

Za vrednost  $\theta = 0$ , posmatrana mreža je potpuno povezan graf sa vezama između svih čvorova. Kako povećavamo vrednost praga  $\theta$  tako odbacujemo sve veći broj veza, a dobijena mreža u jednom trenutku postaje nepovezana, odnosno sadrži više od jedne komponente. Najveća od tih komponenti sadrži najveći procenat čvorova u mreži. Kako povećavamo  $\theta$  tako i veličina najveće komponente  $S$  opada, sve dok njena veličina ne

postane slična veličinama ostalih komponenti. Tačke u kojima se ovo dešava su obeležene crvenom bojom na slici 2. Komponente koje se dobijaju u navedenim tačkama se odlikuju istom ekonomskom delatnošću. Prema tome, najjače veze u grafu, odnosno veze koje imaju najveći koeficijent korelacije jesu one unutar komponenti. Dakle, za navedene tačke praga dobijamo graf koji predstavlja jezgro ekonomskog sistema okarakterisanom jasno razdvojenim ekonomskim kategorijama kroz izdvajanje komponenti. Drugim rečima, dobijamo kostur na koji ukoliko želimo, dalje dodajemo veze unutar i van zajednica smanjivanjem praga i uključivanjem slabijih veza u sistem.

Međutim, mi u ovom radu hoćemo da ispitamo da li postoji simetrija u jezgru ekonomskog sistema. A upravo navedeni pragovi predstavljaju minimalne vrednosti  $\theta$  za koje se dobija maksimalni graf koje predstavlja jezgro odnosno dobijamo jezgro kod koga nisu isključene veze unutar komponenti. Baš te tačke su izabrane jer ukoliko bismo smanjivali dalje prag, uključili bismo veze van komponenti. U slučaju da povećamo prag, izgubili smo veze odnosno informacije iz jezgra.

U tabeli 1: prikazani su rezultati veličina grupa automorfizama za grafove dobijenih za godine 2005, 2008, i 2015. Prema dobijenim rezultatima možemo najpre uočiti veliku zastupljenost simetrije u ekonomskom sistemu u svakoj godini na osnovu veličine grupe automorfizama.

Daljim istraživanjem bi se trebalo ispitati uzrok velike zastupljenosti simetrije u ekonomskom sistemu. U radu [6] je predstavljeno određivanje generatora simetrije u realnim sistemima. Drugim rečima, ispitati da li se radi o lokalnoj ili globalnoj simetriji. Ukoliko su lokalne simetrične strukture glavni uzrok velike zastupljenosti simetrije u ekonomskom sistemu, potrebno je uočiti koje su to strukture i koliko različitih takvih struktura postoji u sistemu. Na primer, ukoliko se ispostavi da je struktura zvezde pristutna u velikoj meri i predstavlja glavni generator simetrije u sistemu, to bi značilo da se jezgro ekonomskog sistema formira tako što se ista lokalna struktura ponavlja u celom sistemu, čime bismo nastanak ili evoluciju ekonomskog sistema sveli na formiranje lokalne strukture zvezde. S obzirom da struktura zvezde predstavlja strukturu vođe i njegovih sledbenika, to bi moglo da znači da se ekonomski sistem formira tako što se oko vođa u sistemu vezuju sledbenici, odnosno isti mikro proces se ponavlja na ceo sistem.

Empirijski graf - godina i prag	Broj čvorova	Broj veza	Veličina grupe automorfizama
Godina 2005, prag 0.68	185	973	$7.58 \times 10^{22}$
Godina 2008, prag 0.82	286	2,092	$1.46 \times 10^{14}$
Godina 2015, prag 0.82	290	1,494	$1.35 \times 10^{30}$

**Tabela 1:** Grupe automorfizama empirijskog grafa u godinama 2005, 2008 i 2015 i za prag  $\theta$  koji rezultuje u jezgru ekonomskog sistema

Drugi rezultat ovog rada se odnosi na smanjenu simetriju u ekonomskom sistemu u kriznom periodu koja se može uočiti u tabeli 1: Naime, vidimo da je veličina grupe automorfizama najmanja u godini 2008 u odnosu na godine pre krize i posle krize. Iako je poređenje empirijskih grafova otežano zbog različitih veličina grafova, odnosno zbog različitog broja čvorova, smanjena simetrija se može tvrditi za godinu 2008 iz sledećih razloga: iako u godini 2005 postoji manji broj čvorova što implicira manji potencijal za simetriju, veličina grupe automorfizama je veća. Dakle, ekonomski sistem u godini 2005 jeste simetričniji u odnosu na godinu 2008. Takođe, u godini 2015 broj čvorova je sličan broju čvorova u godini 2008, međutim veličina grupe automorfizama je značajno veća. Što se tiče tumačenja rezultata smanjenje simetrije u kriznom periodu, takođe je, kao i kod prvog rezultata ovog rada, neophodno sprovesti dalje istraživanje u cilju uvida u generatore simetrije ekonomskog sistema.

Empirijski graf - godina i prag	Broj čvorova	Broj veza	Veličina grupe automorfizama
Godina 2005, prag 0.66	211	1,211	$2.93 \times 10^{18}$
Godina 2008, prag 0.78	368	4,154	$2.2 \times 10^6$
Godina 2015, prag 0.78	401	3,279	$7.28 \times 10^{35}$

**Tabela 2:** Grupe automorfizama empirijskog grafa u godinama 2005, 2008 i 2015 i za maksimalni prag  $\theta$ .

U cilju provere da li bi se simetrija zadržala ukoliko bismo smanjili prag tako da uključivanjem slabijih veza ostvarimo povezivanje komponenti, rezultati veličine grupe automorfizama su prikazani u tabeli 2:.

Smanjivanjem praga, uključili smo više čvorova u sistem, uključili smo veze između komponenti, koje su sada postale zajednice jer su se povezale sa najvećom komponentom, i povećali broj veza unutar zajednica. Prema podacima u tabeli 2:, oba dobijena rezultata prethodno navedena, velika zastupljenost simetrije i smanjivanje simetrije u kriznom periodu, su i ovde potvrđena.

#### 4. ZAKLJUČAK

U ovom radu je korišćena teorija grafova za analizu simetrije u mreži interakcija realnog ekonomskog sistema. Simetrija je ispitivana kroz analizu automorfizama realnog grafa. Veličina grupe automorfizama je korišćena za ispitivanje nivoa zastupljenosti simetrije u sistemu. Automorfizam kao permutacija jeste bijektivno preslikavanje, te mi ispitujemo do koje mere je sistem nepromenljiv u odnosu na primenjenu transformaciju. Ovo nam je značajno jer nam pomaže u razumevanju nastanka strukture realnih sistema. Na primer, ukoliko je utvrđena velika zastupljenost simetrije u sistemu i uzrok tome je mirkostuktura sa malim brojem čvorova koja se ponavlja kroz ceo sistem, nastanak strukture realnog sistema se može svesti na nastanak te mikrostrukture.

Ispitivanje simetrije je sprovedeno na ekonomskom sistemu kao predstavniku realnog sistema. Pokazano je da mreže interakcija ekonomskog sistema imaju visok stepen simetrije. U naučnoj literaturi je potvrđena pristutnost simetrije u nekoliko različitih realnih sistema kao što je genetska struktura ljudske T ćelije, internet mreže i elektro-energetske mreže, te se doprinos ovog rada ogleda u potvrđivanju univerzalnosti zastupljenosti simetrije u realnim sistemima. Takođe je uočeno da se pri promeni stanja ekonomskog sistema menja i stepen zastupljenosti simetrije u sistemu odnosno da je u kriznom periodu ekonomski sistem manje simetričan nego van kriznog perioda. Dalja istraživanja u ovom smeru uključuju ispitivanje uzroka simetrije u ekonomskom sistemu. Na ovaj način bi se uočili generatori simetrije u ekonomskim sistemima, koji mogu biti lokalni i globalni, kao i stepen njihovog učešća u stvaranju simetrije sistema. Njihovom spoznajom bi se doprinelo boljem razumevanju evolucije ekonomskog sistema.

#### LITERATURA

- [1] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4-5), 175-308.
- [2] Barabási, A.-L. (2016). *Network science*. Cambridge university press.
- [3] Cvetković, D. and Simić, S. (1990). *Diskretna matematika: matematika za kompjuterske nauke*. Naučna knjiga.
- [4] Stiglitz, J. E. (2010). *Freefall: America, free markets, and the sinking of the world economy*. WW Norton & Company.
- [5] Garlaschelli, D., Ruzzenenti, F., and Basosi, R. (2010). Complex networks and symmetry i: A review. *Symmetry*, 2(3), 1683-1709.
- [6] MacArthur, B. D., Sánchez-García, R. J., and Anderson, J. W. (2007). On automorphism groups of networks. *arXiv preprint arXiv:0705.3215*.
- [7] Wang, H., Yan, G., and Xiao, Y. (2009). Symmetry in world trade network. *Journal of Systems Science and Complexity*, 22(2), 280-290.
- [8] Cvetkovic, D. and Simic, S. (2002). *Odabrana poglavlja iz diskretne matematike (chosen chapters from discrete mathematics)*. Akademska misao, Belgrade.
- [9] Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L. (1994). Mosaic organization of dna nucleotides. *Physical Review E*, 49(2), 1685.
- [10] Mitrović Dankulov, M., Tadić, B., and Melnik, R. Analysis of worldwide time-series data reveals some universal patterns of evolution of the sars-cov-2 pandemic. *Frontiers in Physics* 10, 936618.
- [11] Živković, J., Tadić, B., Wick, N., and Thurner, S. (2006). Statistical indicators of collective behavior and functional clusters in gene networks of yeast. *The European Physical Journal B-Condensed Matter and Complex Systems*, 50(1-2), 255–258.



## METRIC DIMENSION OF COMPLETE SPLIT GRAPHS

JOZEF KRATICA<sup>1</sup>, MILICA MILIVOJEVIĆ DANAS<sup>2</sup>

<sup>1</sup> Mathematical Institute - Academy of Sciences and Arts, jkratica@mi.sanu.bg.ac.rs

<sup>2</sup> University of Kragujevac - Faculty of Science and Mathematics, milica.milivojevic@kg.ac.rs

**Abstract:** In this paper, the problem of determining the metric dimension for special class of graphs, named complete split graphs  $K^*_{k,n-k}$  is considered. It is stated and proved formula for the metric dimension of this graphs: for  $n - k \geq 2$  and  $k \geq 2$ , as well as for  $k = 1$  and  $n \geq 3$ , metric dimension of  $K^*_{k,n-k}$  is equal to  $n - 2$ , otherwise metric dimension of  $K^*_{k,n-k}$  is equal to  $n - 1$ .

**Keywords:** metric dimension, complete split graph, graph theory, discrete mathematics

### 1. INTRODUCTION

The metric dimension problem was introduced in the seventies by Slater in [1] and Harary and Melter in [2], independently of one another. This NP-hard graph invariant [3] has been very much researched the last five decades and it has applications in many diverse areas.

Let  $G = (V, E)$  be a simple connected undirected graph with vertex set  $V$  and edge set  $E$ . The distance between vertices  $u$  and  $v$ , denoted  $d(u, v)$ , is the length of a shortest  $u - v$  path in graph  $G$ . Then, it can be said that a vertex  $w$  resolves two vertices  $u, v \in V$  if  $d(u, w) \neq d(v, w)$ . A set  $S = \{w_1, \dots, w_k\}$  subset of  $V$  is named a resolving set of graph  $G$ , if every two distinct vertices from  $V$  are resolved by some vertex of  $S$ . A metric basis for graph  $G$  is a resolving set of minimal cardinality. The cardinality of a metric basis for  $G$  is called the metric dimension and is denoted by  $\beta(G)$ . Its applications are in several diverse areas. Applications to the direction of robots in networks are analyzed in [3] and applications to chemistry in [4] and [5], among others.

There are several other variations of metric dimension in the literature, but the motivation for this paper was an already known mixed metric dimension for complete split graph  $K^*_{k,n-k}$ . First, the distance between edge  $uv$  and vertex  $w$  is defined as  $d(uv, w) = \min\{d(u, w), d(v, w)\}$ . Second, similarly as previous, mixed metric dimension is defined in [6] as minimal cardinality of mixed resolving set for graph  $G$ , where every two items (item is vertex or edge) are resolved by some vertex from mixed resolving set. Third, it is easy to see that  $B_M(G) \geq \beta(G)$ .

In [7] exact value for mixed metric dimension of the complete split graphs is found with corresponding proof, i.e. mixed metric dimension for this graphs is equal  $n - 1$  for  $k = 1$  and  $n \geq 3$ , otherwise it is equal to its order  $n$ , i.e.  $\beta_M(K^*_{k,n-k}) = \begin{cases} n - 1, & k = 1 \wedge n \geq 3 \\ n, & \text{otherwise} \end{cases}$ .

Now, some theoretical properties of metric dimension known from literature will be presented, which are used in the next section.

**Theorem 1.** [8] A connected graph  $G$  of order  $n$  has dimension 1 if and only if  $G = P_n$ .

**Theorem 2.** [8] Let  $G$  be a connected graph of order  $n \geq 4$ . Then  $\beta(G) = n - 2$  if and only if  $G = K_{s,t}$  ( $s, t \geq 1$ ),  $G = K_s + \overline{K}_t$  ( $s \geq 1, t \geq 2$ ) or  $G = K_s + (K_1 \cup K_t)$  ( $s, t \geq 1$ ).

Let's note that the graph  $G = K_s + \overline{K}_t$  is the join of a complete graph and an empty graph.

**Theorem 3.** [8] A connected graph  $G$  of order  $n \geq 2$  has dimension  $n - 1$  if and only if  $G = K_n$ .

## 2. METRIC DIMENSION OF COMPLETE SPLIT GRAPHS

In this paper the metric dimension will be studied for an arbitrary complete split graph. These graphs are denoted by  $K^*_{k,n-k}$  and have  $n$  vertices and  $\binom{n}{2} - \binom{n-k}{2}$  edges, where vertex set is  $V(K^*_{k,n-k}) = \{x_i | 1 \leq i \leq k\} \cup \{y_i | 1 \leq i \leq n-k\}$ , while edge set is defined as  $E(K^*_{k,n-k}) = \{x_i x_j | 1 \leq i < j \leq k\} \cup \{x_i y_j | 1 \leq i \leq k; 1 \leq j \leq n-k\}$ . It should be noted that complete split graph  $K^*_{k,n-k}$  can be viewed as a join of a complete graph  $K_k$  and a graph without edges  $\bar{K}_{n-k}$ . Throughout the paper we will assume that  $V_1 = \{x_i | 1 \leq i \leq k\}$  and  $V_2 = \{y_i | 1 \leq i \leq n-k\}$ .

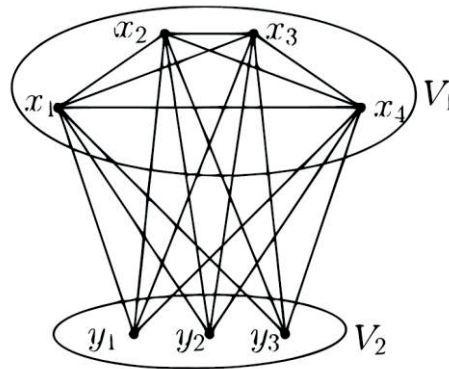


Figure 1: Graph  $K^*_{4,3}$

In the Figure 1 is presented the complete split graph  $K^*_{4,3}$ . Its metric dimension is 5, i.e.  $\beta(K^*_{4,3}) = 5$  which is obtained by total enumeration. One metric basis of  $K^*_{4,3}$  is set  $S = \{x_1, x_2, x_3, y_1, y_2\}$ . Metric coordinate of all vertices, with respect to  $S$ , are presented in Table 1 and as it can be seen all vertices have mutually different metric coordinates.

Table 1: Metric coordinates of vertices of  $K^*_{4,3}$  with respect to  $S$

Type	Item	Metric coordinates
vertex	$x_1$	(0, 1, 1, 1, 1)
	$x_2$	(1, 0, 1, 1, 1)
	$x_3$	(1, 1, 0, 1, 1)
	$x_4$	(1, 1, 1, 1, 1)
	$y_1$	(1, 1, 1, 0, 2)
	$y_2$	(1, 1, 1, 2, 0)
	$y_3$	(1, 1, 1, 2, 2)

In the next Theorem 4 it will be obtained and proved metric dimension of any complete split graph  $K^*_{k,n-k}$ .

**Theorem 4.**  $\beta(K^*_{k,n-k}) = \begin{cases} n-2, & (k = 1 \wedge n \geq 3) \vee (n-k \geq 2 \wedge k \geq 2) \\ n-1, & \text{otherwise} \end{cases}$ .

**Proof.** Several cases will be discussed:

- $n = 2 \wedge k = 1$ . It is easy to see that  $K^*_{1,1} \cong P_2$ . By Theorem 1, we have the following result  $\beta(K^*_{1,1}) = \beta(P_2) = 1$ ;
- $n = 3 \wedge k = 1$ . In a similar way to previous case, it is easy to see that  $K^*_{1,2} \cong P_3$ . By Theorem 1, we have the following result  $\beta(K^*_{1,2}) = \beta(P_3) = 1$ ;

- $n > 3 \wedge k = 1$ . Since  $K_{1,n-1}^* \cong K_{1,n-1}$ , then by Theorem 2, the metric dimension of the complete split graph  $K_{1,n-1}^*$  of order  $n$  is  $n - 2$ , i.e.  $\beta(K_{1,n-1}^*) = n - 2$ ;
- $n - k = 1 \wedge k \geq 2$ . Since  $K_{n-1,1}^* \cong K_n$ , then by Theorem 3, the metric dimension of the complete split graph  $K_{n-1,1}^*$  of order  $n$  is  $n - 1$ , i.e.  $\beta(K_{n-1,1}^*) = \beta(K_n) = n - 1$ ;
- $k \geq 2$  and  $n - k \geq 2$ . Since graph  $K_{k,n-k}^*$  is a connected graph of order  $n \geq 4$ , then by Theorem 2 it follows  $\beta(K_{k,n-k}^*) = n - 2$ .

□

It is interesting to compare metric dimension and mixed metric dimension for these graphs. As it can be seen above, mixed metric dimension and metric dimension differ by two in case of  $n - k \geq 2$  and  $k \geq 2$ . Otherwise, mixed metric dimension and metric dimension differ by one.

### 3. CONCLUSIONS

In this paper, the metric dimension for complete split graphs is considered. It is proved that metric dimension is  $n - 2$  for  $n - k \geq 2$  and  $k \geq 2$ , as well as for  $k = 1$  and  $n \geq 3$ . Otherwise it is equal to  $n - 1$ . Further work can be directed in finding edge metric dimension of complete split graphs, as well as, some other interesting classes of graphs.

### REFERENCES

- [1] Slater, P. J. (1975). Leaves of trees. *Proceeding of the 6th Southeastern Conference on Combinatorics, Graph Theory and Computing, Congressus Numerantium*, 14, 549-559.
- [2] Harary, F., Melter, R. (1976). *On the metric dimension of a graph*. *Ars Combinatoria*, 2, 191-195.
- [3] Khuller, S., Raghavachari, B., Rosenfeld, A. (1996). *Landmarks in graphs*. *Discrete Applied Mathematics*, 70, 217-229.
- [4] Johnson, M. (2007). *Structure-activity maps for visualizing the graph variables arising in drug design*, *Journal of Biopharmaceutical Statistics*, 3, 203-236.
- [5] Chartrand, G., Poisson, C., Zhang, P. (2000). *Resolvability and the upper dimension of graphs*, *Computers and Mathematics with Applications*, 39, 19-28.
- [6] Kelenc, A., Kuziak, D., Taranenko, A., Yero, I. (2017). *Mixed metric dimension of graphs*, *Applied Mathematics and Computation*, 314, 429-438.
- [7] Milivojević Danas, M., Maksimović, Z., Kratica, J., Savić, A. (2020). *Mixed metric dimension of complete split graphs*. XLVII Simpozijum o operacionim istraživanjima, Kraljevo, 20.-23.09, Zbornik radova, 101-104.
- [8] Chartrand, G., Eroh, L., Johnson, M. A., Oellermann, O. R. (2000). *Resolvability in graphs and the metric dimension of a graph*. *Discrete Applied Mathematics*, 105, 99 -113.



## OCCURRENCE OF FIBONACCI AND LUCAS SEQUENCES IN COUNTING RESTRICTED PERMUTATIONS

VLADIMIR BALTIĆ<sup>1</sup>

<sup>1</sup> ATUSS Belgrade, balticvladimir@gmail.com

**Abstract:** In this paper we will count the number of strongly restricted permutations using finite state automata (type of digraph). We will describe the construction of the corresponding automaton. We will illustrate our technique through examples. We will establish the connections between some kind of restricted permutations and Fibonacci, Tribonacci and Lucas sequences.

**Keywords:** Finite state automata, restricted permutations, exact enumeration, recurrences, Fibonacci sequence, Lucas sequence

### 1. INTRODUCTION

At the beginning of this paper we will remind on the basic concepts.

A **finite state automaton**  $M$  is  $M = (T, S, Y, s_0, F)$ , where we define each of the five parts as follows:

- 1) a finite set (alphabet)  $T$  of inputs;
- 2) a finite set  $S$  of (internal) states;
- 3) a subset  $Y$  of  $S$  (final, accepting or “yes” states);
- 4) an initial state (initial or start state)  $s_0$  in  $S$ ;
- 5) a next-state function  $F$  from  $S \times T$  into  $S$ .

An **accessible state** is any state that can be reached from the start state. A **sink state** is any state for which  $f(s, x) = s$  holds for all  $x \in T$ .

The nondeterministic finite automaton is a variant of finite automaton with following characteristic: zero or more than one possible value may exist for state transition (in the deterministic finite automaton the next possible state is uniquely determined).

In [5] you can find more about finite state automata.

Let  $p$  be a permutation of the set  $N_n = \{1, 2, \dots, n\}$ .

A **restricted permutation** is permutation  $p$  in which the positions of the marks after the permutation are restricted. A restricted permutation  $p$  can be specified by an  $n \times n$  zero-one matrix  $A = (a_{ij})$  in which holds:

$a_{ij} = 1$ , if the mark  $j$  is permitted to occupy the  $i$ -th place;

$a_{ij} = 0$ , otherwise.

We can say that it is allowed to be  $p(i) = j$  if and only if  $a_{ij} = 1$ .

We introduce our technique based on finite state automata for counting the number of strongly restricted permutations of  $N_n$  satisfying the condition  $p(i) - i \in I$  (for some set  $I$ ). For the history of this kind of a problems see [1].

Whenever a larger integer precedes a smaller one in a permutation we say that we have an **inversion**. A permutation is called **odd** if the total number of inversions is an odd integer. A permutation is called **even** if the total number of inversions is an even integer.

**Fibonacci sequence** is given with the recurrence:  $F(1) = 1, F(2) = 1, F(n+2) = F(n+1) + F(n)$ .

**Lucas sequence** is given with the recurrence:  $L(1) = 1, L(2) = 3, L(n+2) = L(n+1) + L(n)$ .

From Theorem 2.4.29 in [6] we have:  $F(n-1) + F(n+1) = L(n)$ .

For more properties of these sequences see [6].

## 2. CONSTRUCTION OF FINITE STATE AUTOMATON

For the completeness of this paper we will introduce construction of finite state automaton for counting restricted permutations. For the proofs see [4].

First we will consider restricted permutations satisfying the condition  $-k \leq p(i)-i \leq r$  ( $k$  and  $r$  are arbitrary natural numbers). We notice that choosing the  $i$ -th element of the permutation,  $p(i)$ , depends only on the status of the element  $i$  and preceding  $k$  elements and following  $r$  elements of the permutation under construction (status can be used or unused).

In the case of odd or even permutations states consist of the ordered pairs. Otherwise it is equal to the first coordinate in the case of odd or even permutations.

The first coordinate consist of  $k+1+r$  slots corresponding to status (used/filled or unused/vacant) of the preceding  $r$  elements, momentary element and following  $k$  elements of the permutation under construction. We will denote filled slots by 1 and vacant by 0. Some slots have already been filled and we need to fill one more, then drop the first slot off to give a new state. However, if the first slot is unused, that is the last chance to fill it.

The second coordinate controls the parity of the restricted permutation in the case of odd or even permutations. It is 1 if the permutation is odd, and 0 if it is even.

There is one more sink state, denoted by  $Q$ , corresponding to impossible state for this kind of the restricted permutations.

In this paper we will omit the proof of the Theorem 1, which describes the finite state automaton for counting even restricted permutations (with slightly modifications we construct the finite state automaton for counting odd restricted permutations). You can find the proof in [4].

### Theorem 1.

The set of the internal states  $S$  consists of the states that have in the first coordinate exactly  $k$  ones and one more sink state, denoted by  $Q$ , corresponding to impossible state for this kind of the restricted permutations.

The alphabet  $T$  is given by:

$$T = \{-k, -k+1, \dots, -2, -1, 0, 1, 2, \dots, r-1, r\}.$$

When first  $k$  slots in the first coordinate are 1 (in this case they made one permutation) and 0 in the second coordinate. This is only one "yes" state and also it is the start state.

The next-state function (for  $-k \leq a \leq r$ ) is defined by:

$$F(S_1, a) = Q, \text{ if the first slot is vacant and } a \neq -k;$$

$$F(S_1, a) = Q, \text{ if the } (k+1+a)\text{-th slot is filled};$$

$$F(S_1, a) = S_2, \text{ if the first slot is filled and } (k+1+a)\text{-th slot is vacant.}$$

First coordinate of the state  $S_2$  is obtained from the state  $S_1$  when we put 1 at the  $(k+1+a)$ -th slot, then drop the first slot off and put one 0 (empty slot) at the end. Second coordinate of the state  $S_2$  remains same if we have even number of ones in the positions after  $(k+1+a)$ -th slot and changes (from 0 to 1 and vice-versa) if that number is odd.

Only restricted permutations satisfying the condition  $-k \leq p(i)-i \leq r$ , which are even permutations, are recognized by the finite state automaton  $M = (T, S, Y, S_0, F)$ .

## 3. FIRST EXAMPLE

We will consider restricted permutations satisfying the condition  $-1 \leq p(i)-i \leq 1$  for all  $i \in \mathbb{N}_n$ . We will construct corresponding automaton  $M_1$ .

The alphabet in this case is:  $A = \{-1, 0, 1\}$ .

Internal states are  $S = \{a, b, Q\}$ . We don't have second coordinate, because we don't consider odd or even permutations, but all permutations. In the following table we have the correspondence between the internal states and the 1/0 (filled/vacant) slots:

**Table 1:** The internal states

States	<i>a</i>	<i>b</i>	$\underline{Q}$
filled and vacant slots	100	010	

The start state is *a*. Also, the only one “yes” state is *a*.

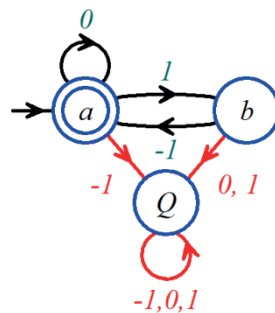
The next-state function  $F(S,a)$  of automaton  $M_1$  is given by Table 2:

**Table 2:** The next-state function  $F(S,a)$

States <i>S</i>	Inputs <i>a</i>		
	-1	0	1
<i>a</i>	$\underline{Q}$	<i>a</i>	<i>b</i>
<i>b</i>	<i>a</i>	$\underline{Q}$	$\underline{Q}$
$\underline{Q}$	$\underline{Q}$	$\underline{Q}$	$\underline{Q}$

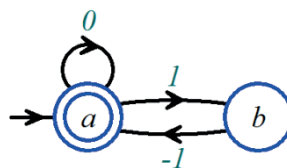
For example, if the state of automaton is *a* corresponding to slots 100. Input -1 may be viewed as causing a change in the state of the automaton from *a* to  $\underline{Q}$ , because the first slot is filled. Input 0 cause change from *a* to *a*: from the state  $a = 100$  we get 110 when we fill momentary element (that’s 1 on the second slot) and after the dropping the first slot off we get 10 and putting 0 (one empty slot) at the end, we finally get 100. Input 1 cause change from *a* to *b* (from the state  $a = 100$  we get 101 and after that 01 and finally  $b = 010$ ). From the state  $b = 010$  we need to fill 0 at the first slot, so  $F(b,-1)=a$  and  $F(b,0)=\underline{Q}$  and  $F(b,1)=\underline{Q}$ .

In the Picture 1 is represented automaton  $M_1$  with the corresponding digraph (it represents the next state function  $F$ ).



**Picture 1:** The automaton  $M_1$ .

For more clarity we can omit edges leading to sink state  $\underline{Q}$  and the state  $\underline{Q}$  – leading to a new automaton  $M_2$ . In a fact,  $M_2$  is nondeterministic automaton and it is represented in the Picture 2.



**Picture 2:** Nondeterministic automaton  $M_2$ .

The state diagram of the automaton  $M_2$  represented in the Picture 2 is an oriented graph. The construction of the restricted permutation corresponds to the forming of the closed path of length  $n$  from vertex *a* to vertex *a* in the graph  $M_2$ . We will denote the number of the closed paths of length  $n$  from vertex *v* to vertex *a* with  $v(n)$ . For the initial conditions we take  $a(0)=1$  and  $b(0)=0$ . This is how we got that  $a(n)$  is equal to the number of the restricted permutations of the set  $N_n$  satisfying the condition  $p(i)-i \in \{-1,0,1\}$ . Each of closed paths in oriented graph  $M_2$  is consisted of elementary paths *a-a* (of length 1) and *a-b-a* (of length 2). This conclusion establishes the bijection between the compositions of number  $n$  into elements from the set  $\{1,2\}$  and the restricted permutations satisfying the condition  $p(i)-i \in \{-1,0,1\}$  for all  $i \in N_n$ .

We have the conclusion that number of the closed path of length  $n+1$  from vertex *a* to vertex *a* is equal to the sum of the number of the closed path of length  $n$  from vertex *a* to vertex *a* and the number of the closed path of length  $n$  from vertex *b* to vertex *a*. Similarly, the number of the closed path of length  $n+1$  from vertex *b* to vertex *a* is equal to the number of the closed path of length  $n$  from vertex *a* to vertex *a*, so we come to the following system of recurrence equations:

$$a(n+1) = a(n) + b(n) \tag{1}$$

$$b(n+1) = a(n) \tag{2}$$

with initial conditions

$$a(0) = 1 \text{ and } b(0) = 0.$$

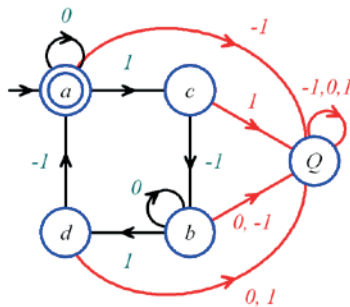
From the equation (2) we find  $b(n) = a(n-1)$ . Putting that into the equation (1) gives homogeneous linear recurrence equation with constant coefficients:

$$a(n+1) = a(n) + a(n-1), \text{ with initial conditions } a(0)=1, a(1)=1.$$

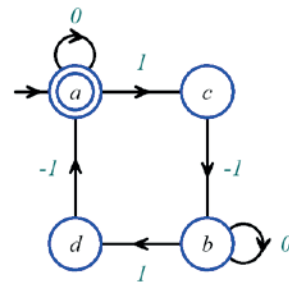
So, we find that  $a(n+1) = F(n+1)$ , where  $F(n+1)$  is  $(n+1)$ -st Fibonacci number.

### 3. SECOND EXAMPLE

In [3] we considered even restricted permutations satisfying the condition  $-1 \leq p(i) - i \leq 1$  for all  $i \in \mathbb{N}_n$ . We constructed corresponding automaton  $M_3$ : and corresponding nondeterministic automaton  $M_4$ .



Picture 3: The automaton  $M_3$ .



Picture 4: Nondeterministic automaton  $M_4$ .

In [3] we find the following connection:

$$a(n) = \frac{1}{2} \cdot (F(n+1) + x(n)),$$

where  $F(n+1)$  represent  $(n+1)$ -st Fibonacci number

$$F(n+1) = \frac{1}{\sqrt{5}} \cdot \left( \frac{1+\sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \cdot \left( \frac{1-\sqrt{5}}{2} \right)^{n+1}$$

and sequence  $x(n)$  is given with:

$$x(n) = \cos \frac{n\pi}{3} + \frac{1}{\sqrt{3}} \sin \frac{n\pi}{2}.$$

The sequence  $x(n)$  has plenty of different interesting interpretations:

$$x(n) = x(n-1) - x(n-2), x(0) = x(1) = 1.$$

$$x(n) = \begin{cases} 1, & n \equiv_6 0,1 \\ 0, & n \equiv_6 2,5 \\ -1, & n \equiv_6 3,4 \end{cases}$$

$$x(n) = \frac{2}{\sqrt{3}} \sin \frac{(n+1)\pi}{2}.$$

$x(n) = U(n, 1/2)$ , where  $U(n, x)$  is Chebyshev polynomial.

The solution to the combinatorial problem:

*How many patterns of  $n$  children in a row are there if the girls must appear in groups of at least three?* is our sequence  $a(n+1)$ .

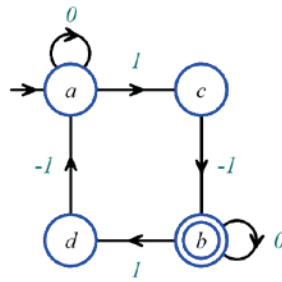
We will omit the proof (see [7]), but we will illustrate it for  $n = 4$ .



Picture 5:  $a(5)=4$  patterns of  $n=4$  children.



From the automaton  $M_4$  we can obtain automaton  $M_5$  which recognize odd restricted permutations satisfying the condition  $-1 \leq p(i) - i \leq 1$  for all  $i \in N_n$ . Everything is same, but the only one “yes” state differs: in automaton  $M_4$  it is  $a$  and in automaton  $M_5$  it is  $b$ .



Picture 6: Nondeterministic automaton  $M_5$ .

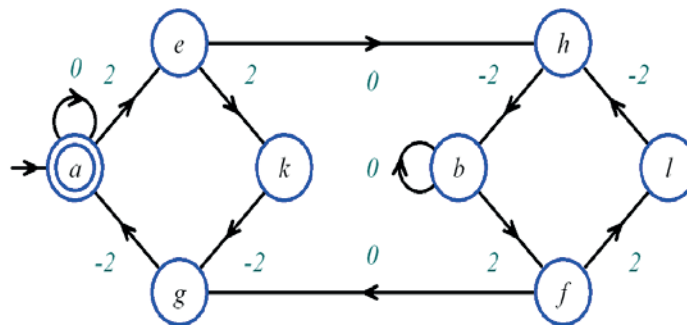
#### 4. THIRD EXAMPLE

In [4] we considered restricted permutations satisfying the condition  $p(i) - i \in \{-2, 0, 2\}$  for all  $i \in N_n$ , and we constructed the corresponding automaton  $M$ . There we find:

$$a(n) = \frac{1}{5} \left( L(n+2) + 2 \cdot \cos \frac{n\pi}{2} + \sin \frac{n\pi}{2} \right).$$

( $L(n)$  denotes  $n$ -th element of Lucas sequence).

In [3] we went further. We considered even (and odd) restricted permutations satisfying the condition  $p(i) - i \in \{-2, 0, 2\}$  for all  $i \in N_n$ . We constructed corresponding nondeterministic automaton  $M_6$ :



Picture 7: Nondeterministic automaton  $M_6$ .

In [3] we the solution which also depends on  $n$ -th element of Lucas sequence:

$$a(n) = \frac{1}{10} \cdot L(n+1) + y(n) + z(n),$$

and initial conditions

$$y(n) = \frac{1}{5}, \frac{1}{10}, -\frac{1}{5}, -\frac{1}{10}, \text{ respectively, for } n \equiv_4 0, 1, 2, 3 \text{ and}$$

$$z(n) = \frac{1}{2} \text{ for } n \equiv_6 0, 1, 2, \quad z(n) = \frac{1}{10} \text{ for } n \equiv_6 3, 4, 5.$$

From the automaton  $M_6$  (represented in the Picture 7) we can establish the bijection between the compositions of number  $n$  into elements from the set  $\{1, 3, 4\}$  which have even number of 3s and the even restricted permutations satisfying  $p(i) - i \in \{-2, 0, 2\}$  for all  $i \in N_n$ .

With slightly modifications from the automaton  $M_4$  we can find the automaton  $M_3$  which recognize only odd restricted permutations satisfying  $p(i) - i \in \{-2, 0, 2\}$  for all  $i \in N_n$ : the only one “yes” state is  $b$  and we can find the connection with the compositions of number  $n$  into elements from the set  $\{1, 3, 4\}$  which odd number of 3s.

#### 4. FOURTH EXAMPLE

In [2] we generalize this technique to the circular case – which is surely the interesting one. We asked for the number of the permutations of the set  $N_n$  such that  $p(i) - i \equiv 0, 1, \dots, k \pmod{n}$ . This has the effect of allowing  $p(n) \in \{1, 2, \dots, k\}$ ,  $p(n-1) \in \{1, 2, \dots, k-1\}$ , and so on,  $p(n-k+1)=1$ . There we find:

$$a(n) = \begin{cases} 1, & n = 0, 1 \\ 2, & n = 2 \\ 2 + L(n), & n \geq 3 \end{cases}$$

(where  $L(n)$  denotes  $n$ -th element of Lucas sequence).

Similarly, but with much more hard calculations, we can solve same problem for  $k = 3$ . Solution in this case is given by:

$$a(n) = 2 + 6T(n) + 8T(n-1) + 2T(n-2), \quad \text{for } n \geq 3.$$

(where  $T(n)$  denotes  $n$ -th element of Tribonacci sequence given with  $T(n+1) = T(n) + T(n-1) + T(n-2)$ ,  $T(0) = 0$ ,  $T(1) = 1$ ,  $T(2) = 1$  – see [6]). Tribonacci sequence is a generalisation of Fibonacci sequence.

Also, we find Tribonacci sequence when we consider restricted permutations satisfying the condition  $-1 \leq p(i) - i \leq 2$  for all  $i \in N_n$  – see [1].

#### 5. CONCLUSION

The finite state automata are powerful tool for generating some combinatorial structures. Tracing that generation leads us to the system of recurrence equations. By this approach we make connections to well known combinatorial sequences (Fibonacci, Lucas, Tribonacci). Searching for the number of restricted permutations is a very current issue, which can be seen in [8], [9] and [10].

#### REFERENCES

- [1] Baltić, V. (2010) On the Number of Certain Types of Strongly Restricted Permutations. *Applicable Analysis and Discrete Mathematics*, 4 (1), 119-135.
- [2] Baltić, V. (2009) Applications of the finite state automata in the enumerative combinatorics. *Proceedings of XXXVI Symposium on Operational Research*, 155-158.
- [3] Baltić, V., Saković, M. (2012) The counting of even and odd restricted permutations with the finite state automata, *Proceedings of XXXIX Symposium on Operational Research*, 217-220.
- [4] Baltić, V. (2012) Applications of the finite state automata for counting restricted permutations and variations. *Yugoslav Journal of Operations Research*, 22(2), 183-198.
- [5] Čangalović, M., Manojlović, V., Baltić, V. (2009) *Diskretne matematičke strukture*, FON.
- [6] Stevanović, D., Baltić, V., Simić, S., Ćirić, M. (2008) *Diskretna matematika – Osnove kombinatorike i teorije grafova*, DMS.
- [7] Conway, J.H., Guy, R.K. (1996), *The book of numbers*, Springer-Verlag, p. 205
- [8] Kitaev, S. (2011), *Patterns in permutations and words*, Springer Verlag.
- [9] Kløve, T. (2009) Generating functions for the number of permutations with limited displacement, *The Electronic Journal of Combinatorics* 16, #R104.
- [10] Mansour, T. (2006), Restricted even permutations and Chebyshev polynomials, *Discrete Mathematics*, 306, 1161-1176.

## SOME PAIRS AND TRIPLETS OF EQUIENERGETIC GRAPHS

IRENA M. JOVANOVIĆ<sup>1</sup>

<sup>1</sup> School of Computing – Union University, Belgrade, Serbia, irenaire@gmail.com

**Abstract:** In this paper, some pairs and triplets of equienergetic and mutually non-cospectral graphs are exposed. By these examples, some of the results present in the literature are slightly improved and supplemented.

**Keywords:** graph, spectrum, graph energy, equienergetic graphs, non-cospectral graphs

### 1. INTRODUCTION

Let  $G$  be a simple graph (i.e. without loops and multiple edges) of order  $n$ , and let  $A$  be the adjacency matrix of  $G$ . The *characteristic polynomial*  $P_G(x) = \det(A - xI)$  of  $G$  is the characteristic polynomial of its adjacency matrix  $A$ , while the (*adjacency*) *eigenvalues*  $\lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$  of  $G$  are the eigenvalues of  $A$ . These eigenvalues form the *spectrum* of  $G$ . If  $\lambda_i(G)$ , for some  $i$ , is the eigenvalue of the multiplicity  $k$ , we will write  $[\lambda_i(G)]^k$ .

The *energy*  $E(G)$  of a graph  $G$  was introduced by I. Gutman in 1978. in his paper [7]:

$$E(G) = \sum_{i=1}^n |\lambda_i(G)|.$$

For the last twenty years, this graph invariant has attracted the attention of researchers, especially in the field of mathematics and chemistry. For more details about graph energy, the reader is referred to the monographs [12] and [17], and the review papers [10] and [11].

Two graphs with the same number of vertices are said to be *equienergetic* if they have the same energy. Since it is obvious that two isomorphic or two *cospectral* (i.e. with the same spectra) graphs are equienergetic, it is of interest to consider non-isomorphic and non-cospectral graphs. The concept of equienergetic graphs was put forward independently by Brankov et al. [3] and Balakrishnan [2], and since then, there are many published papers related to this topic: [1], [4], [8], [9], [13], [14], [15], [16], [18], [19], [21], [22], [23], [24], [26], [20], [27], [28].

The notation common for spectral graph theory is used in the paper (see [5] and [6]). In that way,  $K_n$  is the complete graph on  $n$  vertices, while  $K_{n_1, n_2}$  is the complete bipartite graph of order  $n_1 + n_2$ . If  $n_1 = n_2 = 1$ , we mean that  $K_{1,1}$  is actually  $K_2$ .

The *degree of a vertex*  $v$  in a given graph  $G$  is the number of vertices of  $G$  which are adjacent to  $v$ . If all vertices in  $G$  have equal degrees, say  $r$ , then  $G$  is a  *$r$ -regular graph*. A *strongly regular graph*  $G$  with parameters  $(n, r, e, f)$  is a  $r$ -regular graph on  $n$  vertices in which any two adjacent vertices have exactly  $e$  common neighbours, and any two non-adjacent vertices have exactly  $f$  common neighbours. For more details about strongly regular graphs, see Section 3.6. in [6]. If  $G = SRG(n, r, e, f)$  is a connected strongly regular graph, then the adjacency spectrum of  $G$  (Theorem 3.6.5 from [6]) consists of:  $r$ ,  $[s]^k$  and  $[t]^l$ , where  $s, t = \frac{1}{2} \left( (e - f) \pm \sqrt{\Delta} \right)$ ;  $k, l = \frac{1}{2} \left( n - 1 \mp \frac{2r + (n-1)(e-f)}{\sqrt{\Delta}} \right)$ , and  $\Delta = (e - f)^2 + 4(r - f)$ .

The *complement*  $\bar{G}$  of a graph  $G$  with the vertex set  $V(G)$  is the graph whose the set of vertices is  $V(G)$  and in which two vertices are adjacent if and only if they are not adjacent in  $G$ . If  $G$  is a  $r$ -regular graph of order  $n$  whose spectrum regarding the adjacency matrix is  $r = \lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$ , then following theorem holds:

► **Theorem 1.** (Theorem 2.1.2 from [6]) If  $G$  is a regular graph of degree  $r$  with  $n$  vertices, then:

$$P_{\bar{G}}(x) = (-1)^n \frac{x - n + r + 1}{x + r + 1} P_G(-x - 1).$$

This means that the complement  $\bar{G}$  of  $G$  is  $(n - 1 - r)$ -regular, and that the adjacency spectrum of  $\bar{G}$  is:  $n - 1 - r \geq -1 - \lambda_n(G) \geq -1 - \lambda_{n-1}(G) \geq \dots \geq -1 - \lambda_2(G)$ .

The *line graph*  $L(G)$  of a graph  $G$  is the graph whose vertices are the edges of  $G$ , with two vertices in  $L(G)$  adjacent whenever the corresponding edges in  $G$  have exactly one vertex in common. The *iterated line graphs* of a graph  $G$  are defined recursively as:  $L^2(G) = L(L(G)), L^3(G) = L(L^2(G)), \dots, L^k(G) = L(L^{k-1}(G)), \dots$ . It is assumed that  $L^0(G) \equiv G$  and  $L^1(G) \equiv L(G)$ . The following theorem will be relevant in the next section:

► **Theorem 2.** (Theorem 2.4.1 from [6]) *If  $G$  is a regular graph of degree  $r$ , with  $n$  vertices and  $m = \frac{nr}{2}$  edges, then*

$$P_{L(G)}(x) = (x + 2)^{m-n} P_G(x - r + 2).$$

Now, we state the definition of a very general graph operation called NEPS (i.e. *non-complete extended  $p$ -sum*) as it is done in [6] (see Definition 2.5.1):

► **Definition 3.** Let  $\mathcal{B}$  be a set of non-zero binary  $n$ -tuples, i.e.  $\mathcal{B} \subseteq \{0, 1\}^n / \{(0, \dots, 0)\}$ . The NEPS of graphs  $G_1, \dots, G_n$  with basis  $\mathcal{B}$  is the graph with vertex set  $V(G_1) \times \dots \times V(G_n)$ , in which two vertices, say  $(x_1, \dots, x_n)$  and  $(y_1, \dots, y_n)$ , are adjacent if and only if there exists an  $n$ -tuple  $\beta = (\beta_1, \dots, \beta_n) \in \mathcal{B}$  such that  $x_i = y_i$  whenever  $\beta_i = 0$ , and  $x_i$  is adjacent to  $y_i$  (in  $G_i$ ) whenever  $\beta_i = 1$ .

In this paper, some special cases of NEPS will be considered. Precisely, for  $n = 2$  we will deal with the following familiar operations:

- (i) the *sum*  $G_1 + G_2$  of graphs  $G_1$  and  $G_2$ , when  $\mathcal{B} = \{(0, 1), (1, 0)\}$ , and
- (ii) the *product*  $G_1 \times G_2$  of graphs  $G_1$  and  $G_2$ , when  $\mathcal{B} = \{(1, 1)\}$ .

The spectrum of the NEPS of graphs  $G_1, \dots, G_n$  with basis  $\mathcal{B}$ , is given by the following statement:

► **Theorem 4.** (Theorem 2.5.4 from [6]) *If  $\lambda_{i_1}, \dots, \lambda_{i_k}$  are the eigenvalues of  $G_i$ ,  $i = 1, 2, \dots, n$ , then the spectrum of the NEPS of  $G_1, \dots, G_n$  with basis  $\mathcal{B}$  consists of all possible values  $\Lambda_{i_1, \dots, i_n}$ , where  $\Lambda_{i_1, \dots, i_n} = \sum_{\beta \in \mathcal{B}} \lambda_{i_1}^{\beta_1} \dots \lambda_{i_n}^{\beta_n}$ ,*

$$i_h = 1, \dots, k_h, h = 1, \dots, n.$$

Therefore, if  $\lambda_1(G_1), \lambda_2(G_1), \dots, \lambda_{n_1}(G_1)$  and  $\lambda_1(G_2), \lambda_2(G_2), \dots, \lambda_{n_2}(G_2)$  are the eigenvalues of  $G_1$  and  $G_2$ , respectively, then  $\lambda_i(G_1) + \lambda_j(G_2)$  and  $\lambda_i(G_1) \cdot \lambda_j(G_2)$ , for  $i = 1, 2, \dots, n_1$  and  $j = 1, 2, \dots, n_2$ , are the eigenvalues of  $G_1 + G_2$  and  $G_1 \times G_2$ , respectively.

Let us notice that if  $G_1$  is a  $r_1$ -regular graph and  $G_2$  is a  $r_2$ -regular graph, then  $G_1 + G_2$  is  $(r_1 + r_2)$ -regular graph, while  $G_1 \times G_2$  is  $r_1 \cdot r_2$ -regular graph.

The paper is organized as follows. In Section 2, some pairs, while in Section 3, some triplets of equienergetic and mutually non-cospectral graphs are given.

## 2. CONSTRUCTING SOME PAIRS OF EQUIENERGETIC GRAPHS

The following statement is a slight generalization of Theorem 8 from [4].

► **Theorem 5.** *Let  $G$  be a connected graph of order  $n$  with the eigenvalues  $\lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$ . Then  $G + K_{\frac{m}{2}, \frac{m}{2}}$  and  $G \times K_{\frac{m}{2}, \frac{m}{2}}$ , where  $m \geq 2$  is an even integer, are equienergetic graphs if and only if  $|\lambda_i(G)| \geq \frac{m}{2}$ , for all  $i = 1, 2, \dots, n$ .*

**Proof.** Let us suppose that  $|\lambda_i(G)| \geq \frac{m}{2}$  holds for all  $i = 1, 2, \dots, n$ . By construction,  $G + K_{\frac{m}{2}, \frac{m}{2}}$  and  $G \times K_{\frac{m}{2}, \frac{m}{2}}$  are graphs of the same order, i.e.  $n \cdot m$ . According to Theorem 4, the spectrum of  $G + K_{\frac{m}{2}, \frac{m}{2}}$  is:  $\frac{m}{2} + \lambda_i(G)$ ,  $i = 1, \dots, n$ ;  $[\lambda_i(G)]^{m-2}$ ,  $i = 1, \dots, n$  and  $-\frac{m}{2} + \lambda_i(G)$ ,  $i = 1, \dots, n$ . Let us denote by  $n_p$  the number of positive eigenvalues of  $G$ . Then we compute:

$$\begin{aligned} E(G + K_{\frac{m}{2}, \frac{m}{2}}) &= \sum_{i=1}^n \left| \frac{m}{2} + \lambda_i(G) \right| + \sum_{i=1}^n \left| -\frac{m}{2} + \lambda_i(G) \right| + (m-2)E(G) \\ &= \sum_{i=1}^{n_p} \left( \left| \frac{m}{2} + \lambda_i(G) \right| + \left| -\frac{m}{2} + \lambda_i(G) \right| \right) + (m-2)E(G) \\ &\quad + \sum_{i=n_p+1}^n \left( \left| \frac{m}{2} + \lambda_i(G) \right| + \left| -\frac{m}{2} + \lambda_i(G) \right| \right). \end{aligned}$$

Since  $|\lambda_i(G)| \geq \frac{m}{2}$ , for  $i = 1, 2, \dots, n$ , we have:

$$\begin{aligned} E(G + K_{\frac{m}{2}, \frac{m}{2}}) &= \sum_{i=1}^{n_p} \left( \frac{m}{2} + |\lambda_i(G)| - \frac{m}{2} + |\lambda_i(G)| \right) + \sum_{i=n_p+1}^n \left( |\lambda_i(G)| - \frac{m}{2} + |\lambda_i(G)| + \frac{m}{2} \right) + (m-2)E(G) \\ &= mE(G). \end{aligned} \tag{1}$$

According to Theorem 4, the spectrum of  $G \times K_{\frac{m}{2}, \frac{m}{2}}$  consists of the eigenvalues:  $\frac{m}{2} \cdot \lambda_i(G)$ ,  $i = 1, 2, \dots, n$ ,  $[0]^{(m-2)n}$  and  $-\frac{m}{2} \cdot \lambda_i(G)$ ,  $i = 1, 2, \dots, n$ . Therefore, the energy of  $G \times K_{\frac{m}{2}, \frac{m}{2}}$  is:

$$E(G \times K_{\frac{m}{2}, \frac{m}{2}}) = \sum_{i=1}^n \left| \frac{m}{2} \cdot \lambda_i(G) \right| + \sum_{i=1}^n \left| -\frac{m}{2} \cdot \lambda_i(G) \right| = mE(G). \tag{2}$$

Comparing (1) and (2), we obtain the first part of the proof.

Conversely, let  $G$  be a connected graph with the eigenvalues  $\lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$ . By contradiction, let us suppose that  $|\lambda_i(G)| < \frac{m}{2}$ , for some  $i \in \{1, 2, \dots, n\}$ . Then, if  $\lambda_i(G) \geq 0$ , we have that

$$\left| \lambda_i(G) + \frac{m}{2} \right| + \left| \lambda_i(G) - \frac{m}{2} \right| = m > 2|\lambda_i(G)|.$$

Also, if  $\lambda_i(G) < 0$ , then

$$\left| \lambda_i(G) + \frac{m}{2} \right| + \left| \lambda_i(G) - \frac{m}{2} \right| = m > 2|\lambda_i(G)|.$$

Let us suppose that the eigenvalues of  $G$  are such that:  $|\lambda_1(G)| \geq \frac{m}{2}, |\lambda_2(G)| \geq \frac{m}{2}, \dots, |\lambda_k(G)| \geq \frac{m}{2}$ , and  $|\lambda_{k+1}(G)| < \frac{m}{2}, \dots, |\lambda_n(G)| < \frac{m}{2}$ . Then we have:

$$E(G + K_{\frac{m}{2}, \frac{m}{2}}) = \sum_{i=1}^k \left( \left| \lambda_i(G) + \frac{m}{2} \right| + \left| \lambda_i(G) - \frac{m}{2} \right| \right) + \sum_{i=k+1}^n \left( \left| \lambda_i(G) + \frac{m}{2} \right| + \left| \lambda_i(G) - \frac{m}{2} \right| \right) + (m-2)E(G).$$

As we have seen in the first part of the proof,  $\sum_{i=1}^k \left( \left| \lambda_i(G) + \frac{m}{2} \right| + \left| \lambda_i(G) - \frac{m}{2} \right| \right) = 2 \sum_{i=1}^k |\lambda_i(G)|$ . Also, we have

that  $\sum_{i=k+1}^n \left( \left| \lambda_i(G) + \frac{m}{2} \right| + \left| \lambda_i(G) - \frac{m}{2} \right| \right) > 2 \sum_{i=k+1}^n |\lambda_i(G)|$ . Consequently,

$$E(G + K_{\frac{m}{2}, \frac{m}{2}}) > 2 \sum_{i=1}^k |\lambda_i(G)| + 2 \sum_{i=k+1}^n |\lambda_i(G)| + (m-2)E(G) = mE(G) = E(G \times K_{\frac{m}{2}, \frac{m}{2}}).$$

This contradiction completes the proof. ◀

► **Remark.** For  $m = 2$ , graphs  $G + K_{\frac{m}{2}, \frac{m}{2}}$  and  $G \times K_{\frac{m}{2}, \frac{m}{2}}$  are non-cospectral, as it is proved by Theorem 8 in [4]. For an even  $m$ ,  $m \geq 4$ ,  $G + K_{\frac{m}{2}, \frac{m}{2}}$  and  $G \times K_{\frac{m}{2}, \frac{m}{2}}$  are non-cospectral graphs if  $\lambda_1(G) \neq \frac{m}{m-2} (\leq 2)$ . Indeed, in that case the eigenvalues  $\frac{m}{2} \cdot \lambda_1(G)$  and  $\frac{m}{2} + \lambda_1(G)$  of graphs  $G \times K_{\frac{m}{2}, \frac{m}{2}}$  and  $G + K_{\frac{m}{2}, \frac{m}{2}}$ , respectively, are different eigenvalues.

► **Example 6.** The *Berlekamp-van Lint-Seidel graph*  $G_{BL}$  is a connected strongly regular graph with parameters  $(243, 22, 1, 2)$  and the eigenvalues:  $22, [4]^{132}, [-5]^{110}$ . Graphs  $G_{BL} \times K_{\frac{m}{2}, \frac{m}{2}}$  and  $G_{BL} + K_{\frac{m}{2}, \frac{m}{2}}$ , where  $m \geq 2$  is an even integer, are of the same order (i.e.  $243m$ ), and they are such that  $\lambda_1(G_{BL}) = 22 \neq \frac{m}{m-2}$ , for every  $m > 2$ . Since  $|\lambda_i(G_{BL})| \geq \frac{m}{2}$ , for  $m = 2, 4, 6, 8$  and for all  $i = 1, 2, \dots, 243$ , from Theorem 5 and Theorem 8 from [4], it follows that  $G_{BL} \times K_{\frac{m}{2}, \frac{m}{2}}$  and  $G_{BL} + K_{\frac{m}{2}, \frac{m}{2}}$ , for  $m = 2, 4, 6, 8$ , are non-cospectral equienergetic graphs.

Indeed, let us verify this claim by direct computation for  $m = 8$ . Using Theorem 4, we find that the spectrum of  $G_{BL} \times K_{4,4}$  is:  $88, [20]^{110}, [16]^{132}, [0]^{1458}, [-16]^{132}, [-20]^{110}, -88$ , so we compute  $E(G_{BL} \times K_{4,4}) = 8800$ . In a similar way, since the eigenvalues of  $G_{BL} + K_{4,4}$  are:  $26, [22]^6, 18, [8]^{132}, [4]^{792}, [0]^{132}, [-1]^{110}, [-5]^{660}, [-9]^{110}$ , the energy of this graph is equal to  $E(G_{BL} + K_{4,4}) = 8800$ .

► **Example 7.** The *Hall-Janko graph*  $G_{HJ}$  is a connected strongly regular graph with parameters  $(100, 36, 14, 12)$  and the eigenvalues  $36, [6]^{36}, [-4]^{36}$ . Graphs  $G_{HJ} \times K_{\frac{m}{2}, \frac{m}{2}}$  and  $G_{HJ} + K_{\frac{m}{2}, \frac{m}{2}}$  are of the same order (i.e.  $100m$ ). It holds that  $\lambda_1(G_{HJ}) = 36 \neq \frac{m}{m-2}$ , for every  $m > 2$ . Since  $|\lambda_i(G_{HJ})| \geq \frac{m}{2}$ , for  $m = 2, 4, 6, 8$  and for all  $i = 1, 2, \dots, 100$ , from Theorem 5 and Theorem 8 from [4], it follows that  $G_{HJ} \times K_{\frac{m}{2}, \frac{m}{2}}$  and  $G_{HJ} + K_{\frac{m}{2}, \frac{m}{2}}$ , for  $m = 2, 4, 6, 8$ , are non-cospectral equienergetic graphs.

► **Example 8.** Graphs  $L(K_{n,n}) \times K_{2,2}$  and  $L(K_{n,n}) + K_{2,2}$  are non-cospectral and equienergetic for each  $n \geq 4$ . Namely, since  $K_{n,n}$  is  $n$ -regular graph of order  $2n$ , with the spectrum  $n, [0]^{2n-2}, -n$ , according to Theorem 2, the spectrum of  $L(K_{n,n})$  consists of the eigenvalues  $2n - 2, [n - 2]^{2n-2}, [-2]^{(n-1)^2}$ . Graphs  $L(K_{n,n}) \times K_{2,2}$  and  $L(K_{n,n}) + K_{2,2}$  are of the equal order, i.e.  $4n^2$ . Also,  $\lambda_1(L(K_{n,n})) = 2n - 2 \neq 2$ . Since  $|\lambda_i(L(K_{n,n}))| \geq 2$ , for all  $i = 1, 2, \dots, 4n^2$  and  $n \geq 4$ , from Theorem 5 it follows that  $L(K_{n,n}) \times K_{2,2}$  and  $L(K_{n,n}) + K_{2,2}$  are non-cospectral equienergetic graphs.

This claim can be verified by direct computation, as well. Using Theorem 4, we find that the spectrum of  $L(K_{n,n}) \times K_{2,2}$  is:  $4n - 4, [2n - 4]^{2n-2}, [4]^{(n-1)^2}, [0]^{2n^2}, [-4]^{(n-1)^2}, [-2n + 4]^{2n-2}, -4n + 4$ , and therefore  $E(L(K_{n,n}) \times K_{2,2}) = 16(n - 1)^2$ . In a similar way, the spectrum of  $L(K_{n,n}) + K_{2,2}$  consists of the following eigenvalues:  $2n, [2n - 2]^2, 2n - 4, [n]^{2n-2}, [n - 2]^{2(2n-2)}, [n - 4]^{2n-2}, [0]^{(n-1)^2}, [-2]^{2(n-1)^2}$  and  $[-4]^{(n-1)^2}$ , wherefrom it follows that  $E(L(K_{n,n}) + K_{2,2}) = 16(n - 1)^2$ .

► **Example 9.** Let  $G$  be a  $r$ -regular graph,  $r \geq 4$ , of order  $n$  and with the spectrum:  $\lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$ . In [25], it is obtained that the eigenvalues of  $L^2(G)$  are:  $\lambda_i(G) + 3r - 6, i = 1, 2, \dots, n, [2r - 6]^{\frac{n}{2}(r-2)}$  and  $[-2]^{\frac{n}{2}(r-2)}$ , and by Theorem 1 from the same paper, it is stated that all negative eigenvalues of the considered graph are equal to  $-2$ . Graphs  $L^2(G) \times K_{2,2}$  and  $L^2(G) + K_{2,2}$  are of the same order, i.e.  $2nr(r - 1)$ . It holds that  $\lambda_1(L^2(G)) = 4r - 6 \neq 2$ , for  $r \geq 4$ . Also,  $|\lambda_i(L^2(G))| \geq 2$ , for all  $i = 1, 2, \dots, \frac{nr}{2}(r - 1)$ . Therefore, according to Theorem 5,  $L^2(G) \times K_{2,2}$  and  $L^2(G) + K_{2,2}$  are non-cospectral equienergetic graphs.

► **Proposition 10.** Let  $G_1$  and  $G_2$  be two non-cospectral connected  $r$ -regular graphs of order  $n$ , where  $r \leq \lfloor \frac{n}{2} \rfloor$ , with the eigenvalues  $\lambda_1(G_i) \geq \lambda_2(G_i) \geq \dots \geq \lambda_n(G_i), i = 1, 2$ . Then  $\underbrace{K_n + K_n + \dots + K_n}_r + G_1$  and  $\underbrace{K_n + K_n + \dots + K_n}_r + G_2$  are equienergetic non-cospectral graphs.

**Proof.** Using Theorem 4, we find that the spectrum of  $\underbrace{K_n + K_n + \dots + K_n}_r$  consists of the following eigenvalues:  $[(r - j) \cdot n - r]^{\binom{r}{j}(n-1)^j}$ , for  $j = 0, 1, \dots, r$ .  $\underbrace{K_n + K_n + \dots + K_n}_r + G_i, i = 1, 2$ , is  $(rn)$ -regular graph of order  $n^{r+1}$ , whose spectrum, according to Theorem 4, is:  $[(r - j) \cdot n - r + \lambda_k(G_i)]^{\binom{r}{j}(n-1)^j}$ , where  $j = 0, 1, \dots, r, k = 1, 2, \dots, n$  and  $i = 1, 2$ . Using the fact  $\lambda_k(G_i) \in [-r, r]$ , for all  $k = 1, 2, \dots, n$  and  $i = 1, 2$ , and the assumption  $r \leq \lfloor \frac{n}{2} \rfloor$ , we find that for  $j = 0, 1, \dots, r - 1$ , the following holds:  $(r - j)n - r + \lambda_k(G_i) \geq (r - j)n - 2r \geq (r - r + 1)n - 2r \geq n - 2 \cdot \lfloor \frac{n}{2} \rfloor \geq 0$ . Also,  $-r + \lambda_k(G_i) \leq -r + r = 0$ , for all  $k = 1, 2, \dots, n$  and  $i = 1, 2$ . Therefore, since  $\sum_{k=1}^n \lambda_k(G_i) = 0$ , for  $i = 1, 2$ , we compute:

$$\begin{aligned} E(\underbrace{K_n + K_n + \dots + K_n}_r + G_i) &= \sum_{j=0}^r \sum_{k=1}^n \binom{r}{j} (n-1)^j |(r-j)n - r + \lambda_k(G_i)| \\ &= \sum_{j=0}^{r-1} \binom{r}{j} (n-1)^j \sum_{k=1}^n ((r-j)n - r + \lambda_k(G_i)) + (n-1)^r \sum_{k=1}^n (r - \lambda_k(G_i)) \\ &= \sum_{j=0}^{r-1} \binom{r}{j} (n-1)^j \left( n((r-j)n - r) + \sum_{k=1}^n \lambda_k(G_i) \right) \\ &\quad + (n-1)^r \left( nr - \sum_{k=1}^n \lambda_k(G_i) \right) \\ &= \sum_{j=0}^{r-1} \binom{r}{j} (n-1)^j ((r-j)n^2 - nr) + nr(n-1)^r, \end{aligned}$$

wherefrom it follows that  $\underbrace{K_n + K_n + \dots + K_n}_r + G_1$  and  $\underbrace{K_n + K_n + \dots + K_n}_r + G_2$  are equienergetic graphs.

Since,  $G_i, i = 1, 2$ , are connected  $r$ -regular graphs, eigenvalues  $-r + \lambda_k(G_i)$ , for  $k = 1, 2, \dots, n$ , are less than zero if  $\lambda_k(G_i) \neq r$ , i.e. for  $k = 2, 3, \dots, n$ . Since  $G_1$  and  $G_2$  are non-cospectral graphs, then  $\lambda_l(G_1) \neq \lambda_l(G_2)$ , for some  $l \in \{2, 3, \dots, n\}$ . This means that  $-r + \lambda_l(G_1) \neq -r + \lambda_l(G_2)$ , and hence  $\underbrace{K_n + K_n + \dots + K_n}_r + G_1$  and  $\underbrace{K_n + K_n + \dots + K_n}_r + G_2$  are non-cospectral graphs. ◀

► **Example 11.** Graphs  $K_3 + K_2$  and  $K_{3,3}$  are connected 3-regular graphs of order 6. Since their spectra are:  $3, 1, [0]^2, [-2]^2$  and  $3, [0]^4, -3$ , respectively, they are non-cospectral graphs. Then, according to Proposition 10,  $K_6 + K_6 + K_6 + K_3 + K_2$  and  $K_6 + K_6 + K_6 + K_{3,3}$  are equienergetic and non-cospectral graphs. Indeed, the eigenvalues of  $K_6 + K_6 + K_6$  are as follows:  $15, [9]^{15}, [3]^{75}, [-3]^{125}$ . So, the spectrum of  $K_6 + K_6 + K_6 + K_3 + K_2$  is:

18, 16,  $[15]^2$ ,  $[13]^2$ ,  $[12]^{15}$ ,  $[10]^{15}$ ,  $[9]^{30}$ ,  $[7]^{30}$ ,  $[6]^{75}$ ,  $[4]^{75}$ ,  $[3]^{150}$ ,  $[1]^{150}$ ,  $[0]^{125}$ ,  $[-2]^{125}$ ,  $[-3]^{250}$  and  $[-5]^{250}$ , and we have  $E(K_6 + K_6 + K_6 + K_3 + K_2) = 4500$ . The spectrum of  $K_6 + K_6 + K_6 + K_{3,3}$  consists of the eigenvalues: 18,  $[15]^4$ ,  $[12]^{16}$ ,  $[9]^{60}$ ,  $[6]^{90}$ ,  $[3]^{300}$ ,  $[0]^{200}$ ,  $[-3]^{500}$ , and  $[-6]^{125}$ , so we find  $E(K_6 + K_6 + K_6 + K_{3,3}) = 4500$ .

### 3. CONSTRUCTING SOME TRIPLETS OF EQUIENERGETIC GRAPHS

In this section, we give a slight generalization of some examples of triplets of equienergetic and mutually non-cospectral graphs exposed in [4]. For this purpose we need the following corollary:

► **Corollary 12.** (Corollary 3.1 from [16]) *Let  $G$  be a  $r$ -regular graph of order  $n$  whose the adjacency eigenvalues are:  $\lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$ . Let us denote  $I_{-1} = \{i \in \{1, 2, \dots, n\} : \lambda_i(G) \in [-2, -1]\}$  and  $I_{+1} = \{i \in \{1, 2, \dots, n\} : \lambda_i(G) \in [0, 1]\}$ , and let us suppose that in the spectrum of  $G$  there are  $n_{-2}$  eigenvalues which are greater than or equal to  $-2$ , and that among them there are  $n_0$  non-negative eigenvalues. Then  $E(G + K_2) = E(\overline{G + K_2})$  if and only if*

$$n_{-2} + n_0 + \sum_{i \in I_{-1} \cup I_{+1}} \lambda_i(G) + |I_{-1}| - |I_{+1}| = r + 2,$$

where  $|I|$  denotes the cardinality of the set  $I$ .

Theorem 5 and Corollary 12 imply the following statement:

► **Proposition 13.** *Let  $G \cong K_n$  be a connected  $r$ -regular graph of order  $n$  with the following eigenvalues  $\lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$ , among which there are  $n_{-2}$  eigenvalues which are greater than or equal to  $-2$ , and  $n_0$  non-negative eigenvalues. Let us denote  $I_{-1} = \{i \in \{1, 2, \dots, n\} : \lambda_i(G) \in [-2, -1]\}$  and  $I_{+1} = \{i \in \{1, 2, \dots, n\} : \lambda_i(G) \in [0, 1]\}$ .  $(G + K_2, \overline{G + K_2}, G \times K_2)$  is a triplet of equienergetic and mutually non-cospectral graphs if and only if  $|\lambda_i(G)| \geq 1$ , for all  $i = 1, 2, \dots, n$ , and  $\sum_{i \in I_{-1} \cup I_{+1}} \lambda_i(G) = -n_{-2} - n_0 - |I_{-1}| + |I_{+1}| + r + 2$ ,*

where  $|I|$  denotes the cardinality of the set  $I$ .

► **Remark.** When  $|\lambda_i(G)| \geq 1$ , for all  $i = 1, 2, \dots, n$ , the set  $I_{+1}$  is empty, so  $|I_{+1}| = 0$ .

► **Remark.** By Theorem 8 from [4], it has been proved that  $G + K_2$  and  $G \times K_2$  are non-cospectral graphs, for an arbitrary connected graph  $G$ .

Graph  $G + K_2$  is a connected  $(r + 1)$ -regular graph of order  $2n$ . According to Theorem 1,  $\overline{G + K_2}$  is a  $(2n - r - 2)$ -regular graph. It holds that  $\lambda_1(G + K_2) \neq \lambda_1(\overline{G + K_2})$ . Indeed, if  $r + 1 = 2n - r - 2$ , it means that  $2(r + 1) + 1 = 2n$ , which is a contradiction since  $n$  and  $r$  are integers. Therefore,  $G + K_2$  and  $\overline{G + K_2}$  are non-cospectral graphs.

As we have seen,  $\overline{G + K_2}$  is a  $(2n - r - 2)$ -regular graph, while  $G \times K_2$  is a  $r$ -regular graph. Hence, their the largest eigenvalues are not equal eigenvalues, since, according to the assumption,  $r \neq n - 1$ . This means that  $\overline{G + K_2}$  and  $G \times K_2$  are non-cospectral graphs. ◀

► **Example 14.** Let  $G = SRG(d^2(d + 2), d(d^2 + d - 1), d(d^2 - 1), d(d^2 - 1))$ , where  $d > 2$ , be a connected strongly regular graph. By direct computation, it can be obtained that the spectrum of  $G$  is:  $d(d^2 + d - 1)$ ,  $[d]^{\frac{d}{2}(d^2 + d - 1)}$ , and  $[-d]^{\frac{1}{2}(d^3 + 3d^2 + d - 2)}$ . By Proposition 3.1 in [16], it has been proved that  $G + K_2$  and  $\overline{G + K_2}$  are equienergetic and non-cospectral graphs, such that  $E(G + K_2) = E(\overline{G + K_2}) = 2d(d^3 + 3d^2 + d - 2)$ . This means that the eigenvalues of  $G$  satisfy the following relation:  $n_{-2} + n_0 + \sum_{i \in I_{-1} \cup I_{+1}} \lambda_i(G) + |I_{-1}| - |I_{+1}| = r + 2$ .

Since  $d > 2$ ,  $|\lambda_i(G)| \geq 1$  holds for all  $i = 1, 2, \dots, d^2(d + 2)$ , which means that  $G \times K_2$  is equienergetic with  $G + K_2$ , and hence with  $\overline{G + K_2}$ . Indeed, the eigenvalues of  $G \times K_2$  are:  $d(d^2 + d - 1)$ ,  $[d]^{d^3 + 2d^2 - 1}$ ,  $[-d]^{d^3 + 2d^2 - 1}$ , and  $-d(d^2 + d - 1)$ , so its energy is equal to  $E(G \times K_2) = 2d(d^3 + 3d^2 + d - 2)$ . As the largest eigenvalues of graphs  $G + K_2$ ,  $\overline{G + K_2}$  and  $G \times K_2$  are  $d(d^2 + d - 1) + 1$ ,  $d^3 + 3d^2 + d - 2$  and  $d(d^2 + d - 1)$ , respectively, these graphs are mutually non-cospectral.

As observed in [16], the strongly regular graph  $G$  with parameters  $(640, 568, 504, 504)$  (for  $d = 8$ ), exists. The spectrum of this graph consists of the following eigenvalues: 568,  $[8]^{284}$ ,  $[-8]^{355}$ . Therefore, we have  $n_{-2} = n_0 = 285$ , and  $I_{-1} = I_{+1} = \emptyset$ . Since,  $|\lambda_i(G)| \geq 1$ , for all  $i = 1, 2, \dots, 640$  and  $568 + 2 - n_{-2} - n_0 = 0$ , from Proposition 13, it follows that  $G \times K_2$ ,  $G + K_2$  and  $\overline{G + K_2}$  are equienergetic and mutually non-cospectral graphs.

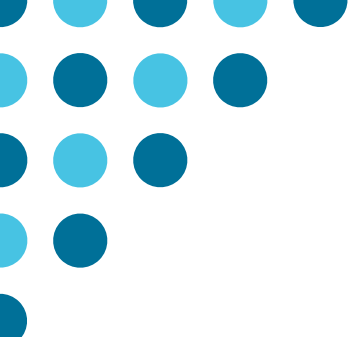
### Acknowledgement

I. M. Jovanović thanks the Serbian Ministry of Education, Science and Technological Development, for support through the Mathematical Institute of the Serbian Academy of Sciences and Arts.

## REFERENCES

- [1] Ashraf F., On energy of trees with perfect matching, *MATCH Communications in Mathematical and in Computer Chemistry*, 82 (2019), 439–442.
- [2] Balakrishnan R., The energy of a graph, *Linear Algebra and its Applications*, 387 (2004), 287–295.
- [3] Brankov V., Stevanović D., Gutman I., Equienergetic chemical trees, *Journal of the Serbian Chemical Society*, 69 (2004), 549–553.
- [4] Bonifacio A. S., Vinagre C. T. M., de Abreu N. M. M., Constructing pairs of equienergetic and non-cospectral graphs, *Applied Mathematics Letters*, 21 (2008), 338–341.
- [5] Cvetković D. M., Doob M., Sachs H., *Spectra of Graphs – Theory and Application*, Johann Ambrosius Barth Verlag, Heidelberg, 1995.
- [6] Cvetković D., Rowlinson P., Simić S., *An Introduction to the Theory of Graph Spectra*, Cambridge University Press, Cambridge, 2010.
- [7] Gutman I., The energy of a graph, *Ber. Math. Statist. Sect. Forschungsz. Graz*, 103 (1978), 1–22.
- [8] Gutman I., Open problems for equienergetic graphs, *Iranian Journal of Mathematical Chemistry*, 6 (2015), 185–187.
- [9] Gutman I., Firoozabadi S. Z., Pena J. A., Rada J., On the energy of regular graphs, *MATCH Communications in Mathematical and in Computer Chemistry*, 57 (2007), 435–442.
- [10] Gutman I., Furtula B., Survey of graph energies, *Mathematics Interdisciplinary Research*, 2 (2017), 85–129.
- [11] Gutman I., Furtula B., *Energies of Graphs, Survey, Census, Bibliography*, Center for scientific research of the Serbian Academy of Sciences and Arts and the University of Kragujevac, Kragujevac, 2019.
- [12] Gutman I., Li X., *Graph Energies - Theory and Applications*, University of Kragujevac, Kragujevac, 2016.
- [13] Indulal G., Vijayakumar A., On a pair of equienergetic graphs, *MATCH Communications in Mathematical and in Computer Chemistry*, 55 (2006), 83–90.
- [14] Indulal G., Vijayakumar A., Equienergetic self-complementary graphs, *Czechoslovak Mathematical Journal*, 58 (2008), 911–919.
- [15] Jog S. R., Hampiholi P. R., Joshi A. S., On the energy of some graphs, *Annals of Pure and Applied Mathematics*, 17 (2018), 15–21.
- [16] Jovanović I. M., Zogić E., Some spectral characterizations of equienergetic regular graphs and their complements, *MATCH Communications in Mathematical and in Computer Chemistry*, 86 (2021), 559–575.
- [17] Li X., Shi Y., Gutman I., *Graph Energy*, Springer, New York, 2012.
- [18] Liu J., Liu B., On a pair of equienergetic graphs, *MATCH Communications in Mathematical and in Computer Chemistry*, 59 (2008), 275–278.
- [19] Nikiforov V., Remarks on the energy of regular graphs, arXiv:1604.04275v3, 2016.
- [20] Ramane H. S., Ashoka K., Parvathalu B., Patil D., Gutman I., On complementary equienergetic strongly regular graphs, *Discrete Mathematics Letters*, 4 (2020), 50–55.
- [21] Ramane H. S., Gutman I., Walikar H. B., Halkarni S. B., Another class of equienergetic graphs, *Kragujevac Journal of Mathematics*, 26 (2004), 15–18.
- [22] Ramane H. S., Gutman I., Walikar H. B., Halkarni S. B., Equienergetic complement graphs, *Kragujevac Journal of Science*, 27 (2005), 67–74.
- [23] Ramane H. S., Parvathalu B., Patil D., Ashoka K., Graphs equienergetic with their complements, *MATCH Communications in Mathematical and in Computer Chemistry*, 82 (2019), 471–480.
- [24] Ramane H. S., Walikar H. B., Construction of equienergetic graphs, *MATCH Communications in Mathematical and in Computer Chemistry*, 57 (2007), 203–210.
- [25] Ramane H. S., Walikar H. B., Rao S. B., Acharya B. D., Hampiholi P. R., Jog S. R., Gutman I., Spectra and energies of iterated line graphs of regular graphs, *Applied Mathematics Letters*, 18 (2005), 679–682.
- [26] Ramane H. S., Walikar H. B., Rao S. B., Acharya B. D., Hampiholi P. R., Jog S. R., Gutman I., Equienergetic graphs, *Kragujevac Journal of Mathematics*, 26 (2004), 5–13.
- [27] Shalini S. G. B., Mayamma J., New results on equienergetic graphs of small order, *International Journal of Computational and Applied Mathematics*, 12 (2017), 595–602.
- [28] Tura F., More on equienergetic threshold graphs, arXiv: 1807.00627v1, 2018.





# HEURISTIKE

---

# HEURISTICS





## ALTERNATIVE LOCAL SEARCH ALGORITHMS IN GREEDY HEURISTIC ALGORITHMS FOR THE P-MEDIAN PROBLEM

LEV KAZAKOVTSSEV<sup>1</sup>, IVAN ROZHNOV<sup>2</sup>, GUZEL SHKABERINA<sup>3</sup>, ALENA STUPINA<sup>4</sup>, IGOR MASICH<sup>5</sup>

<sup>1</sup> Siberian Federal University – Laboratory “Hybrid Methods of Modelling and Optimization in Complex Systems”, levk@bk.ru

<sup>2</sup> Siberian Federal University – Laboratory “Hybrid Methods of Modelling and Optimization in Complex Systems”, ris2005@mail.ru

<sup>3</sup> Siberian Federal University – Laboratory “Hybrid Methods of Modelling and Optimization in Complex Systems”, z\_guzel@gmail.com

<sup>4</sup> Siberian Federal University – Laboratory “Hybrid Methods of Modelling and Optimization in Complex Systems”, h7677hm@gmail.com

<sup>5</sup> Siberian Federal University – Laboratory “Hybrid Methods of Modelling and Optimization in Complex Systems”, imasich@sfu-kras.ru

---

**Abstract:** Progress in the development of location theory and clustering of objects and data, the most popular of which are based on solving  $p$ -median and similar problems ( $k$ -means,  $k$ -medoid), is mainly aimed at improving the performance of algorithms. It is known that in local search algorithms, the result depends on the choice of the initial solution. Local search algorithms, as well as evolutionary algorithms, demonstrate fairly accurate results in this case. Various algorithms based on the use of greedy agglomerative procedures are able to obtain very accurate results that are difficult to improve by other methods. The greedy agglomerative procedure starts the search with a solution with an excessive number of demand points and combines the removal of the demand points with local search procedures. Usually, in the case of the  $p$ -median problem, the most widely used procedures are the Lloyd ( $k$ -means) procedure combined with the Weiszfeld algorithm for solving Weber ( $1$ -median) problem. In our research, we investigate the efficiency of various versions of hybridization of the  $k$ -means and Weiszfeld procedures embedded into a greedy agglomerative algorithm.

**Keywords:**  $p$ -median, location problem, clustering, Greedy Heuristic Method



## BEE COLONY OPTIMIZATION FOR MULTI-LABEL FEATURE SELECTION

LUKA MATIJEVIĆ<sup>1</sup>

<sup>1</sup> Mathematical Institute of the Serbian Academy of Sciences and Arts, luka@mi.sanu.ac.rs

**Abstract:** In this paper, we consider the problem of feature selection for multi-label data. Multi-label feature selection is a process of finding the appropriate subset of features that allows multi-label classifiers to find better solutions in a shorter amount of time. For this purpose, we developed the Bee Colony Optimization algorithm based on mutual information and compared it with other metaheuristics from literature, i.e. Ant Colony Optimization and Memetic Algorithm. After testing it on several benchmark instances, we concluded that our approach outperforms the other two methods.

**Keywords:** Combinatorial optimization, Metaheuristics, Mutual information, Classification

### 1. INTRODUCTION

In recent years, machine learning (ML) and data mining techniques have become invaluable tools in business, medicine, in banking and finance, and many other professional areas. With the evergrowing amount of data and its dimensionality, it is becoming increasingly important to properly select the appropriate subset of features that will allow the aforementioned techniques to provide users with good quality predictions in a reasonable amount of time.

The *Feature Selection* (FS) problem aims at reducing the dimensionality of the data by removing less relevant features. *Multi-label FS* is a more general case of FS, in the sense that each object in the data can have multiple labels associated with it. There are two ways of approaching this problem: transforming the multi-label data into single-label data and applying classical FS, or constructing algorithms that can directly deal with multi-label data. In both cases, there are three main types of methods: *filter*, *wrapper*, and *embedded* methods. Filter methods use statistical techniques to reduce the number of features, without evaluating the result with some specific ML model. In contrast, wrapper methods use ML models to evaluate every considered solution. Embedded methods reduce the number of features during the learning process itself. In this paper, we are only interested in wrapper methods, or, more precisely metaheuristics.

The main contribution of this paper is in the development of the *Bee Colony Optimization* (BCO) algorithm for the multi-label FS problem. We also presented a stochastic way of adding and removing features from the current solution, that can be utilized in other metaheuristics as well.

This paper is organized as follows: In Section 2 we present the formulation of the problem, followed by the relevant literature in Section 3. In Section 4 we provide a detailed description of our method, and in Section 5 we present the results of experimental evaluation.

### 2. PROBLEM FORMULATION

*Multi-label Feature Selection* (MLFS) problem can be formulated in the following manner:

► **Definition 1.** Let  $\mathcal{D}$  be a dataset where each row has a finite set of features  $S = \{s_1, s_2, \dots, s_n\}$  and  $\mathcal{M}$  is a specific machine learning model. The objective is to determine a feature subset  $S_k$  of size  $k$  ( $k < n$ ) where

$$\max_{S_k \subset S} \text{Accuracy}(\mathcal{M}(\mathcal{D}[S_k])) , \quad (1)$$

$\mathcal{D}[S_k]$  being the transformation of the original dataset. Only the features from  $S_k$  are present in  $\mathcal{D}[S_k]$ .

The function *accuracy* can be defined as follows.

► **Definition 2.** Let us assume that dataset  $\mathcal{D}$  consists of  $n$  test instances  $(x_i, Y_i)$ ,  $i = 1 \dots n$ , where  $Y_i$  is a subset of labels associated with instance  $x_i$ . We will denote the predicted set of labels as  $Z_i$  for each test instance. We define classification accuracy as:

$$Accuracy = \frac{1}{n} \sum_{i=1}^n I(Z_i = Y_i) \quad (2)$$

where  $I(true) = 1$  and  $I(false) = 0$ .

We can substitute the function *Accuracy* for any other appropriate multi-label metric, such as *Hamming-Loss* or *Precision*. Some of these metrics are described in a paper by Tsoumakas et al. (2009) [16].

### 3. RELATED WORK

There have been several attempts during recent years to apply metaheuristic methods to multi-label feature selection.

One of the first papers published on this topic was by Zhang et al. (2009) [18], who used the *Principal Component Analysis (PCA)* to reduce the number of irrelevant features before applying a *Genetic algorithm (GA)*. Shao et al. (2013) [14] used mutation-based *simulated annealing (SA)*, combined with GA and *greedy hill-climbing* algorithm. Similarly to [18], Yu et al. (2014) [17] discarded irrelevant features by using a forward search strategy, before applying the GA. In research by Lee and Kim (2015) [8], a *Memetic algorithm (MA)* is proposed, using a local search based on *mutual information*. Jungjit and Freitas (2015) [7] used a basic GA but introduced a different fitness function based on the correlation between features and labels and between pairs of features. *Particle Swarm Optimization (PSO)* is utilized by Zhang et al. (2017) [19] for MLFS, combined with a local learning strategy. In a study by Dowlatshahi et al. (2017) [5], the authors proposed a novel approach called *Epsilon-Greedy Swarm Optimizer*. Paniri et al. (2019) [13] utilized the *Ant Colony Optimization (ACO)* algorithm for MLFS.

### 4. PROPOSED METHOD

In this paper, we propose the *Bee Colony Optimization* based on the improvement concept (**BCOi**) for finding a satisfactory solution to MLFS. BCO is a stochastic, nature-inspired, population based metaheuristic, first introduced by Lučić and Teodorović [9, 10, 11], and since then it has been successfully applied to many optimization problems [1, 3, 4]. A more detailed description of the algorithm and its application can be found in papers [2, 15].

Firstly, we adopt the function  $Q(f)$  from paper [8] for evaluating the influence of features on the current solution (Equation 3). In this function,  $f$  stands for a specific feature we want to evaluate,  $Y$  is a set of all labels, and  $S_k$  is the set of already selected features. The function  $I(f, h)$  is approximated mutual information, presented in Equation 4.  $H(f)$  represents the entropy of variable  $f$ , while  $H(f, h)$  is the joint entropy of variables  $f$  and  $h$ . With this in mind, we can construct stochastic methods for adding and removing features from the solution.

$$Q(f) = \sum_{y_j \in Y} I(f, y_j) - \sum_{f_j \in S_k} I(f, f_j) \quad (3)$$

$$I(f, h) = H(f) + H(h) - H(f, h) \quad (4)$$

When adding a new feature, we first evaluate every feature  $f$  that is not present in the current solution, by calculating the  $Q(f)$ . Instead of adding the feature with the best (highest) value of  $Q(f)$ , we opted for a stochastic approach, inspired by the GRASP metaheuristic [6]. Therefore, we detect  $\alpha$  best features and choose one of them at random to add it to the current solution. This method is presented in Algorithm 1. Likewise, we remove a feature from the current solution in a similar manner, by calculating  $Q(f)$  for all of the features present in the current solution, detecting  $\alpha$  of those with the lowest value, and choosing one at random (Algorithm 2). The value of  $\alpha$  is provided as a hyperparameter.

**Algorithm 1** Procedure that adds a feature into the set of selected features

---

```

1: procedure ADD_FEATURE(features,  $\alpha$ )
2:    $A \leftarrow F \setminus \text{features}$ 
3:   for  $\forall f \in A$  do
4:     Evaluate  $Q(f)$ 
5:   end for
6:    $f \leftarrow$  randomly choose one of  $\alpha$  best evaluated features
7:    $\text{features} \leftarrow \text{features} \cup \{f\}$ 
8:   return features
9: end procedure

```

$\triangleright F$  is the set of all the possible features  
 $\triangleright$  Higher  $Q(f)$  indicates a better feature

---

**Algorithm 2** Procedure that removes a feature from the set of selected features

---

```

1: procedure DEL_FEATURE(features, alpha)
2:   for  $\forall f \in \text{features}$  do
3:     Evaluate  $Q(f)$ 
4:   end for
5:    $f \leftarrow$  randomly choose one of  $\alpha$  worst evaluated features
6:    $\text{features} \leftarrow \text{features} \setminus \{f\}$ 
7:   return features
8: end procedure

```

$\triangleright$  Lower  $Q(f)$  indicates a worse feature

---

In Algorithm 3, we provided an overall structure of our BCOi algorithm. First, we initialize all the bees with a starting solution (Line 2). The starting solution is generated by randomly choosing one feature and then calling *ADD\_FEATURE* procedure ( $n - 1$ ) times, where  $n$  is the preset number of features in the solution. After each bee is initialized with a solution and each solution is evaluated, the best solution among them is memorized (Line 4). The evaluation is performed by transforming the dataset so that it consists only of features present in the solution, invoking the *Multi-label K-nearest neighbors* algorithm on the testing part of the dataset, and applying the aforementioned metric to the results. The main part of the algorithm consists of two steps, repeated iteratively: forward pass (Lines 6-12) and backward pass (Lines 13-16). During the forward pass, each bee transforms its solution, evaluates it, and if it proves better than the current best solution, the newfound solution becomes the new current best solution. The transformation of a solution is done by applying *ADD\_FEATURE* function  $k$  times, followed by invoking *DEL\_FEATURE* function  $k$  times. The argument  $k$  is determined dynamically, based on the number of consecutive iterations in which none of the bees found a new best solution. The idea for this approach comes from the *Variable Neighborhood Search* metaheuristic (VNS) [12], where the size of the neighborhood is increased after each non-improving iteration. In the backward pass of the algorithm, each bee decides whether it is going to stay loyal to its solution and explore it further, or to adopt a solution from some other bee. The probability of a bee staying loyal to its solution is calculated as:

$$p_i = \frac{v_i - v_{min}}{v_{max} - v_{min}} \quad (5)$$

where  $v_i$  is the quality of the solution of the  $i$ -th bee,  $v_{max}$  and  $v_{min}$  are qualities of the best and the worst solutions among all the bees, respectively. This way, we can guarantee that at least one bee is going to stay loyal to its solution. After each bee decided on its loyalty, the bees that abandoned their solution have to choose some loyal bee and adopt its solution (Line 16). The probability of each loyal bee being selected is calculated as:

$$p_i = \frac{v_i}{\sum_{v_j \in B_l} v_j} \quad (6)$$

where  $v_i$  is the quality of the solution of the  $i$ -th bee and  $B_l$  is the set of loyal bees. Therefore, a loyal bee with a better solution has a higher chance of recruiting disloyal bees. Finally, in order to prevent the algorithm from being stuck in the local optimum, we introduce the concept of reinitialization after a certain number of iterations without any improvement (Line 23).

**Table 1:** The selected parameter values

ACO	MA	BCOi
number_of_ants = 25	population_size = 15	number_of_bees = 30
$\beta = 0.8$	$v = 500$	$\alpha = 5$
$\rho = 0.1$	$h = 15$	$k_{max} = 5$
	crossover_probability = 0.5	
	mutation_probability = 0.1	

**Algorithm 3** Bee Colony Optimization

```

1: procedure BCO(number_of_bees,  $\alpha$ ,  $k_{max}$ )
2:   bees  $\leftarrow$  initialize(number_of_bees,  $\alpha$ )
3:    $k \leftarrow 1$ 
4:   best_solution  $\leftarrow$  find_best(bees)
5:   while stopping criterion is not met do
6:     for  $\forall b \in$  bees do ▷ Forward pass
7:       transform( $b, \alpha, k$ )
8:       value  $\leftarrow$  evaluate( $b$ )
9:       if value > evaluate(best_solution) then
10:        best_solution  $\leftarrow$   $b$ 
11:      end if
12:    end for
13:    for  $\forall b \in$  bees do ▷ Backward pass
14:      decide_loyalty( $b$ )
15:    end for
16:    recruitment(bees)
17:    if there was an improvement then
18:       $k \leftarrow 1$ 
19:    else
20:       $k \leftarrow k + 1$ 
21:    end if
22:    if  $k > k_{max}$  then
23:      bees  $\leftarrow$  initialize(number_of_bees,  $\alpha$ )
24:    end if
25:  end while
26:  return best_solution
27: end procedure

```

## 5. EXPERIMENTAL EVALUATION

To evaluate our approach, we implemented three algorithms in total: ACO presented in [13], MA presented in [8], and our own method (BCOi). All three algorithms were written in Python programming language and executed on a personal laptop with an Intel i7-10750H CPU and 32GB of RAM, under the Ubuntu 20.04 operating system. The algorithms were tested on four benchmark datasets for multi-label classification, available at <https://www.uco.es/kdis/mlresources/>. Since some of the datasets contained features with continuous values, those features were first discretized by dividing the interval into 10 bins. Each test was repeated 30 times with a different value for the random number generator seed.

The values of parameters were determined by using the *iRace*<sup>1</sup> package for R programming language with a budget of 200 tests. The meaning behind the parameters for ACO and MA can be found in the papers in which these methods were originally proposed for MLFS. In Table 1 we present the obtained values for each parameter.

As a stopping criterion, we used a limit of 500 calls to the fitness function, for all three algorithms, as the evaluation procedure is highly expensive in terms of computational power, especially for larger datasets.

<sup>1</sup> <https://cran.r-project.org/web/packages/irace/index.html>



**Table 2:** Average Accuracy over 30 independent runs

Dataset	Num. instances	Num. features	Num. labels	ACO	MA	BCOi
Emotions	593	72	6	0.2574 (0.0107)	0.3054 (0.0066)	<b>0.3115 (0.0095)</b>
Birds	645	260	19	0.5108 (0.0039)	0.5012 (0.0045)	<b>0.5158 (0.0068)</b>
Yeast	2417	130	14	0.1586 (0.0037)	0.1747 (0.0073)	<b>0.1759 (0.0079)</b>
Scene	2407	294	6	0.4153 (0.0151)	0.5099 (0.0052)	<b>0.5279 (0.0110)</b>

Furthermore, we wanted to emphasize the differences in performance between metaheuristics while keeping the fitness function budget relatively low. We suggest that in practice this limit should be set higher if the user has the necessary resources, in order for algorithms to obtain better quality solutions. Larger limits have been tested on the smallest dataset *Emotions* (1000, 1500, and 2000 iterations), and we exhibited the same trend as the one presented in Figure 1.

In Table 2 we presented the obtained results for each dataset under consideration. The first column represents the name of the dataset, the second, third, and fourth show the number of instances, features, and labels respectively, whereas the three last columns present the average accuracy over 30 runs, with standard deviation presented in parenthesis. We can clearly see that BCOi outperformed the other two algorithms in terms of average accuracy for all datasets.

In Figure 1 we present a boxplot showing the performance of each algorithm over 30 independent runs for each dataset. Furthermore, we performed *Wilcoxon pair-wise statistical test* comparing BCOi with the other two algorithms, for each dataset separately. The results are presented in Table 3. Additionally, we calculated the 95% confidence interval of the location parameter for each of the tests performed and presented it below p-values in Table 3. With the significance level of  $\alpha = 0.05$ , we concluded that BCOi was statistically different from ACO in 3 out of 4 cases, and in 2 out of 4 cases compared to MA. These results support conclusions based on Figure 1.

**Figure 1** Boxplots showcasing the difference in the accuracy of tested algorithms on four datasets**Table 3:** p-values of Wilcoxon pair-wise statistical tests and 95% confidence interval for location parameter comparing BCOi to other proposed methods

Dataset	ACO	MA
Emotions	<b>0.0001717</b> [0.04438175, 0.06411505]	0.1609 [-0.00155461, 0.01450039]
Birds	0.06312 [-0.0004934487, 0.0107584934]	<b>0.0005501</b> [0.008805312, 0.020857824]
Yeast	<b>0.0003197</b> [0.009839083, 0.024617684]	0.4268 [-0.006605605, 0.007226751]
Scene	<b>0.0001242</b> [0.1034662, 0.1247954]	<b>0.001345</b> [0.01040258, 0.02699941]

## 6. CONCLUSION

In this paper, we examined metaheuristic approaches to multi-label feature selection problems. We presented a version of Bee Colony Optimization based on mutual information and compared it to methods already present in the literature, i.e. Ant Colony Optimization algorithm and Memetic algorithm. Experimental evaluation was performed on four benchmark datasets.

We demonstrated that BCO was statistically better than ACO in 3/4 cases and better than MA in 2/4 cases, while performing relatively the same in relation to the rest of the cases. Future research should include testing these algorithms on larger datasets and potentially finding an alternative to calculating the mutual information, given that it is a computationally expensive task.

## Acknowledgement

This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, Agreement No. 451-03-9/2021-14/200029 and by the Science Fund of the Republic of Serbia, Grant "AI4TrustBC: Advanced Artificial Intelligence Techniques for Analysis and Design of System Components Based on Trustworthy BlockChain Technology".

## REFERENCES

- [1] Davidović, T., Ramljak, D., Šelmić, M., & Teodorović, D. (2011). *Bee colony optimization for the p-center problem*. *Computers & Operations Research*, 38(10), 1367-1376.
- [2] Davidović, T., Teodorović, D., & Šelmić, M. (2015). *Bee colony optimization-part I: the algorithm overview*. *Yugoslav Journal of Operations Research*, 25(1), 33-56.
- [3] Davidović, T., Šelmić, M., Teodorović, D., & Ramljak, D. (2012). *Bee colony optimization for scheduling independent tasks to identical processors*. *Journal of heuristics*, 18(4), 549-569.
- [4] Dimitrijević, B., Teodorović, D., Simić, V., & Šelmić, M. (2012). *Bee colony optimization approach to solving the anticovering location problem*. *Journal of Computing in Civil Engineering*, 26(6), 759-768.
- [5] Dowlatshahi, M. B., Derhami, V., & Nezamabadi-pour, H. (2017). *Ensemble of filter-based rankers to guide an epsilon-greedy swarm optimizer for high-dimensional feature subset selection*. *Information*, 8(4), 152.
- [6] Feo, T. A., & Resende, M. G. (1995). *Greedy randomized adaptive search procedures*. *Journal of global optimization*, 6(2), 109-133.
- [7] Jungjit, S., & Freitas, A. A. (2015, April). *A new genetic algorithm for multi-label correlation-based feature selection*. In *23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (pp. 285-290).
- [8] Lee, J., & Kim, D. W. (2015). *Memetic feature selection algorithm for multi-label classification*. *Information Sciences*, 293, 80-96.
- [9] Lučić, P., & Teodorović, D. (2001, June). *Bee system: modeling combinatorial optimization transportation engineering problems by swarm intelligence*. In *Preprints of the TRISTAN IV triennial symposium on transportation analysis* (pp. 441-445).
- [10] Lučić, P., & Teodorović, D. (2002, November). *Transportation modeling: an artificial life approach*. In *14th IEEE International Conference on Tools with Artificial Intelligence, 2002.(ICTAI 2002)*. *Proceedings*. (pp. 216-223). IEEE.
- [11] Lučić, P., & Teodorović, D. (2003). *Computing with bees: attacking complex transportation engineering problems*. *International Journal on Artificial Intelligence Tools*, 12(03), 375-394.
- [12] Mladenović, N., & Hansen, P. (1997). *Variable neighborhood search*. *Computers & operations research*, 24(11), 1097-1100.
- [13] Paniri, M., Dowlatshahi, M. B., & Nezamabadi-Pour, H. (2020). *MLACO: A multi-label feature selection algorithm based on ant colony optimization*. *Knowledge-Based Systems*, 192, 105285.
- [14] Shao, H., Li, G., Liu, G., & Wang, Y. (2013). *Symptom selection for multi-label data of inquiry diagnosis in traditional Chinese medicine*. *Science China Information Sciences*, 56(5), 1-13.
- [15] Teodorović, D., Šelmić, M., & Davidović, T. (2015). *Bee Colony Optimization-part II: The application survey*. *Yugoslav Journal of Operations Research*, 25(2), 185-219.
- [16] Tsoumakas, G., Katakis, I., & Vlahavas, I. (2009). *Mining multi-label data*. *Data mining and knowledge discovery handbook*, 667-685.
- [17] Yu, Y., & Wang, Y. (2014, October). *Feature selection for multi-label learning using mutual information and GA*. In *International Conference on Rough Sets and Knowledge Technology* (pp. 454-463). Springer, Cham.
- [18] Zhang, M. L., Peña, J. M., & Robles, V. (2009). *Feature selection for multi-label naive Bayes classification*. *Information Sciences*, 179(19), 3218-3229.
- [19] Zhang, Y., Gong, D. W., Sun, X. Y., & Guo, Y. N. (2017). *A PSO-based multi-objective multi-label feature selection method in classification*. *Scientific reports*, 7(1), 1-12.

## COMPARATIVE ANALYSIS OF GREEDY CONSTRUCTIVE HEURISTICS FOR SOLVING CAPACITATED VEHICLE ROUTING PROBLEM

ANĐELA ĐORĐEVIĆ<sup>1</sup>, MILAN STANOJEVIĆ<sup>2</sup>

<sup>1</sup> University of Belgrade – Faculty of Organizational Sciences andjela.djordjevi@fon.bg.ac.rs

<sup>2</sup> University of Belgrade – Faculty of Organizational Sciences milan.stanojevic@fon.bg.ac.rs

**Abstract:** *The capacitated vehicle routing problem (CVRP) is the simplest and most studied member of the VRP family in the area of combinatorial optimization. Three different algorithms for solving CVRPs are presented in this paper. One of these algorithms is Clarke – Wright, which is a well-known savings algorithm for VRPs. The next is a modified, enhanced version of Clarke – Wright algorithm - the Enhanced savings algorithm. The last one is the Nearest Unvisited Neighbor algorithm, adapted for solving CVRP. The performance of these algorithms regarding the total route length and the required execution time were compared. The numerical reports of this research are also given in this paper.*

**Keywords:** *Capacitated Vehicle Routing Problem, Clarke – Wright Algorithm, Nearest Unvisited Neighbor, Enhanced Savings Algorithm*

### 1. INTRODUCTION

The vehicle routing problem dates back to 1959 when it was first introduced by Dantzig and Ramser [5]. At that time the problem was presented as a generalized traveling salesman problem (TSP). Namely, it was originally designed with a solution algorithm for the problem of delivering gasoline to service stations [1]. The main objective of a vehicle routing problem is to find a set of vehicle routes, with each vehicle starting and ending at the depot, in a manner that all customers' needs are fulfilled, at a minimum total route length [2] [3]. The vehicle routing problem is an *NP*-hard problem, which means that it can not be solved in polynomial time. The fact that variants of VRP are *NP*-hard was determined by Lenstra and Rinnooy Kan in 1981 [6]. Since 1959 hundreds of scientific papers were devoted to solutions for the basic version of VRP, known as the CVRP (Capacitated Vehicle Routing Problem). The only constraint that this version of VRP considers is the vehicle capacity.

CVRP is a basic version of VRP with an additional constraint on the capacity of vehicles i.e. the total demand of customers serviced by one vehicle should not exceed the capacity of the vehicle [3]. The main objective of CVRP is to find a set of routes so that each vehicle starts and ends at the depot, each customer is visited only once, and the total length of all routes is minimal [7].

In this paper, three greedy constructive heuristics for solving CVRPs are presented and compared. Those heuristics are: 1) Clarke – Wright savings algorithm (CWA) [4], 2) adapted version of Nearest unvisited neighbor algorithm (NUN) [9], and 3) Enhanced savings algorithm (ESA) [10]. The heuristics are implemented in a software that was developed in Lua programming language environment. The heuristics were tested on a set of VRP instances from the library of standard VRP problems VRPLIB [12]. These instances were generated by the members of the Operations Research Group, which is part of the Department of Electrical, Electronic, and Information Engineering Guglielmo Marconi (DEI), and can be found under the Section: Library and Codes on the above - mentioned website.

This paper is organized in the following manner. In Section 2, basic VRP characteristics are presented and capacitated VRP is explained. Section 3 describes greedy constructive heuristics that are used. The comparative analysis of the VRP using three different heuristics is presented in Section 4. Concluding thoughts are given in Section 5.

### 2. CAPACITATED VEHICLE ROUTING PROBLEM

The vehicle routing problems are frequently referred to as truck dispatching problems. The main goal is to find reliable means of obtaining optimal routes to different locations [1]. Many organizations in the fields of distribution, transportation, and logistics with complex operations are often encountering these types of problems.

VRP refers to the distribution of goods by vehicles using an existing road network to serve customers [10]. The goal is to minimize the total route length (i.e., a weighted function of the set of routes and their length or travel time) to serve all the customers. The set of routes needs to satisfy several constraints which are determined by the nature of the transported goods, the quality of the service level, and the characteristics of the customers and the vehicles [2].

The road network is defined by a graph whose vertices represent the customers and the depot. In VRP, each customer is located at one node, and each customer must be visited exactly once. The length of each arc represents the shortest paths between customers. It is assumed that the graph is complete, and arcs that connect two vertices have the same lengths [10].

In real life environment, VRP problems are more complex with the addition of several constraints. These constraints might be the capacity of a vehicle, the time windows of deliveries (or pick-ups), number of depots, number of allowed trips, nature of demands, nature of service, and many others [1]. When the vehicle capacity constraints are introduced to a basic version of VRP, CVRP is established.

The VRP that was introduced by Dantzig and Ramser in 1959 is also believed to be a capacitated VRP [8]. The CVRP is a significant and widely studied problem in the area of combinatorial optimization. The traveling salesman problem (TSP) and the bin packing problem are integrated into CVRP. The CVRP involves one depot, a homogeneous fleet of vehicles, and a set of customers located at nodes. All vehicles have the same capacity and in some types of CVRP, we assume that there is not a limited number of vehicles. To each vertex, except the depot, a demand of the corresponding customer is assigned. Length of arcs represent the cost of the transportation along them [10]. In the CVRP, the demands are deterministic, known in advance, and can not be split between different routes [2]. The vehicles start and end their visits at the central depot,

i.e. each route starts and ends at the depot [1]. The length of a route is the sum of the lengths of arcs that belong to the route [10]. The goal is to determine a feasible set of vehicle routes that minimize the total traveling distance and/or the total number of vehicles used [8]. The shortest path should be found with the satisfaction of the capacity constraint of each vehicle [3]. The CVRP deals with the maximum weight or volume that each vehicle can load [3]. In general, CVRPs are solved on a complete graph, otherwise, arcs with virtually infinite lengths are added in place of non-existent arcs.

Let  $G = (V, E)$  denote a complete and undirected graph where  $V = \{0, \dots, n\}$  is the vertex set and  $E$  is the set of arcs. Vertex set  $V_c = \{1, \dots, n\}$  corresponds to  $n$  customers, whereas vertex 0 corresponds to the depot. A nonnegative length,  $c_{ij}$  is associated with each arc  $(i, j) \in E$ . Each customer  $i \in V_c$  is associated with a known nonnegative demand,  $q_i$ , to be delivered (the depot having a fictitious demand  $q_0 = 0$ ). A set of  $m$  identical vehicles, each with capacity  $Q$ , is available at the depot. Without loss of generality we assume that  $q_i \leq Q$  for each  $i \in V_c$  [2].

The goal of CVRP is to find  $m$  simple routes with minimum total length, defined as the sum of the lengths of the arcs belonging to the routes, and such that [2]:

- each route starts and ends at the depot node,
- each customer node is visited by exactly one route, and
- the sum of the demands of the nodes visited by a route does not exceed the vehicle capacity  $Q$ .

### 3. IMPLEMENTED AND TESTED CONSTRUCTIVE HEURISTICS

In this section, the following heuristics: CWA, NUN, and ESA are going to be explained. What these three heuristics have in common is that they are all constructive and greedy.

#### 3.1. Clarke – Wright algorithm (CWA)

Clarke – Wright algorithm, also known as the savings algorithm was first introduced in 1964 [4].

The first step of this algorithm is an initialization, i.e. the elementary routes are generated, and the total route lengths are calculated for each node. Also, for every two different nodes, savings should be calculated. Savings are sorted in a nonascending order. Savings are calculated using the following formula:

$$S_{i,j} = c_{0i} + c_{0j} - c_{ij}, \forall (i,j) \in E, i \neq j. \quad (1)$$

In order to realize a saving, the nodes must belong to different routes, and the total demand of those routes must not exceed the vehicle capacity. If all requirements are met, merging of the two routes is possible. The saving is deducted from the total route length. Arcs from the old routes,  $(i, 0), (0, j)$  are eliminated and new arc,  $(i, j)$  is added.

#### 3.2. Nearest unvisited neighbor (NUN)

This heuristics was originally developed for TSP. For this research, a modified version of the nearest unvisited neighbor algorithm adapted for CVRP is used. The algorithm starts from the depot and the distances between the depot and all nodes (neighbors) are considered. Then, if a vehicle capacity constraint is satisfied, a route is constructed with the depot as a starting point and the nearest node (neighbor) as the next point. The next step is checking distances from the current nearest node and the remainder of the nodes. The next node in the route is the node (neighbor) that is the closest to the current one and that does not exceed the vehicle capacity constraint. Nodes are added to the route as long as the vehicle capacity

requirement is met. The ending point in the route is the same as the starting point – the depot. When a route is constructed, nodes that belong to that route are eliminated from the set of unvisited neighbors. When vehicle capacity constraint is exceeded, the algorithm goes back to the first step. The algorithm runs as long as there are unvisited neighbors.

The nearest unvisited neighbor algorithm is not reliable for uncomplete graphs. It can result in an unallowable solution.

### 3.3. Enhanced savings algorithm (ESA)

The enhanced savings algorithm is a new, modified version of the Clarke – Wright savings algorithm. Enhanced savings can be used in many algorithms where standard savings are typically used. ESA combines the enhanced savings calculation with a dynamically updated savings list [10].

In comparison to CWA, ESA allows dynamic recalculation of some savings in each iteration. In ESA, two routes  $q = (0, \dots, i, j, \dots, 0)$  and  $r = (0, k, \dots, l, 0)$  are merged into the route  $s = (0, \dots, i, k, \dots, l, j, \dots, 0)$ . The arc  $a = (i, j)$  is replaced by the path  $(i, k, \dots, l, j)$ . The enhanced saving obtained by such a merging is:

$$S_{r,a} = c_{ij} + c_{0k} + c_{l0} - c_{ik} - c_{lj}. \quad (2)$$

The merging used in CWA is limited to the cases where the arc  $a$  is the first or the last arc of route  $q$ , thus it has a form similar to:

$$S_{r,a} = c_{i0} + c_{0k} - c_{ik}. \quad (3)$$

The formula (3) is used for a CWA merging in which  $a$  is the last arc in  $q$ , and  $j = 0$  [10].

Both CWA and ESA have elements of local search. Due to enhanced savings and mergings, ESA searches larger neighborhoods than CWA in each iteration [10].

## 4. COMPARATIVE ANALYSIS

The above-mentioned heuristics were implemented in software for solving VRP. The software was developed in the Lua programming language environment. Lua is extremely fast for an interpreted language. It has many characteristics of modern programming languages like: first-class functions, closure, co-routines, etc. It is very easy to use it with other programming languages, such as C/C++. Lua provides meta-mechanisms for implementing the needed features, instead of incorporating features within the language [11].

CWA, NUN, and ESA were compared by the time required for the execution, and by the total route length (TRL) that they create. The results of this comparison are presented in Table 1.

**Table 1:** CWA, NUN, and ESA – comparative analysis

Instances	CWA		NUN		ESA		CWA vs NUN	ESA vs CWA
	TRL	Time	TRL	Time	TRL	Time		
1	2.1	2.2	3.1	3.2	4.1	4.2	5	6
A-n32-k5.vrp	832	0.013	1145	0.014	832	0.023	37.62%	0.00%
A-n33-k5.vrp	691	0.013	977	0.013	693	0.028	41.39%	-0.29%
A-n33-k6.vrp	766	0.018	1042	0.018	745	0.033	36.03%	2.82%
A-n34-k5.vrp	804	0.024	982	0.018	803	0.047	22.14%	0.12%
A-n36-k5.vrp	806	0.023	1077	0.019	835	0.038	33.62%	-3.47%
A-n37-k5.vrp	706	0.022	948	0.019	702	0.037	34.28%	0.57%
A-n37-k6.vrp	1000	0.026	1334	0.024	970	0.039	33.40%	3.09%
A-n38-k5.vrp	767	0.027	1084	0.022	763	0.039	41.33%	0.52%
A-n39-k5.vrp	900	0.025	1070	0.022	907	0.046	18.89%	-0.77%
A-n39-k6.vrp	855	0.024	1145	0.024	874	0.039	33.92%	-2.17%
A-n44-k6.vrp	975	0.030	1391	0.022	1005	0.049	42.67%	-2.99%
A-n45-k6.vrp	986	0.033	1485	0.025	966	0.050	50.61%	2.07%
A-n45-k7.vrp	1203	0.034	1428	0.025	1217	0.050	18.70%	-1.15%
A-n46-k7.vrp	940	0.034	1335	0.022	939	0.051	42.02%	0.11%
A-n48-k7.vrp	1110	0.036	1475	0.026	1110	0.057	32.88%	0.00%
A-n53-k7.vrp	1087	0.036	1444	0.025	1128	0.066	32.84%	-3.63%
A-n54-k7.vrp	1175	0.038	1421	0.028	1206	0.069	20.94%	-2.57%

A-n55-k9.vrp	1109	0.043	1500	0.031	1116	0.071	35.26%	-0.63%
A-n60-k9.vrp	1555	0.044	2033	0.032	1555	0.090	30.74%	0.00%
A-n61-k9.vrp	1094	0.045	1377	0.03	1045	0.083	25.87%	4.69%
A-n62-k8.vrp	1343	0.043	1849	0.031	1336	0.090	37.68%	0.52%
A-n63-k10.vrp	1395	0.049	1799	0.033	1423	0.099	28.96%	-1.97%
A-n63-k9.vrp	1656	0.047	2173	0.031	1652	0.099	31.22%	0.24%
A-n64-k9.vrp	1619	0.048	1968	0.033	1620	0.103	21.56%	-0.06%
A-n65-k9.vrp	1005	0.048	1466	0.037	1010	0.096	45.87%	-0.50%
A-n69-k9.vrp	1000	0.054	1456	0.037	1013	0.108	45.60%	-1.28%
A-n80-k10.vrp	1930	0.068	2335	0.035	1929	0.151	20.98%	0.05%
B-n31-k5.vrp	681	0.019	887	0.019	693	0.034	30.25%	-1.73%
B-n34-k5.vrp	806	0.021	886	0.02	806	0.033	9.93%	0.00%
B-n35-k5.vrp	981	0.024	1304	0.021	978	0.032	32.93%	0.31%
B-n38-k6.vrp	837	0.024	1268	0.121	835	0.039	51.49%	0.24%
B-n39-k5.vrp	566	0.026	1058	0.024	567	0.049	86.93%	-0.18%
B-n41-k6.vrp	894	0.028	1080	0.022	901	0.042	20.81%	-0.78%
B-n43-k6.vrp	757	0.025	965	0.023	792	0.047	27.48%	-4.42%
B-n44-k7.vrp	934	0.032	1274	0.024	934	0.049	36.40%	0.00%
B-n45-k5.vrp	755	0.027	1051	0.024	752	0.047	39.21%	0.40%
B-n45-k6.vrp	725	0.032	1001	0.023	722	0.050	38.07%	0.42%
B-n50-k7.vrp	751	0.035	1036	0.027	751	0.063	37.95%	0.00%
B-n50-k8.vrp	1354	0.036	1527	0.031	1396	0.060	12.78%	-3.01%
B-n51-k7.vrp	1035	0.039	1333	0.028	1083	0.065	28.79%	-4.43%
B-n52-k7.vrp	763	0.034	1276	0.028	760	0.060	67.23%	0.39%
B-n56-k7.vrp	732	0.038	1248	0.027	730	0.073	70.49%	0.27%
B-n57-k7.vrp	1238	0.040	1866	0.029	1242	0.077	50.73%	-0.32%
B-n57-k9.vrp	1579	0.044	1877	0.03	1597	0.102	18.87%	-1.13%
B-n63-k10.vrp	1580	0.049	2123	0.033	1600	0.118	34.37%	-1.25%
B-n64-k9.vrp	1104	0.049	1297	0.037	1043	0.097	17.48%	5.85%
B-n66-k9.vrp	1195	0.048	1485	0.038	1168	0.100	24.27%	2.31%
B-n67-k10.vrp	1254	0.051	1891	0.035	1311	0.100	50.80%	-4.35%
B-n68-k9.vrp	1412	0.054	1856	0.031	1494	0.109	31.44%	-5.49%
B-n78-k10.vrp	1241	0.063	1728	0.037	1244	0.132	39.24%	-0.24%
E-n101-k14.vrp	1142	0.102	1658	0.05	1132	0.216	45.18%	0.88%
E-n101-k8.vrp	897	0.094	1176	0.047	884	0.209	31.10%	1.47%
E-n22-k4.vrp	387	0.012	464	0.019	387	0.017	19.90%	0.00%
E-n23-k3.vrp	651	0.011	760	0.016	577	0.018	16.74%	12.82%
E-n30-k3.vrp	533	0.018	627	0.018	532	0.025	17.64%	0.19%
E-n33-k4.vrp	843	0.017	1010	0.022	843	0.037	19.81%	0.00%
E-n76-k10.vrp	879	0.064	1221	0.037	896	0.133	38.91%	-1.90%
E-n76-k14.vrp	1062	0.070	1353	0.043	1068	0.130	27.40%	-0.56%
E-n76-k7.vrp	743	0.057	1098	0.034	724	0.123	47.78%	2.62%
E-n76-k8.vrp	794	0.063	1081	0.042	778	0.123	36.15%	2.06%
P-n101-k4.vrp	755	0.093	1021	0.04	752	0.200	35.23%	0.40%
P-n16-k8.vrp	482	0.014	497	0.021	482	0.017	3.11%	0.00%
P-n19-k2.vrp	237	0.033	250	0.014	224	0.017	5.49%	5.80%
P-n20-k2.vrp	237	0.029	309	0.015	234	0.016	30.38%	1.28%
P-n21-k2.vrp	233	0.013	257	0.014	236	0.015	10.30%	-1.27%
P-n22-k2.vrp	237	0.012	261	0.016	251	0.016	10.13%	-5.58%
P-n22-k8.vrp	590	0.017	801	0.021	590	0.025	35.76%	0.00%
P-n23-k8.vrp	537	0.020	746	0.022	537	0.023	38.92%	0.00%
P-n40-k5.vrp	525	0.024	682	0.023	520	0.042	29.90%	0.96%
P-n45-k5.vrp	571	0.032	755	0.022	571	0.048	32.22%	0.00%
P-n50-k10.vrp	760	0.037	907	0.032	758	0.061	19.34%	0.26%
P-n50-k7.vrp	587	0.034	794	0.026	585	0.059	35.26%	0.34%

P-n50-k8.vrp	651	0.037	820	0.028	650	0.065	25.96%	0.15%
P-n51-k10.vrp	771	0.043	904	0.031	774	0.064	17.25%	-0.39%
P-n55-k10.vrp	735	0.042	983	0.03	731	0.069	33.74%	0.55%
P-n55-k15.vrp	992	0.048	1217	0.035	992	0.077	22.68%	0.00%
P-n55-k7.vrp	628	0.037	754	0.029	623	0.072	20.06%	0.80%
P-n55-k8.vrp	637	0.052	828	0.026	614	0.071	29.98%	3.75%
P-n60-k10.vrp	527	0.058	696	0.037	526	0.071	32.07%	0.19%
P-n60-k15.vrp	1022	0.062	1170	0.041	1027	0.084	14.48%	-0.49%
P-n65-k10.vrp	815	0.052	1116	0.031	828	0.097	36.93%	-1.57%
P-n70-k10.vrp	867	0.055	1010	0.032	891	0.106	16.49%	-2.69%
P-n76-k4.vrp	642	0.056	833	0.032	656	0.156	29.75%	-2.13%
P-n76-k5.vrp	697	0.058	962	0.033	691	0.117	38.02%	0.87%
<b>AVG</b>	/	<b>0.039</b>	/	<b>0.029</b>	/	<b>0.071</b>	<b>31.66%</b>	<b>-0.06%</b>

Column 1 contains instances our software was tested on. Columns 2.1 and 2.2 contain obtained results regarding total route length and required time for execution, using the CWA in the software, respectively. The same goes for Columns 3.1 and 3.2, and 4.1 and 4.2, the difference is in the used algorithm. Last rows of Columns 2.2, 3.2, and 4.2 represent average execution times for CWA, NUN, and ESA, respectively. According to Table 1, NUN has the least execution time. The second is CWA.

Column 5 compares the performances of CWA and NUN concerning the obtained total route lengths given in Columns 2.1 and 3.1. The results in Column 5 were calculated using the following formula:

$$\frac{\text{Column 3.1} - \text{Column 2.1}}{\text{Column 2.1}} \quad (3)$$

The last row of Column 5 shows that on this set of instances CWA obtained 31.66% better results than NUN, concerning the total route length.

The total route lengths obtained using ESA and CWA are compared in Column 6, using the following formula:

$$\frac{\text{Column 2.1} - \text{Column 4.1}}{\text{Column 4.1}} \quad (4)$$

Since the result in the last row of Column 6 is negative (-0.06%), that means that CWA performed better on this set of instances regarding total route length than ESA.

In the comparison to total route length, CWA is significantly better than NUN. Since it was assumed that ESA has significantly better performances than NUN, too, these algorithms were not compared. This research has proven that NUN has better results in terms of execution time, because it is a fairly simple algorithm, while in terms of route length it gives much worse results than ESA and CWA. The initial hypothesis was that ESA has better performances than CWA. However, based on the results from Table 1, CWA is better than ESA, compared to the execution time and total route length. For 13 instances (15.48%), CWA and ESA give the same results in terms of total route length. For 37 instances (44.05%), ESA gives better results in terms of total route length.

## 5. CONCLUSION

In this paper, three greedy constructive heuristics for solving CVRPs are presented. Their performances regarding CVRP were tested in the software we created in the Lua programming environment.

Besides being greedy and constructive, the Clarke-Wright algorithm and the Enhanced savings algorithm both have elements of local search. CWA calculates all savings just in the initialization step, while in ESA, the savings list is updated after each merging. CWA and ESA have the same computational complexity  $O(n^2 \log n)$ . A Modified version of the Nearest unvisited neighbor for CVRP that was used in this research has complexity  $O(n^2)$ . The  $n$  stands for the number of customers in CVRP.

The three above-mentioned algorithms were tested on a set of instances for CVRP. We can conclude that NUN has the least execution time, which was expected due to its simplicity. CWA has a similar execution time, while ESA has a significantly longer execution time. Regarding obtained total route length results, NUN has the worst performance. Compared to CWA and ESA, NUN achieves drastically worse results. On average, CWA achieves slightly better results than ESA.

If the required execution time is crucial no matter the quality of the results regarding the total route length, NUN can be the algorithm of choice. On the other hand, if the required execution time is not as important as the obtained total route length we can conclude that CWA has the best performance on the set of instances, that heuristics were tested on.

## REFERENCES

- [1] Adewumi, A.O., Adeleke, O.J. (2016). A survey of recent advances in vehicle routing problems. *Int J Syst Assur Eng Manag*, DOI 10.1007/s13198-016-0493-4.
- [2] Baldacci, R., Toth, P., Vigo, D. (2010). Exact algorithms for routing problems under vehicle capacity constraints. *Ann Oper Res*, 175: 213–245 DOI 10.1007/s10479-009-0650-0.
- [3] Carwalo, T., Thankappan, J., Patil, V. (2017). Capacitated Vehicle Routing Problem. *2017 2nd International Conference on Communication Systems, Computing and IT Applications (CSCITA)*, 978-1-5090-4381-1/17/\$31.00 © 2017 IEEE.
- [4] Clarke, G., Wright, J.V. (1964). Scheduling of vehicles from a central depot to a number of delivery points. *Oper. Res.* 12 (1964) 568–581.
- [5] Dantzig, G.B., Ramser, J.H. (1959). The truck dispatching problem. *Manag. Sci.* 6 (1) (1959) 80–91.
- [6] Lenstra J.K., Rinnooy Kan A.H.G. (1981). Complexity of vehicle routing and scheduling problems. *Networks*, 11(2):221–227.
- [7] Mamat, N.J.Z., Jaaman, S.H., Ahmad, R.R. (2016). Vehicle routing problem and capacitated vehicle routing problem frameworks in fund allocation problem. *AIP Conference Proceedings*, 1784, 050001; DOI: 10.1063/1.4966820.
- [8] Pichpibul, T., Kawtummachai, R. (2013). A Heuristic Approach Based on Clarke-Wright Algorithm for Open Vehicle Routing Problem. *The Scientific World Journal*, Volume 2013, Article ID 874349, 11 pages.
- [9] Raya, L., Saud, S.N. (2020). A Comparative Study Between The Nearest – Neighbour Algorithm And Its Variants For Solving The Euclidean Traveling Salesman Problem. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 17(10), 938-945 ISSN: 1567-214x.
- [10] Stanojević, M., Stanojević, B., Vujošević, M. (2013). Enhanced savings calculation and its applications for solving capacitated vehicle routing problem. *Applied Mathematics and Computation*, 219:10302–10312.
- [11] Stanojević, M., Stanojević, B. (2020). Mathematical Optimization in Lua Programming Language Environment. *XLVII International Symposium on Operational research - SYM-OP-IS 2020*, ISBN: 978-86-7395-429-5.
- [12] Vigo, D., Operations Research Group Bologna – VRPLIB: A Vehicle Routing Problem LIBrary. <http://or.dei.unibo.it/library/vrplib-vehicle-routing-problem-library>, last access: 21.06.2022.



## UNAPREĐENA METODA PROMENLJIVIH OKOLINA ZA FAZI KLASTEROVANJE NA KOMPLEKSNIM MREŽAMA IMPROVED VARIABLE NEIGHBORHOOD SEARCH FOR FUZZY CLUSTERING ON COMPLEX NETWORKS

FILIP VIDOJEVIĆ<sup>1</sup>, LAZAR MRKELA<sup>2</sup>, DUŠAN DŽAMIĆ<sup>2</sup>, MIROSLAV MARIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Matematički fakultet, filip\_vidojevic@matf.bg.ac.rs

<sup>2</sup> Univerzitet Metropolitan – Fakultet za informacione tehnologije, lazar.mrkela@metropolitan.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, dusan.dzamic@fon.bg.ac.rs

<sup>4</sup> Univerzitet u Beogradu – Matematički fakultet, maricm@matf.bg.ac.rs

**Rezime:** U okviru ovog rada predstavljena je nova VNS metoda za rešavanje problema fazi klasterovanja na kompleksnim mrežama. Za razliku od disjunktog klasterovanja, gde jedan čvor može samo u celini pripadati nekom klasteru, kod fazi klasterovanja čvor može delimično pripadati različitim klasterima. Dakle, dimenzija rešenja problema fazi klasterovanja je nekoliko puta veća nego u slučaju disjunktog klasterovanja, što dodatno usložnjava problem klasterovanja na kompleksnoj mreži. Kao lokalna pretraga, implementirana je efikasna metoda brze optimizacije fazi modularnosti, koja se u literaturi pokazala kao metoda sa najmanjim vremenom izvršavanja do konvergencije. U fazi razmrđavanja, implementirane su i upoređene tri različite okoline, zasnovane na slučajnom izboru određenog broja čvorova čije se pripadnosti klasterima menjaju. Razvijena metoda testirana je na poznatim skupovima podataka Američki univerzitetski fudbal, Zaharijev karate klub i Mreži delfina. Eksperimentalni rezultati su pokazali da kombinacija sve tri okoline u fazi razmrđavanja daje najbolje rezultate na svim skupovima podataka.

**Ključne reči:** Kompleksne mreže, fazi klasterovanje, modularnost, metoda promenljivih okolina

**Abstract:** In this paper, we present a new VNS method for fuzzy clustering on complex networks. Unlike disjoint clustering, where one node can only entirely belong to a cluster, in fuzzy clustering a node can partially belong to different clusters. Thus, the solution dimensionality of the fuzzy clustering algorithms is several times larger than in the case of disjoint clustering, which further complicates the clustering problem on complex networks. As a local search, an efficient method of fast fuzzy modularity optimization was implemented, which proved in the literature to be the method with the lowest execution time until convergence. In the shaking phase, three different environments were implemented and compared, based on a random selection of a number of nodes whose membership degrees change. The developed method was tested on the well-known data sets American University Football, Zachary Karate Club and Dolphin Network. Experimental results have shown that the combination of all three neighborhoods in the shaking phase gives the best results on all data sets.

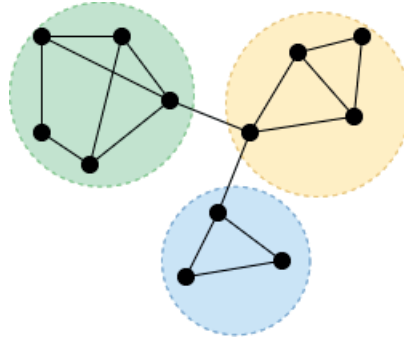
**Keywords:** Complex networks, fuzzy clustering, modularity, variable neighborhood search

### 1. UVOD

Potreba za klasterovanjem javlja se u različitim sferama života, počev od analize društvenih mreža i nalaženja zajednica u njima, do analize metaboličkih mreža i određivanja funkcija proteina. Kada karakteristike i ponašanje mreže ne mogu biti izvedeni iz karakteristika i ponašanja njenih pojedinačnih komponenti, mrežu nazivamo kompleksnom. Te komponente mogu biti predstavljene čvorovima grafa, gde ivice predstavljaju njihove interakcije. Mapiranje mreže na graf omogućava primenu teorije grafova i nauke o mrežama [3].

U kompleksnim mrežama, često je moguće uočiti grupe čvorova koji su više povezani među sobom nego sa ostatkom mreže (slika 1). Takve grupe čvorova nazivamo klasterima (zajednicama, modulima). Čvorovi koji pripadaju istom klasteru uglavnom imaju ista svojstva ili uloge. Na primer, u društvenim mrežama ljudi u istoj zajednici mogu imati istu karijeru, isti hobi, isto državljanstvo, i sl. Međutim, jedna osoba može imati više hobija ili dva državljanstva. Stoga, javlja se potreba za otkrivanjem preklapajućih klastera u kompleksnim mrežama.

Oblik klasterovanja u kojem svaki čvor može istovremeno pripadati većem broju klastera naziva se fazi (preklapajuće) klasterovanje, pri čemu se čvor koji je delimično sadržan u dva ili više klastera naziva fazi



Slika 1 Klasteri u grafu.

(preklapajućim) čvorom. Fazi klasterovanje je posebno važno u situacijama kada mreža sadrži ivicu koja predstavlja most, odnosno vezu između dva ili više klastera.

Iako na prvi pogled intuitivan, klasterovanje na grafu nije precizno definisan problem. Postoji mnogo primera kompleksnih mreža u realnom svetu, pri čemu svaka od njih ima određene specifičnosti koje je razlikuju od ostalih. Stoga, mnoge metode klasterovanja su bazirane na ovim specifičnostima, pa je gotovo nemoguće dati univerzalnu definiciju klasterovanja [5].

## 2. FAZI KLASTEROVANJE NA KOMPLEKSNIM MREŽAMA

### 2.1. Pojam klastera i particija u fazi klasterovanju

Da bi bilo moguće identifikovati preklapajuće klasterove u grafu, neophodno je definisati koncepte klastera i particija grafa u fazi slučaju. Za razliku od disjunktog slučaja, gde je jedno rešenja vektor oznaka klastera, kod fazi klasterovanja rešenje predstavlja matrica  $c \times n$ , gde je  $c$  broj klastera u grafu. Stoga, dimenzija rešenja značajno raste, pa je potrebno razviti efikasne algoritme sa velikom brzinom izvršavanja, kao i strukture podataka koje će ti algoritmi koristiti. Za reprezentaciju kompleksne mreže, korišćićemo graf  $G = (V, E)$ ,  $|V| = n$ ,  $|E| = m$ , gde je  $V$  skup čvorova i  $E$  skup ivica grafa  $G$ .

Neka je  $c$  ceo broj,  $1 < c < n$ , gde je  $n$  broj čvorova grafa  $G$ .  $C$ -particija skupa  $V$  je skup od  $c \cdot n$  vrednosti  $u_{ki}$  koje formiraju matricu dimenzije  $c \times n$ . Postoji tri skupa  $c$ -particija:

$$M_{pcn} = \left\{ \mathbf{U} \in \mathbb{R}^{c \times n}; 0 \leq u_{ki} \leq 1, \forall k, i; \sum_{k=1}^c u_{ki} > 0, \forall i; 0 < \sum_{i=1}^n u_{ki} < n, \forall k \right\}, \quad (1)$$

$$M_{fcn} = \left\{ \mathbf{U} \in M_{pcn}; \sum_{k=1}^c u_{ki} = 1, \forall i \right\}, \quad (2)$$

$$M_{hcn} = \left\{ \mathbf{U} \in M_{fcn}; u_{ki} \in \{0, 1\}, \forall k, i \right\}. \quad (3)$$

Izrazima (1) - (3) definisani su skupovi, redom, *nedegenerisanih posibilističkih*, *fazi (probabilističkih)* i *strogih*  $c$ -particija skupa  $V$  [6]. Može se uočiti da važi  $M_{hcn} \subset M_{fcn} \subset M_{pcn}$ . Iako se slučaj disjunktog klasterovanja definiše zasebno, može se posmatrati kao specijalan slučaj prethodne definicije (ukoliko se particije ograniče na skup  $M_{hcn}$ ).

Kada su uvedene neophodne definicije, problem pronalaženja preklapajućih klastera podrazumeva određivanje  $c$ -particija  $\mathbf{U} \in M_{fcn}$ . Postavlja se pitanje: koja od particija iz skupa  $M_{fcn}$  najbolje opisuje raspodelu ivica u grafu  $G$ ? Da bismo dali odgovor na to pitanje, potreban je mehanizam za određivanje kvaliteta particija.

### 2.2. Modularnost

Neka je  $G = (V, E)$  graf i  $\mathbf{U}$  jedna  $c$ -particija skupa  $V$ . Funkcija  $f : M_{pcn} \rightarrow \mathbb{R}$  naziva se ocenom kvaliteta particija iz  $M_{pcn}$ . Ideja iza ocena kvaliteta ogledase u dodeljivanju većih vrednosti boljim particijama. Međutim, često se događa da ocena kvaliteta dodeli veće vrednosti lošijim particijama. Stoga, u literaturi postoji puno ocena kvaliteta (modularnost, sprovodljivost,  $E$ -funkcija [3, 4]).

Jedna od najpopularnijih ocena kvaliteta je Njuman-Girvanova modularnost, koja je inicijalno predložena 2004 [1]. Smisao funkcije modularnosti ogleda se u poređenju unutrašnje gustine ivica klastera  $C$  i očekivane gustine ivica u nekim nultim modelima grafova, odnosno grafovima sa veštački generisanom unutrašnjom strukturom. Nulti modeli nemaju tendenciju ka formiranju klastera, pa služe za testiranje novih metoda klasterovanja [5].

Neka je  $G = (V, E)$  graf i  $C_1, C_2, \dots, C_c$  disjunktni neprazni podskupovi skupa  $V$ , čija je unija skup  $V$ . Modularnost je data jednakošću:

$$Q = \sum_{k=1}^c \left[ \frac{l_k}{m} - \left( \frac{s_k}{2m} \right)^2 \right], \quad (4)$$

gde je  $l_k$  ukupan broj ivica unutar klastera  $k$ , a  $s_k$  suma stepena svih čvorova klastera  $k$ . U jednakosti 4, prvi deo izraza unutar velikih zagrada predstavlja odnos između broja ivica unutar klastera i ukupnog broja ivica u grafu, dok drugi deo predstavlja očekivanu vrednost istog odnosa u grafu sa čvorovima istog stepena a slučajno postavljenim granama.

Navedena definicija modularnosti podrazumeva strogu podelu na podskupove  $C_i, i = 1, 2, \dots, c$ , tj. klaster  $C_i$  sadrži čvor u potpunosti, ili ga ne sadrži uopšte. Postavlja se pitanje kakvu ulogu imaju stroge  $c$ -particije u izračunavanju modularnosti [6]. Havens i saradnici su proširili definiciju modularnosti na fazi slučaj i dobili:

$$Q_g = \frac{\text{tr}(\mathbf{UBU}^T)}{\|\mathbf{A}\|}, \quad \mathbf{U} \in M_{pcn}, \quad (5)$$

gde je  $B$  matrica modularnosti (definicija i svojstva se mogu naći u [2]) i  $A$  matrica susedstva grafa  $G$ . Dodatno, ukoliko je matrica  $U$  fazi  $c$ -particija,  $Q_g$  se naziva fazi modularnost matrice  $U$ .

### 3. METODA PROMENLJIVIH OKOLINA ZA FAZI KLASTEROVANJE NA KOMPLEKSNIM MREŽAMA

Metoda promenljivih okolina predstavlja metaheurističku metodu vođenu jednim rešenjem koju je predložio Mladenović 1995. godine [10]. Glavni deo algoritma sastoji se u definisanju niza okolina  $\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_{k_{max}}$ . U svakoj iteraciji algoritma, iz okoline  $\mathcal{N}_k$  bira se jedno slučajno rešenje, koje postaje polazno rešenje za fazu lokalne pretrage. Nakon lokalne pretrage, lokalni minimum se upoređuje sa trenutno najboljim rešenjem. Ukoliko je trenutno rešenje bolje, ono postaje novo najbolje rešenje i pretraga se vraća u prvu okolinu. U suprotnom, prelazi se u sledeću iteraciju algoritma.

Modifikacija metode promenljivih okolina za slučaj fazi klasterovanja data je algoritmom 1. Na početku, generiše se početno rešenje na slučajan način i izračunava modifikovana matrica modularnosti grafa  $G$  koja je neophodna za fazu lokalne pretrage. U svakoj iteraciji, vrši se razmrđavanje trenutnog rešenja, tako što se na slučajan način bira  $k$  čvorova čiji se stepeni pripadanja klasterima menjaju. Način promene stepena pripadanja zavisi od tipa okoline, koji se dobija kao parametar algoritma. Definisani su sledeći tipovi okolina:

- $\mathcal{N}_1(\mathbf{U}, k, \text{Swap})$  - sastoji se od svih particija koje se dobijaju zamenom stepena pripadanja za dva slučajno izabrana klastera, svakom od  $k$  slučajno izabranih čvorova.
- $\mathcal{N}_2(\mathbf{U}, k, \text{Normal})$  - sastoji se od svih particija koje se dobijaju generisanjem slučajnog vektora iz multivarijantne logističke normalne raspodele, za svaki od  $k$  slučajno izabranih čvorova. Pomenuta normalna raspodela kao srednju vrednost koristi vektor pripadanja trenutno izabranog čvora, dok je varijansa pozitivno semidefinitna matrica  $\Sigma$  i predstavlja parametar algoritma. Matrica  $\Sigma$  se može dobiti kao proizvod  $AA^T$ , pri čemu je  $A$  kvadratna matrica dimenzije  $(c-1) \times (c-1)$  i sastoji od vrednosti iz intervala  $[0, 1]$ . Na taj način dobijaju se vektori koji su "blizu" polaznom vektoru.
- $\mathcal{N}_3(\mathbf{U}, k, \text{Random})$  - sastoji se od svih particija koje se dobijaju zamenom vektora stepena pripadanja slučajno generisanim vektorom, za svaki od  $k$  slučajno izabranih čvorova.

Pored tri pomenute okoline, implementiran je i slučaj kombinacije sve tri okoline (*Combined*). Prvo se bira rešenje iz jedne okoline i vrši se lokalna pretraga. Ukoliko nema poboljšanja, prelazi se na sledeću, za isto  $k$ . Okoline su birane u redosledu *Swap*, *Normal*, *Random*. Ukoliko posle sve tri okoline nema poboljšanja,  $k$  se povećava za 1.

U fazi lokalne pretrage, implementiran je algoritam brze maksimizacije fazi modularnosti (engl. *Fast Fuzzy Modularity Maximization*) [9], koji iterativno poboljšava trenutno rešenje tako što maksimizuje fazi modularnost  $Q_g$ . S obzirom da se lokalni optimum dostigne u malom broju iteracija, ovaj algoritam je poslužio kao idealan kandidat za fazu lokalne pretrage. Parametri algoritma lokalne pretrage su modifikovana matrica modularnosti ( $\tilde{\mathbf{B}}$ ), kao i broj iteracija jedne lokalne pretrage ( $I_f$ ).

**Algoritam 1:** *Fazi-VNS*


---

**Ulaz:** Graf:  $G$ ; Fazi modularnost:  $Q_g$ ; Parametri:  $I_v, I_f, Okoline, K_{max}$   
**Izlaz:** Particiona matrica:  $U$

```

1  $\tilde{B} \leftarrow$  Modifikovana matrica modularnosti grafa  $G$ ;
2  $U \leftarrow$  SlučajnoRešenje();
3  $k \leftarrow 1$ ;
4  $vreme \leftarrow 0$ ;
5 while  $vreme < n/10$  do
6   for  $k = 1, 2, \dots, K_{max}$  do
7     for  $t$  in  $Okoline$  do
8        $U' \leftarrow$  Razmrdavanje( $U, k, t$ );
9        $U' \leftarrow$  FFMM( $U', \tilde{B}, I_f$ );
10      if  $Q_g(U') > Q_g(U)$  then
11         $U \leftarrow U'$ ;
12         $k \leftarrow 1$ ;
13        break;
14      end
15    end
16  end
17  AžurirajVreme( $vreme$ );
18 end
19 return  $U$ 

```

---

#### 4. EKSPERIMENTALNI REZULTATI

Za testiranje razvijene metode korišćeni su poznati skupovi podataka:

- Zaharijev karate klub - mreža koja opisuje odnose između članova karate kluba, pri čemu čvorovi predstavljaju članove kluba, a ivice njihov odnos, ukoliko postoji. Mreža se sastoji od 34 čvorova i 78 ivica [1].
- Mreža delfina - mreža delfina i njihovih interakcija u fjordovima na Novom Zelandu u intervalu od 7 godina. Ova mreža se sastoji od 62 čvora i 159 grana [7].
- Američki univerzitetski fudbal - mreža sa 115 čvorova i 613 grana, pri čemu čvorovi predstavljaju timove, a grane odigrane utakmice između timova [8].

U tabeli 1 su prikazani rezultati na pomenutim mrežama, za svaki tip okoline pojedinačno, kao i za kombinaciju sve tri okoline. Vrednosti fazi modularnosti su date kao prosek 50 pokretanja algoritma za svaki tip okoline zasebno, kao i za kombinovani slučaj (najbolje vrednosti su podebljane). Parametar lokalne pretrage, odnosno broj iteracija algoritma brze maksimizacije fazi modularnosti je postavljen na 3. U cilju korektnog poređenja, vreme izvršavanja svakog od algoritama ograničeno je na  $n/10$ . Parametar  $K_{max}$  je postavljen na 4. Za parametar matrice kovarijanse, uzeta je vrednost 0.5.

Očekivano, faza razmrdavanja značajno utiče na rezultat klasterovanja. Takođe, iz rezultata se može zaključiti da kombinovana okolina daje najbolje rezultate na svim mrežama. Ukoliko posmatramo okoline pojedinačno, najbolje rezultate u proseku daje *Swap* okolina.

**Tabela 1:** Rezultati implementirane metode na različitim mrežama. Prikazano je poređenje različitih okolina sa njihovom kombinacijom, kao i rezultati koje obična lokalna pretraga postiže.

Mreža	LocalSearch	Random	Swap	Normal	Combined
Zaharijev karate klub	0.239681	0.42234	0.422827	0.399778	<b>0.425541</b>
Mreža delfina	0.492774	0.506456	0.516572	0.509999	<b>0.518944</b>
Američki koledž fudbal	0.509527	0.518089	0.528	0.524368	<b>0.528754</b>
Prosek	0.413994	0.482295	0.489133	0.478048	<b>0.491079</b>

#### 5. ZAKLJUČAK

U ovom radu je implementirana metoda promenljivih okolina za fazi klasterovanje na kompleksnim mrežama. Definisano je 3 tipa okolina, kao i njihova kombinacija. Novorazvijena metoda testirana je na poznatim skupovima podataka (Američki univerzitetski fudbal, Zaharijev karate klub, Mreža delfina). Pored međusobnog poređenja, predložene okoline upoređene su i sa algoritmom koji se sastoji samo od lokalne pretrage. Pokazalo se da kombinacija sve tri definisane okoline postiže najbolje rezultate na svim pomenutim skupovima podataka.

Buduća istraživanja mogu da uključe efikasniju implementaciju faze lokalne pretrage, radi testiranja na kompleksnim mrežama velikih dimenzija. Takođe, dalji rad može da obuhvati i implementaciju drugih metaheuristika poput GRASP metode.

## LITERATURA

- [1] Girvan, M., & Newman, M. E. J., (2004). Finding and evaluating community structure in networks. *Physical Review*.
- [2] Newman, M., (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review*, 74 3 Pt 2:036104.
- [3] Džamić, D., (2021). Nove metode klasterovanja na kompleksnim mrežama. Doktorska disertacija, Matematički fakultet, Univerzitet u Beogradu.
- [4] Džamić, D., Pei, J., & Marić, M., (2018). Exponential quality function for community detection in complex networks. *International Transactions in Operational Research*.
- [5] Fortunato, S., (2009). Community detection in graphs. *Physics Reports*.
- [6] Havens, T., Bezdek, J., Leckie, C., Ramamohanarao, K., & Palaniswami, M., (2013). A soft modularity function for detecting fuzzy communities in social networks. *IEEE Transaction on Fuzzy Systems*.
- [7] Lusseau, D., Schneider, K., Boisseau, O. J., Haase, P., Slooten, E., & Dawson, S. M., (2003). The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, vol. 54, pp. 396-405.
- [8] Girvan, M., & Newman, M. E. J., (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826.
- [9] Yazdanparast, A., Havens, T., & Jamalabdollahi, M. (2020). Soft overlapping community detection in large-scale networks via fast fuzzy modularity maximization. *IEEE Transactions of Fuzzy Systems*.
- [10] Mladenovic, N. (1995). A variable neighborhood algorithm - a new metaheuristic for combinatorial optimization. *Optimization Days*, Vol. 112.



## METODA PROMENLJIVIH OKOLINA ZA PROBLEM P-REZERVNOG CENTRA A VARIABLE NEIGHBOURHOOD SEARCH METHOD FOR THE P-NEXT CENTER PROBLEM

JELENA TASIĆ<sup>1</sup>, ZORICA STANIMIROVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Matematički fakultet, {jelena\_tasic, zoricast}@matf.bg.ac.rs

**Rezime:** U ovom radu razmatrana je jedna modifikacija problema  $p$ -centra, pod nazivom problem  $p$ -rezervnog centra (engl.  $p$ -next center), kod koga se pretpostavlja da se centri mogu iznenada zatvoriti usled nekog incidenta ili nedostatka osoblja. U slučaju iznenadnog zatvaranja centra, potrošači se preusmeravaju na rezervni centar, odnosno, otvoreni centar najbliži tom zatvorenom centru. Kako je razmatrani problem NP-težak, korišćen je metaheuristički pristup za njegovo rešavanje zasnovan na metodi promenljivih okolina. Testiranje je izvršeno na skupu  $p$ med instanci do 200 korisnika i 50 centara. Dobijeni rezultati su upoređeni sa poznatim rešenjima iz literature i pokazuju efikasnost i pouzdanost predložene metode pri rešavanju problema  $p$ -rezervnog centra.

**Ključne reči:** Metoda promenljivih okolina,  $p$ -rezervni centar, metaheuristika, kombinatorna optimizacija

**Abstract:** This paper considers a modification of the  $p$ -center problem, denoted as the  $p$ -next center, in which it is assumed that a center may suddenly close due to some kind of incident or lack of staff. If that turns out to be the case, customers are being directed to the backup center - the one closest to the closed center. As the considered problem is NP-hard, a metaheuristic approach based on the variable neighborhood search method is proposed. Computational study is performed over the set of  $p$ med test instances with up to 200 users and 50 centers. The obtained experimental results are compared with the optimal or best-known solutions from the literature, showing the efficiency and stability of the proposed method when solving the  $p$ -next center problem.

**Keywords:** Variable neighborhood search method,  $p$ -next center, metaheuristic, combinatorial optimization

### 1. UVOD I MATEMATČKA FORMULACIJA

Problem  $p$ -centra je jedan od najpoznatijih i najviše proučavanih problema kombinatorne optimizacije. Uveo ga je Hakimi 1965. godine u [3], a u radu [6] je dokazano da je je NP-težak. Neophodno je odabrati lokacije za uspostavljanje  $p$  centara i alocirati svakog korisnika nekom od njih, tako da se minimizuje maksimalno rastojanje od korisnika do njemu pridruženog centra. Problem  $p$ -centra i njegove modifikacije su do sada proučavani u brojnim radovima i razvijen je širok spektar metoda za njegovo rešavanje [2].

U ovom radu proučavan je problem  $p$ -rezervnog centra (engl.  $p$ -next center problem, PNC), koji predstavlja jednu modifikaciju problema  $p$ -centra. Problem PNC uzima u obzir da neki od uspostavljenih centara može postati neupotrebljiv usled tehničkih problema, prirodnih katastrofa, nedostatka radnika itd. Osim lokacija centara, neophodno je odrediti i lokacije rezervnih (backup) centara, na koje se korisnici preusmeravaju ukoliko primarni centar više nije u funkciji. Cilj je minimizovati maksimalno rastojanje lokacije korisnika do njemu najbližeg (primarnog) centra uvećano za rastojanje primarnog centra do njemu najbližeg (rezervnog) centra. Ovo se može i drugačije formulisati, pod pretpostavkom da nije unapred poznato da li je neki primarni centar aktivan ili ne. Ako se pri poseti ispostavi da primarni centar nije u funkciji, korisnik produžava do rezervnog centra, tako da treba minimizovati maksimalno rastojanje lokacije do njoj najbližeg rezervnog centra.

Problem PNC je prvi put izložen u radu [1], a motivacija za njegovo uvođenje je realna situacija sa problemom u snabdevanju strujom usled zemljotresa u Španiji. Autori su predložili nekoliko matematičkih formulacija problema PNC i uporedili njihovu efikasnost. Do sada, problem je rešavan GRASP i VNS metodama, kao i njihovom hibridizacijom koje su predložene u [8]. U radu [10] implementirana je Filtrirana VNS metoda, dok je u radu [7] korišćen evolutivni pristup za rešavanje problema PNC.

Ostatak rada je organizovan na sledeći način. U sekciji 2 data je matematička formulacija problema  $p$ -rezervnog centra. U sekciji 3 opisani su detalji predložene metode promenljivih okolina za njegovo rešavanje. Rezultati testiranja izloženi su u sekciji 4, dok su zaključak i smernice daljeg istraživanja dati u sekciji 5.

## 2. MATEMATIČKA FORMULACIJA PROBLEMA PNC

Neka je  $A$  skup korisnika i istovremeno skup potencijalnih lokacija za uspostavljanje centara,  $p$  je broj centara koje treba uspostaviti i neka su rastojanja između lokacija zadata matricom  $D = (d_{ij})_{n \times n}$ . Celobrojna linearna matematička formulacija problema PNC predložena u radu [1] koristi dva skupa binarnih promenljivih:

$$y_j = \begin{cases} 1, & \text{ukoliko je uspostavljen centar na lokaciji } j, \\ 0, & \text{inače.} \end{cases}$$

$$x_{ij} = \begin{cases} 1, & \text{ukoliko je } j \text{ najbliži centar za lokaciju } i \text{ (različit od } i), \\ 0, & \text{inače.} \end{cases}$$

Treba primetiti da  $x_{ij} = 1$  ima različito značenje u zavisnosti da li je na lokaciji  $i$  uspostavljen centar ili ne. Ukoliko jeste uspostavljen centar na lokaciji  $i$ , onda je  $j$  zapravo rezervni centar za centar  $i$ . Ukoliko  $i$  nije centar,  $x_{ij} = 1$  znači da je lokacija  $i$  pridružena primarnom centru  $j$ . Matematički se problem PNC može formulirati na sledeći način [1]:

$$\min w \tag{1}$$

$$\sum_{j \in A} y_j = p, \tag{2}$$

$$\sum_{j \in A, j \neq i} x_{ij} = 1, \quad i \in A, \tag{3}$$

$$x_{ij} \leq y_j \quad i, j \in A, i \neq j, \tag{4}$$

$$y_j + \sum_{k \in A, d_{ik} > d_{ij}} x_{ik} \leq 1 \quad i, j \in A, i \neq j, \tag{5}$$

$$w \geq \sum_{k \in A, k \neq j} d_{jk} x_{jk}, \quad j \in A, \tag{6}$$

$$w \geq d_{ij}(x_{ij} - y_i) + \sum_{k \in A, k \neq j} d_{jk} x_{jk}, \quad i, j \in A, i \neq j, \tag{7}$$

$$x_{ij} \in \{0, 1\}, \quad i, j \in A, i \neq j, \tag{8}$$

$$y_j \in \{0, 1\}, \quad j \in A. \tag{9}$$

Pomoćna kontinualna promenljiva  $w$  predstavlja maksimalno rastojanje lokacije do njoj najbližeg rezervnog centra. Uslov (2) garantuje da je uspostavljeno tačno  $p$  centara. Uslovima (3) obezbeđuje se da je svaki korisnik pridružen tačno jednom centru, odnosno da za svaki centar postoji tačno jedan rezervni centar. Uslovima (4) onemogućuje se pridruživanje korisnika centru koji nije uspostavljen. Uslovima (5) garantuje se da je svaka lokacija pridružena njoj najbližem otvorenom centru. Uslovi (6) postavljaju  $w$  za gornju granicu rastojanja lokacije do njoj najbližeg rezervnog centra. Uslovi (7) obezbeđuju da se sve lokacije sa istim primarnim centrom preusmeravaju ka istom rezervnom centru i da je ukupno rastojanje u tom slučaju jednako zbiru rastojanja od lokacije do primarnog centra i rastojanju od primarnog do rezervnog centra.

## 3. METODA PROMENLJIVIH OKOLINA ZA REŠAVANJE PROBLEMA PNC

Metoda promenljivih okolina (engl. Variable Neighborhood Search-VNS) je metaheuristička metoda koju su predložili Hansen i Mladenović u radu [4]. Glavna ideja metode je promena okoline u kojoj se vrši lokalna pretraga, a osim ovoga najčešće korišćenu verziju metode-osnovnu metodu promenljivih okolina (Basic VNS) karakteriše i stohastički element, tzv. faza razmrdavanja. Osnovne faze BVNS metode su: razmrdavanje, lokalna pretraga tekuće okoline i promena okoline (pomeraj). U fazi inicijalizacije potrebno je generisati početno rešenje i definisati okoline koje će biti korišćene za lokalnu pretragu, kao i okoline koje će se koristiti u fazi razmrdavanja. Zatim se iterativno smenjuju faze razmrdavanja, poboljšanja i pomeraja do zadovoljenja odabranog kriterijuma zaustavljanja (maksimalan broj iteracija, maksimalno vreme izvršavanja, maksimalan broj iteracija bez popravke rešenja itd).

U ovom radu, za rešavanje problema PNC implementirana je osnovna varijanta metode promenljivih okolina (engl. Variable Neighborhood Search-BVNS). Slična varijanta VNS metode je upotrebljena prilikom rešavanja problema p-medijane u radu [5], kao i prilikom rešavanja standardnog problema p-centra u radu [9].

Predložena BVNS metoda koristi celobrojnu reprezentaciju rešenja. Kako je  $n$  ukupan broj lokacija i  $p$  broj (primarnih i rezervnih) centara koje treba odrediti, rešenje problema PNC dato je vektorom  $x$  dužine  $p$



koji sadrži lokacije primarnih centara. Kada se iz datog vektora pročitaju lokacije na kojima su uspostavljeni primarni centri, svaki korisnik se pridruži svom najbližem uspostavljenom centru, a za svaki primarni centar određuje se lokacija rezervnog centra (uspostavljeni centar koji mu je najbliži).

Početno rešenje  $x^*$  se generiše na slučajnan način (celobrojni vektor dužine  $p$ ). Okoline  $N_k$ ,  $k \in N$  korišćene u BVNS metodi su definisane na sledeći način:  $N_k(x) = \{x' \in S : d(x, x') = k\}$ , pri čemu je  $S$  skup svih dopustivih rešenja, a  $d(x, x')$  rastojanje između  $x$  i  $x'$  koje predstavlja broj uspostavljenih lokacija u kojima se ova dva rešenja razlikuju. U fazi razmrđavanja korišćene su okoline  $N_k$ ,  $k = 1, 2, \dots, k_{max}$ , dok je u fazi lokalne pretrage korišćena okolina  $N_1$ . Kardinalnost ove okoline je  $p(n-p)$  i traženo je rešenje  $x''$  koje daje najbolju popravku. Za svaku lokaciju koja nije deo trenutnog rešenja  $x'$ , određuje se centar trenutnog rešenja koji ona treba da zameni, tako da vrednost funkcije cilja bude najmanja moguća. Tom prilikom se pamte i ažuriraju rastojanja od korisnika do njihovih rezervnih centara, a najveće od tih rastojanja predstavlja vrednost funkcije cilja za rešenje koje bi se dobilo posmatranom zamenom. Kada se odredi najbolji par, vrši se zamena, čime je pretraga okoline  $N_1$  završena. Za kriterijum zaustavljanja uzet je maksimalan broj iteracija  $iter_{max}$ , odnosno broj ponavljanja VNS ciklusa koji se sastoji od faze razmrđavanja, poboljšanja i pomeraja. Struktura implementirane BVNS metoda za problem PNC je prikazana Algoritmom 1.

---

#### Algoritam 1 BVNS algoritam za PNC

---

- 1: Inicijalizacija: odabrati skup okolina  $N_k$ ,  $k = 1, 2, \dots, k_{max}$  i parametar  $iter_{max}$  za kriterijum zaustavljanja.
  - 2: Konstruisati početno rešenje  $x^*$ .
  - 3: **while** kriterijum zaustavljanja nije ispunjen **do**
  - 4:     Postaviti  $k = 1$
  - 5:     **while**  $k \leq k_{max}$  **do**
  - 6:         Odrediti  $x'$  slučajnim izborom tačke iz okoline  $N_k(x^*)$  (**Razmrđavanje**)
  - 7:         Odrediti  $x''$  lokalnom pretragom počevši od tačke  $x'$  (**Lokalna pretraga**)
  - 8:         Ukoliko je novo rešenje  $x''$  bolje u smislu funkcije cilja od rešenja  $x^*$ , ažurirati  $x^*$  i  $k: x^* = x''$  i  $k = 1$ ,
  - 9:         inače  $k = k + 1$  (**Prihvatanje rešenja**)
- 

## 4. EKSPERIMENTALNI REZULTATI

U eksperimentalnoj studiji korišćene su instance pmed1-pmed4 i pmed6-pmed8, iz kojih je preuzeto prvih  $n$  čvorova. Instance su preuzete sa <http://people.brunel.ac.uk/~mastjjb/jeb/info.html>, a podeljene su u tri grupe prema broju čvorova: male instance (do 50 čvorova), srednje (do 120 čvorova) i velike (do 200 čvorova). Za ulazne podatke iz svake od pmed instanci, neophodno je najpre primeniti Flojd-Varšalov algoritam da bi se dobila matrica rastojanja i zatim pristupiti rešavanju problema.

Testiranja su rađena na računaru sa AMD Ryzen 5 3500U procesorom i 8GB RAM memorije. U cilju dobijanja optimalnih rešenja (ako je moguće), korišćen je egzaktan rešavač CPLEX 20.1. Vreme rada rešavača ograničeno na 2 sata za sve test instance. Predložena BVNS metaheuristika implementirana je u programskom jeziku C++ i koristi parametre date u Tabeli 1.

**Tabela 1:** Vrednosti parametara za instance malih, srednjih i velikih dimenzija

Instance	$iter_{max}$	$k_{max}$	Broj pokretanja
male	30	$\lfloor p/2 \rfloor$	20
srednje	100	10	10
velike	120	10	5

Tabele 2-5 sadrže eksperimentalne rezultate na skupu malih, srednjih i velikih instanci, respektivno. Za svaku instancu, pored imena, date su njene dimenzije ( $n$  i  $p$ ). Zatim je prikazano optimalno rešenje  $opt$  egzaktnog rešavača (ako je poznato), vreme za koje je rešenje nađeno ( $t$  u sekundama) i broj iteracija  $iter$ . U slučaju kada nije nađeno optimalno rešenje posle 2 sata, u tabeli je data dobijena gornja granica označena znakom "\*". Naredne kolone se odnose na rezultate BVNS metaheuristike i sadrže sledeće vrednosti:

- $best$  – najbolje dobijeno rešenje u 20 izvršavanja,
- $t_{best}$  – prosečno vreme za koje BVNS prvi put dostigne najbolje/optimalno rešenje (u sekundama),
- $t_{tot}$  – srednje ukupno vreme izvršavanja (u sekundama),
- $agap$  – srednje odstupanje od najboljeg/optimalnog rešenja (u procentima),
- $\sigma$  – standardnu devijaciju tog odstupanja.

**Tabela 2:** Rezultati za  $p$ -med instance dimenzija  $10 \leq n \leq 50$ ,  $5 \leq p \leq 20$ 

Instanca	n	p	CPLEX			BVNS				
			best	t(s)	Iter	best	$t_{best}(s)$	$t_{tot}(s)$	agap(%)	$\sigma$
pmed1	10	5	84	0.06	44	<b>84</b>	0.00	0.02	0.00	0.00
pmed1	20	5	120	1.17	921	<b>120</b>	0.01	0.10	0.00	0.00
pmed1	20	10	95	0.28	129	<b>95</b>	0.00	0.16	0.00	0.00
pmed1	30	5	126	6.11	2998	<b>126</b>	0.02	0.21	0.00	0.00
pmed1	30	10	95	1.5	750	<b>95</b>	0.05	0.49	0.00	0.00
pmed1	40	5	144	35.48	33188	<b>144</b>	0.06	0.40	3.09	0.94
pmed1	40	20	89	3.14	316	<b>89</b>	0.01	0.76	0.00	0.00
pmed1	50	10	110	24.45	9481	<b>110</b>	0.17	0.77	1.18	1.15
pmed1	50	20	89	7.67	490	<b>89</b>	0.40	1.80	1.12	1.12
pmed2	10	5	121	0.09	100	<b>121</b>	0.01	0.03	0.29	1.26
pmed2	20	10	99	0.70	243	<b>99</b>	0.02	0.14	0.00	0.00
pmed2	20	5	147	1.83	1050	<b>147</b>	0.00	0.06	0.00	0.00
pmed2	30	5	169	8.77	4782	<b>169</b>	0.02	0.14	0.06	0.26
pmed2	30	10	110	2.39	785	<b>110</b>	0.03	0.33	0.00	0.00
pmed2	40	5	164	37.92	38064	<b>164</b>	0.04	0.20	0.00	0.00
pmed2	40	20	96	2.92	462	<b>96</b>	0.31	1.26	0.31	0.74
pmed2	50	10	140	97.22	266697	<b>140</b>	0.22	1.51	0.00	0.00
pmed2	50	20	99	16.52	37429	<b>99</b>	0.56	2.67	0.05	0.22
pmed3	10	5	77	0.05	19	<b>77</b>	0.00	0.02	0.00	0.00
pmed3	20	5	145	1.09	961	<b>145</b>	0.01	0.09	0.00	0.00
pmed3	20	10	77	0.28	84	<b>77</b>	0.01	0.18	0.00	0.00
pmed3	30	5	157	6.80	3184	<b>157</b>	0.06	0.24	0.32	1.39
pmed3	30	10	122	3.09	1444	<b>122</b>	0.03	0.48	0.00	0.00
pmed3	40	5	157	49.84	37434	<b>157</b>	0.10	0.41	0.67	1.74
pmed3	40	20	77	5.53	636	<b>77</b>	0.13	1.68	0.00	0.00
pmed3	50	10	125	97.94	273537	<b>125</b>	0.67	1.62	1.08	1.61
pmed3	50	20	87	7.75	891	<b>87</b>	0.43	2.00	0.11	0.50

**Tabela 3:** Rezultati za  $p$ -med instance dimenzija  $75 \leq n \leq 120$ ,  $20 \leq p \leq 50$ 

Instanca	n	p	CPLEX			BVNS				
			best	t(s)	Iter	best	$t_{best}(s)$	$t_{tot}(s)$	agap(%)	$\sigma$
pmed6	75	20	75	256.16	269139	<b>75</b>	2.60	6.06	1.2	0.93
pmed6	75	30	69	59.30	1341	<b>69</b>	1.98	3.88	0.87	0.71
pmed6	100	20	73	2767.52	903116	75	3.29	10.21	2.67	2.46
pmed6	100	30	67	181.92	7010	<b>67</b>	1.59	8.08	3.13	2.05
pmed6	100	50	67	119.75	756	<b>67</b>	0.08	2.44	0.00	0.00
pmed6	120	20	78*	39872	8990418	<b>77</b>	4.05	9.98	5.71	5.55
pmed6	120	50	67	109.25	1180	<b>67</b>	1.00	5.33	0.00	0.00
pmed7	75	20	69	51	2042	<b>69</b>	2.80	6.25	1.16	2.32
pmed7	75	30	69	35.62	408	<b>69</b>	0.28	4.53	0.00	0.00
pmed7	100	20	66	1231.31	1176375	<b>66</b>	6.05	11.71	0.30	0.61
pmed7	100	30	59	106.48	20268	<b>59</b>	2.81	8.59	3.90	3.39
pmed7	100	50	59	64.23	1110	<b>59</b>	0.19	2.53	0.00	0.00
pmed7	120	20	66	4124.23	1629296	67	2.86	10.31	1.49	0.94
pmed7	120	50	59	108.09	2268	<b>59</b>	1.29	5.17	1.19	3.56
pmed8	75	20	78	603.80	241943	<b>78</b>	0.64	6.50	3.21	1.31
pmed8	75	30	63	50	11206	65	1.57	4.62	5.38	5.97
pmed8	100	20	75	958.14	745627	76	3.49	13.06	2.76	1.49
pmed8	100	30	61	312.05	268277	63	3.70	9.69	7.78	8.68
pmed8	100	50	58	63.36	710	<b>58</b>	0.58	3.78	1.55	2.37
pmed8	120	20	78*	36515.2	8457198	<b>78</b>	2.39	10.96	0.64	0.64
pmed8	120	50	58	137.91	2694	<b>58</b>	1.55	5.14	0.00	0.00

**Tabela 4:** Rezultati za  $p$ -med instance dimenzija  $150 \leq n \leq 200$ ,  $20 \leq p \leq 50$ 

Instanca	n	p	CPLEX			BVNS				
			best	t(s)	Iter	best	$t_{best}(s)$	$t_{tot}(s)$	agap(%)	$\sigma$
pmed6	150	20	83*	27309.9	2038753	<b>77</b>	28.52	90.21	1.56	1.51
pmed6	150	30	68*	37360.1	5793732	70	37.34	90.37	2.29	2.32
pmed6	150	50	56	277.73	5384	<b>56</b>	9.45	34.32	1.79	3.57
pmed6	150	80	56	309.42	3367	<b>56</b>	0.10	14.83	0.00	0.00
pmed6	200	20	116*	14782.2	269124	<b>77</b>	91.15	171.29	2.08	2.11
pmed6	200	30	76*	17125.8	423657	<b>70</b>	38.91	88.80	4.29	4.69
pmed6	200	50	49	15945.5	1314118	55	45.60	97.34	13.09	10.56
pmed6	200	80	49	1115.72	9858	<b>49</b>	13.7	38.51	0.82	0.99
pmed7	150	20	67*	32063.3	5690530	68	18.33	63.11	0.29	0.59
pmed7	150	30	59	1199.48	215948	<b>59</b>	13.18	46.03	8.47	4.67
pmed7	150	50	59	282.73	7398	<b>59</b>	3.02	43.84	0.00	0.00
pmed7	200	20	75*	24872.8	542016	<b>68</b>	39.82	139.13	1.18	1.10
pmed7	200	30	64*	22770.4	697833	67	27.93	108.71	1.49	0.94
pmed7	200	50	50	8113.89	2081827	55	29.67	65.03	10.18	8.18
pmed7	200	80	46	1283.41	13176	49	20.93	45.93	5.312	4.40
pmed8	150	20	76*	41399.5	5493184	<b>74</b>	33.07	117.70	1.35	2.70
pmed8	150	30	58	1066.39	47329	<b>58</b>	10.97	59.84	7.24	4.28
pmed8	150	50	58	294	8538	<b>58</b>	3.14	31.83	0.00	0.00
pmed8	200	20	93*	14789.7	339573	<b>84</b>	14.54	150.99	0.71	1.43
pmed8	200	30	75*	14715.1	321585	<b>70</b>	69.24	117.38	3.43	2.32
pmed8	200	50	68	1121.56	13138	<b>68</b>	7.95	66.25	0.00	0.00

Za sve instance malih dimenzija predložena metaheuristika dostiže optimalna rešenja, a prosečno vreme izvršavanja je kraće nego kod CPLEX-a. Parametar  $agap$  ne prelazi vrednost 3.09%, dok  $\sigma$  ne prelazi vrednost 1.74% (Tabela 2). Za instance srednjih dimenzija, CPLEX nije našao optimalno rešenje u 2 slučaja, dok BVNS nalazi isto ili bolje rešenje od CPLEX-a. Na ostalim instancama srednjih dimenzija, BVNS nije dostigao optimalno rešenje 5 puta. U proseku, parametar  $agap$  ne prelazi vrednost 7.78%, a parametar  $\sigma$  ne prelazi 8.68% (Tabela 3). Za instance velikih dimenzija CPLEX nije našao optimalno rešenje u 10 slučajeva. Za 7 od tih 10 instanci BVNS nalazi bolje rešenje od CPLEX-a, a u 3 slučaja BVNS ne dostiže optimalna rešenja. Parametar  $agap$  ne prelazi 13.09%, dok parametar  $\sigma$  ne prelazi 10.56% (Tabela 4). Za sve testirane instance, vreme izvršavanja BVNS-a je bilo kraće u odnosu na vreme CPLEX rešavača.

Na osnovu dobijenih rezultata može se primetiti uticaj odnosa parametara  $n$  i  $p$ . Čak i kod instanci malih dimenzija, smanjivanje odnosa ovih parametara dovodi do dužeg vremena izvršavanja i velikog broja iteracija. Rešavačem u dva slučaja iz grupe srednjih instanci nije nađeno optimalno rešenje u zadatom vremenu, dok kod instanci velikih dimenzija optimalno rešenje nije nađeno u 10 slučajeva. Nasuprot tome, za vrednosti parametra  $p$  koje su bliske  $n/2$  vreme izvršavanje i broj iteracija se smanjuju, čak i za instance velikih dimenzija.

Rezultati BVNS metode na instancama velikih dimenzija su upoređeni sa rezultatima Filtered VNS metode iz [10], hibridnog GRASP-VND pristupa iz [8] i evolutivnog pristupa iz [7]. Iz Tabele 5 se može videti da je BVNS metaheuristika našla bolja rešenja od Filtered VNS metode u 14 od 21 slučaja, u ostalim slučajevima nađeno je isto rešenje. Što se se hibridnog algoritma iz rada [8] tiče, BVNS je našla bolja rešenja 15 od 21 puta, a u preostalim slučajevima nađeno je isto rešenje kao i hibridnim algoritmom. Neke od instanci rešavane su i u radu [7]. BVNS u 9 od 18 slučajeva nalazi isto rešenje kao i evolutivni algoritam iz [7], ali u značajno kraćem vremenu. U ostalim slučajevima EA nalazi bolje rešenja. Treba primetiti da su vremena izvršavanja EA data u Tabeli 5 prikazana u minutima (m), satima (h) i danima (d).

**Tabela 5:** Poređenje rezultata BVNS, Filtered VNS, GRASP-VNS i EA na instancama većih dimenzija

Instanca	n	p	BVNS		FVNS		GRASP-VNS		EA	
			best	$t_{tot}(s)$	best	t(s)	best	t(s)	best	t
pmed6	150	20	<b>77</b>	90.21	80	222.75	79	33.86	77	7.78 d
pmed6	150	30	70	90.37	74	250.42	71	77.19	<b>64</b>	5.17 h
pmed6	150	50	<b>56</b>	34.32	60	240.37	62	200.17	56	2.42 m
pmed6	150	80	<b>56</b>	14.83	56	68.05	56	479.01	56	1.30 m
pmed6	200	20	<b>77</b>	171.29	81	224.1	79	49.78	-	-
pmed6	200	30	70	88.80	77	254	72	150.29	<b>69</b>	6.05 d
pmed6	200	50	55	97.34	70	139.42	68	493.55	<b>49</b>	48.73 m
pmed6	200	80	<b>49</b>	38.51	54	224.1	54	1151.12	49	5.99 m
pmed7	150	20	68	63.11	68	226.25	69	22.90	<b>67</b>	10.13 h
pmed7	150	30	<b>59</b>	46.03	63	254.15	62	66.10	59	3.92 m
pmed7	150	50	<b>59</b>	43.84	59	209.71	59	206.33	59	1.69 m
pmed7	200	20	<b>68</b>	139.13	75	195.7	73	44.05	-	-
pmed7	200	30	67	108.71	67	222.79	68	129.25	<b>60</b>	22.00 h
pmed7	200	50	55	65.03	63	211.34	63	505.03	<b>50</b>	55.14 m
pmed7	200	80	49	45.93	52	210.42	52	1213.66	<b>46</b>	5.70 m
pmed8	150	20	74	117.70	74	244.01	74	23.76	<b>70</b>	11.21 h
pmed8	150	30	<b>58</b>	59.84	63	218.5	61	64.84	58	4.38 m
pmed8	150	50	<b>58</b>	31.83	58	263.65	58	194.59	58	1.61 m
pmed8	200	20	84	150.99	88	222.01	<b>84</b>	41.21	-	-
pmed8	200	30	70	117.38	84	226.27	77	122.14	<b>69</b>	1.39 h
pmed8	200	50	<b>68</b>	66.25	68	247.91	68	394.95	68	5.82 m

#### 4.1. Zaključak

U ovom radu je razmatran problem  $p$ -rezervnog centra (PNC) koji uzima u obzir mogućnost otkaza nekog od uspostavljenih centara i u tom slučaju pridružuje korisnike rezervnom centru. Implementirana je osnovna varijanta metode promenljivih okolina (BVNS) čiji su elementi prilagođeni karakteristikama problema. Testiranje je izvršeno na pmed instancama iz literature, a dobijeni rezultati upoređeni su sa rezultatima egzaktnog rešavača CPLEX i postojećim metaheuristikama iz literature. Na osnovu prikazanih rezultata se može zaključiti da je predložena BVNS metoda pogodna za rešavanje problema PNC. Pravci daljeg istraživanja su testiranje na pmed instancama većih dimenzija kao i implementacija drugih varijanti VNS metode za rešavanje PNC problema.

#### LITERATURA

- [1] Albareda-Sambola, M., Hinojosa, Y., Marín, A., and Puerto, J. (2015). When centers can fail: A close second opportunity. *Computers & Operations Research*, 62:145–156.
- [2] Daskin, M. S. (2011). *Network and discrete location: models, algorithms, and applications*. John Wiley&Sons.
- [3] Hakimi, S. (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problem. *Operations Research*, 13(3):462–475.
- [4] Hansen, P. and Mladenović, N. (1997a). Variable neighborhood search. *Computers & Operations Research*.
- [5] Hansen, P. and Mladenović, N. (1997b). Variable neighborhood search for the  $p$ -median. *Location Science*, 5(4):207–226.
- [6] Kariv, O. and Hakimi, S. (1979). An algorithmic approach to network location problems. part 1: The  $p$ -centers. *SIAM Journal on Applied Mathematics*, 37(3):513–538.
- [7] Londe, M. A., Andrade, C. E., and Pessoa, L. S. (2021). An evolutionary approach for the  $p$ -next center problem. *Expert Systems with Applications*, 175:114728.
- [8] Lopez-Sanchez, A., Sanchez-Oro, J., and Hernandez-Dóaz, A. (2018). GRASP and VNS for solving the  $p$ -next center problem. *Computers & Operations Research*.
- [9] Mladenović, N., Labbe, M., and Hansen, P. (2003). Solving the  $p$ -center problem with tabu search and variable neighborhood search. *Networks An International Journal*, 42(1):48–64.
- [10] Ristić, D., Mladenović, N., Todosijević, R., and Urosević, D. (2021). Filtered variable neighborhood search method for the  $p$ -next center problem. *International Journal for Traffic and Transport Engineering*.

## PRIMENA KARUSEL HEURISTIKE NA PROBLEM RUTIRANJA VOZILA

### APPLICATION OF CAROUSEL GREEDY HEURISTIC TO VEHICLE ROUTING PROBLEM

JANA VUGDELIJA<sup>1</sup>, DRAGANA MAKAJIĆ-NIKOLIĆ<sup>2</sup>

<sup>1</sup> University of Belgrade – Faculty of Organizational Sciences vugdelijajana@gmail.com

<sup>2</sup> University of Belgrade – Faculty of Organizational Sciences dragana.makajic-nikolic@fon.bg.ac.rs

**Rezime:** Predmet ovog rada je rešavanje problema rutiranja sa homogenom flotom vozila ograničenog kapaciteta pomoću Karusel heuristike. Karusel heuristika je unapređenje pohlepnog (greedy) algoritma, koje na lak i jednostavan način omogućava poboljšanje kvaliteta početnog rešenja dobijenog nekim pohlepnim algoritmom. Rezultati primene će biti ilustrovani na grupi test primera različitih dimenzija.

**Cljučne reči:** rutiranje vozila, homogena flota, Karusel, heuristika.

**Abstract:** The subject of this paper is solving the capacitated vehicle routing problem with homogeneous fleet using Carousel greedy heuristics. Carousel greedy is the enhanced greedy algorithm, which in an easy and simple way enables the improvement of the quality of the initial solution obtained by some Greedy procedure. The results of the implementation will be illustrated on a group of benchmarks of different dimensions.

**Keywords:** vehicle routing, homogeneous fleet, Carousel greedy, heuristics

## 1. UVOD

Problem rutiranja vozila sa ograničenim kapacitetom (*Capacitated Vehicle Routing Problem - CVRP*) je dobro poznat problem koji su postavili Dancig i Ramser 1959. godine [10]. Predmet ovog rada je CVRP sa homogenim voznim parkom, koji je bio rešavan različitim pristupima. U bazi Scopus se u periodu od 2002. do 2022. godine može naći 85 radova kojima se u naslovu, apstraktu i ključnim rečima nalaze reči *homogeneous* i *capacitated vehicle routing*. Pošto je u pitanju NP-težak problem, uglavnom je rešavan heurističkim algoritmima. Neki od najcitiranijih radova koriste pristupe: hibridnog algoritma koji kombinuje Iterativno lokalno pretraživanje i VNS [20], Evolutivni algoritmi [17], GRASP (*Greedy Randomized Adaptive Search Procedure*) [11], VND (*Variable Neighborhood Descent*) [9] itd. Poslednjih godina se primenjuju algoritmi poput: Granularnog tabu pretraživanja [12], specijalnih heuristika [13, 19], ASA (*Augmented Savings Algorithm*) [18] itd.

U ovom radu će se za rešavanje CVRP sa homogenim voznim parkom koristiti Karusel heuristika (*Carousel greedy-CG*), koja je prvi put objavljena u [5]. CG je konstruktivna dvoparameterska heuristika u kojoj se početno rešenje, dobijeno nekim pohlepnim algoritmom, degradira (postaje nedopustivo) i zatim obnavlja primenom istog pohlepnog algoritma. Parametri CG su  $\beta$ , kojim se određuje procenat rešenja koji će biti degradiran, odnosno broj promenljivih koje će biti uklonjene iz rešenja, i  $\alpha$ , kojim se određuje broj iteracija koji obezbeđuje da svaka od promenljivih iz početnog rešenja bude uklonjena i zamenjena nekom drugom promenljivom bar jedamput.

Najveći broj primena Karusel heuristike obavili su i objavili njeni autori Cerone, Ceruli i Golden, zajedno sa drugim autorima. Problemi u čijem rešavanju su koristili ovu heuristiku su: minimalno pokrivanje čvorova grafa [5], maksimizacija životnog veka mreže senzora (*Maximum Network Lifetime Problem - MLP*) [4, 8], distribucija u lancu snabdevanja sa više fabrika, više distributera i više kupaca [7], razapinjuće stablo [6], problem trgovačkog putnika [3]. Pored njih, Kong i drugi su primenili Karusel heuristiku u rešavanju problema detekcije zajednica (*Community Detection*) [15] a korišćena je i u rešavanju problema pogađanja skupova [16]. Najnoviji rad u kome je primenjena Karusel heuristika je odnosi na rešavanje problema ranca sa penalima koji se plaćaju ukoliko se neki od predefinisanih parova predmeta nađe u rešenju (*Knapsack Problem with Forfeits - KPF*) [2].

Rad se sastoji iz četiri sekcije. Nakon uvodnog dela, u Sekciji 2 je opisan algoritam Karusel heuristike za rešavanje CVRP i Klark-Rajt algoritam koji je korišćen za dobijanje početnog rešenja i obnavljanje degradiranog rešenja. U Sekciji 3 su prikazani rezultati eksperimenata na test primerima, a u Sekciji 4 zaključna razmatranja.

## 2. KARUSEL HEURISTIKA

Algoritam predstavljen u ovom radu implementiran je u programskom jeziku Pajton (Python). Pohlepni Klark-Rajt (CW) algoritam ušteda generiše početno rešenje iz kojeg se zatim izbacivanjem određenog procenta primenjenih ušteda dobija parcijalno (nedopustivo) rešenje. Karusel algoritam ovako dobijeno parcijalno rešenje iterativno modifikuje izbacivanjem jedne i dodavanjem nove uštede u svakoj iteraciji. Na kraju se aktuelno parcijalno rešenje dopunjava iz skupa preostalih ušteda, čime se dobija konačno rešenje.

Ceo algoritam se može predstaviti na sledeći način:

### Procedura 1 - CW

Ulazni podaci: koordinate čvorova (depo i kupci), tražnja svakog kupca, kapacitet vozila, broj raspoloživih kamiona.

1. Kreirati matricu rastojanja  $D_{ij}$  - izračunati euklidsko rastojanje između svaka dva čvora  $d_{ij}$ .
2. Kreirati početne rute oblika depo-kupac-depo za svakog kupca (0-i-0).
3. Za svaka dva kupca izračunati uštedu koja bi se ostvarila spajanjem ta dva čvora u istu rutu:
 
$$s_{ij} = d_{i0} + d_{j0} - d_{ij}$$
4. Uštede sortirati u nerastući niz.
5. Redom za svaku uštedu proveriti da li se može primeniti:
  - Da li su čvorovi na koje se ušteda odnosi na prvom ili poslednjem mestu u svojim trenutnim rutama
  - Da li se pomenuti čvorovi ne nalaze u istoj ruti
  - Da li je zbir tražnje trenutnih ruta pomenutih čvorova manji ili jednak kapacitetu vozila
6. Ukoliko su sva tri uslova ispunjena primeniti uštedu i ažurirati skup ruta spajanjem pomenute dve rute.
7. Ukoliko je broj na ovaj način dobijenih ruta manji ili jednak raspoloživom broju kamiona:
  - Izračunati vrednost funkcije cilja (pređeni put) sabiranjem dužina grana na svim rutama.
8. U suprotnom, dobijeno rešenje je nedopustivo i funkcija cilja ima vrednost  $+\infty$ .

Izlazni podaci: skup dobijenih ruta, niz primenjenih ušteda, vrednost funkcije cilja.

### Algoritam

Ulazni podaci: koordinate čvorova (depo i kupci), tražnja svakog kupca, kapacitet vozila, broj raspoloživih kamiona,  $a$ ,  $b$ ,  $\max\_br\_pokusaja$ .

1. Primenom CW odrediti početno rešenje.
2. Svaku uštedu iz početnog rešenja zadržati u parcijalnom rešenju sa verovatnoćom  $(1-b)$ .
3. Ponoviti dok broj izmena nije jednak  $a$  ili broj pokušaja nije jednak  $\max\_br\_pokusaja$ :
  - 3.1. Na slučajan način izbaciti iz parcijalnog rešenja jednu od najstarijih 5 ušteda.
  - 3.2. Na slučajan način izabrati jednu od 5 najboljih ušteda van parcijalnog rešenja.
  - 3.3 Proveriti da li se ušteda može primeniti, kao u Procedura 1-CW korak 5.
  - 3.4. Ukoliko su ispunjeni uslovi primeniti izabranu uštedu.
  - 3.5. U suprotnom vratiti izbačenu uštedu.
4. Za svaku uštedu iz skupa preostalih ušteda, počev od najveće:
  - 4.1. Ukoliko su ispunjeni uslovi za dodavanje, dodati u parcijalno rešenje.
5. Za svaku rutu iz konačnog rešenja:
  - 5.1. Primeniti CW proceduru nad skupom čvorova rute kao ulaznim skupom.
  - 5.2. Ukoliko je dobijena kraća ruta od prvobitne, rutu iz rešenja zameniti novogenerisanom.
6. Ukoliko je broj na ovaj način dobijenih ruta manji ili jednak raspoloživom broju kamiona:
  - Izračunati vrednost funkcije cilja (pređeni put) sabiranjem dužina grana na svim rutama.
7. U suprotnom, dobijeno rešenje je nedopustivo i funkcija cilja ima vrednost  $+\infty$ .

Izlazni podaci: skup dobijenih ušteda, vrednost funkcije cilja, vizuelni prikaz rešenja.

Interesantno je napomenuti da se, s obzirom na princip dodavanja ušteda koji uključuje faktor slučajnosti, dešavalo da neke od ruta dobijenih u konačnom rešenju imaju tačke samopresecanja. Samim tim dobijeno rešenje neće biti optimalno i dužine takvih ruta bi se mogle smanjiti. Iz ovog razloga dodata je procedura koja za svaku rutu konačnog rešenja fiksira skup čvorova koji se u njoj nalaze, uštede između tih čvorova sortira u nerastući niz i pokušava da ih primeni redom, počevši od najveće. Proces se logički svodi na primenu Klark-Rajt algoritma samo na čvorove rute i u velikoj većini slučajeva pomaže da se na rutama izbegnu tačke samopresecanja.

### 3. REZULTATI EKSPERIMENATA

Za potrebe eksperimenta korišćeni su test primeri dati u [1], konkretno set A i set P koji sadrže ukupno 51 test primer. Kako je faktor slučajnosti prisutan pri kreiranju parcijalnog rešenja, kao i pri izboru uštede koja će iz parcijalnog rešenja biti izbačena i uštede koja će se zatim dodati, algoritam je pokrenut više puta. Rezultati prikazani tabeli 1 i razmatrani kroz dalju diskusiju predstavljaju najbolja rešenja dobijena kroz 20 pokušaja, pri čemu se korak generisanja početnog rešenja nije ponavljao, već su samo koraci 2-7 ponovljeni 20 puta. Ukupno vreme rada algoritma (51 put po 20 pokušaja) bilo je približno 190 sekundi. Eksperimenti su izvršavani na računaru sa 64-bitnim operativnim sistemom, procesorom AMD Ryzen 7 4700U, 2.00 GHz i 16 GB RAM memorije.

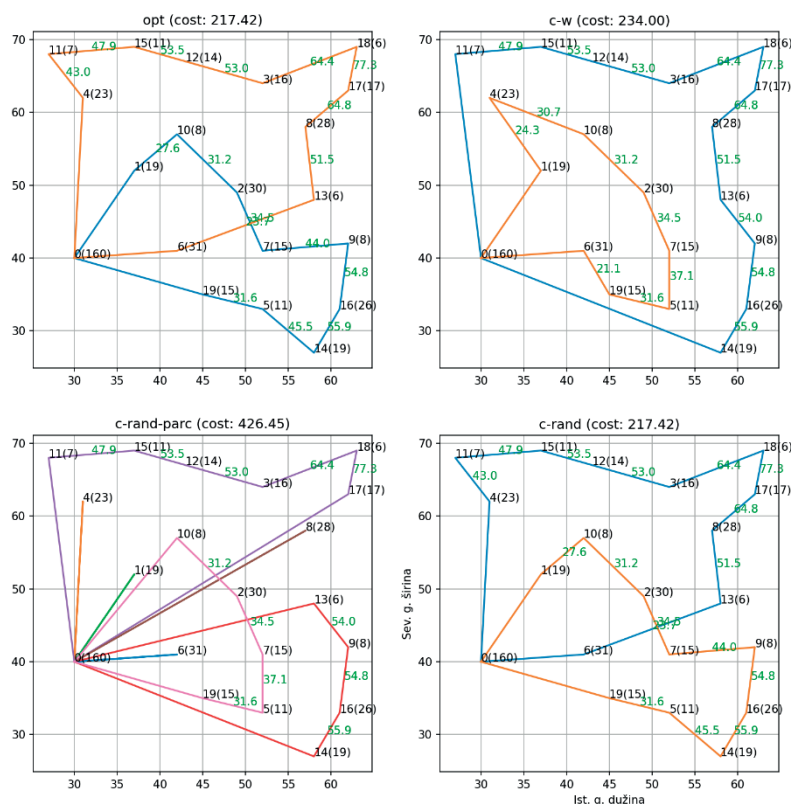
**Tabela 1:** Rezultati eksperimenata

Primer	Klark-Rajt rešenje	Karusel rešenje	Optimalno rešenje	Relativno odstupanje Karusel rešenja od opt. (%)
P-n101-k4	765.38	764.79	692.28	10.47
P-n16-k8	nedopustivo	451.34	451.34	0
P-n19-k2	237.89	227.3	212.66	6.89
P-n20-k2	234	217.42	217.42	0
P-n21-k2	236.19	220.5	212.71	3.66
P-n22-k2	239.5	238.68	217.85	9.56
P-n22-k8	nedopustivo	nedopustivo	601.42	/
P-n23-k8	nedopustivo	nedopustivo	531.17	/
P-n40-k5	518.37	501.16	461.73	8.54
P-n45-k5	572.95	552.7	512.79	7.78
P-n50-k10	nedopustivo	730.22	699.56	4.38
P-n50-k7	597.03	594.55	559.86	6.19
P-n50-k8	nedopustivo	nedopustivo	634.85	/
P-n51-k10	nedopustivo	nedopustivo	742.48	/
P-n55-k10	nedopustivo	727.77	697.81	4.29
P-n55-k15	nedopustivo	nedopustivo	991.48	/
P-n55-k7	618.68	613.21	570.27	7.53
P-n55-k8	631.67	634.03	592.17	7.07
P-n60-k10	800.19	800.19	748.94	6.84
P-n60-k15	nedopustivo	nedopustivo	971.58	/
P-n65-k10	851.67	851.67	796.67	6.9
P-n70-k10	nedopustivo	nedopustivo	830.02	/
P-n76-k4	688.34	672.98	598.22	12.5
P-n76-k5	709.38	709.26	635.04	11.69
A-n32-k5	843.69	843.69	787.81	7.09
A-n33-k5	712.05	692.28	662.76	4.45
A-n33-k6	nedopustivo	754.97	742.83	1.63
A-n34-k5	nedopustivo	790.63	781.3	1.19
A-n36-k5	828.47	806.78	802.13	0.58
A-n37-k5	707.81	707.81	672.59	5.24
A-n37-k6	976.61	976.61	952.22	2.56
A-n38-k5	nedopustivo	787.83	734.18	7.31
A-n39-k5	901.99	884.56	828.99	6.7
A-n39-k6	863.08	860.01	833.2	3.22
A-n44-k6	nedopustivo	998.2	939.33	6.27

Primer	Klark-Rajt rešenje	Karusel rešenje	Optimalno rešenje	Relativno odstupanje Karusel rešenja od opt. (%)
A-n45-k6	nedopustivo	nedopustivo	944.88	/
A-n45-k7	1199.98	1193.62	1147.22	4.04
A-n46-k7	939.74	939.74	918.46	2.32
A-n48-k7	1112.82	1112.82	1074.34	3.58
A-n53-k7	1099.45	1099.45	1013.31	8.5
A-n54-k7	1201.2	1201.2	1171.78	2.51
A-n55-k9	1099.84	1099.84	1074.46	2.36
A-n60-k9	1421.88	1420.51	1355.8	4.77
A-n61-k9	nedopustivo	nedopustivo	1039.08	/
A-n62-k8	1352.81	1333.94	1294.28	3.06
A-n63-k10	1352.48	1352.48	1313.73	2.95
A-n63-k9	nedopustivo	nedopustivo	1622.14	/
A-n64-k9	nedopustivo	1501.66	1400.83	7.2
A-n65-k9	nedopustivo	1276.54	1181.69	8.03
A-n69-k9	1210.78	1207.08	1165.99	3.52
A-n80-k10	1860.94	1860.94	1766.5	5.35

Kao što se iz predstavljenih rezultata može videti, rešenja dobijena korišćenjem Klark-Rajt algoritma nisu uvek bila dopustiva. U pojedinim slučajevima ova rešenja prekoračivala su ograničenje koje se odnosi na raspoloživ broj vozila, odnosno imala su prevelik broj ruta, ali su i dalje korišćena za generisanje parcijalnog rešenja. Od ukupno 51 primera, Klark-Rajt rešenja bila su nedopustiva u 19 slučajeva, što čini 37,25%. Karusel algoritam je kod 9 ovakvih slučajeva uspeo da dođe do dopustivog krajnjeg rešenja, dok za 10 test primera nije dobijeno dopustivo rešenje. Za svaki od test primera preuzeto je i optimalno rešenje sa [14]. Predstavljani algoritam je za 2 primera došao do optimalnog rešenja, oba iz P skupa, P-n16-k8 i P-n20-k2.

Na slici 1 je prikazan proces dobijanja rešenja.



Slika 1: Vizuelni prikaz dobijenog rešenja



Prilikom testiranja generisani su i vizuelni prikazi rezultata u formi koja je predstavljena na slici 1, gde su za primer P-n20-k2 prikazani redom: optimalno rešenje, rešenje dobijeno Klark-Rajt algoritmom, korišćeno parcijalno rešenje i krajnje rešenje dobijeno primenom Karusel heuristike. U gornjem levom uglu, na prvoj slici, nalazi se vizualni prikaz optimalnog rešenja za dati problem, koje je peuzeto sa [14]. U gornjem desnom uglu prikazano je početno rešenje dobijeno primenom pohlepnog Klark-Rajt algoritma ušteta. U donjem levom uglu može se videti parcijalno rešenje iz kojeg je dobijeno konačno rešenje predstavljeno na poslednjoj slici. Parcijalno rešenje prikazano na trećoj slici dobijeno je uklanjanjem na slučajan način izabranih 40% ušteta iz Klark-Rajt rešenja sa prethodne slike. Krajnje rešenje dobijeno je primenom Karusel algoritma na parcijalno rešenje i u slučaju primera sa slike postignut je optimalan rezultat, odnosno može se videti da su rute na prvoj i poslednjoj slici iste, kao i funkcija cilja prikazana u zagradama iznad pojedinačnih slika.

Sumiranje dobijenih rezultata je prikazano u tabeli 2. Treba napomenuti da su pri računanju prosečnog relativnog odstupanja od optimalnog rešenja uzeti u obzir samo primeri kod kojih je algoritam došao do dopustivog rešenja.

**Tabela 2:** Zbirni rezultati postignuti primenom Karusel heuristike

Dostignuto optimalno rešenje	2/51 (3,92%)
Dobijeno nedopustivo rešenje	10/51 (19,61%)
Prosečno relativno odstupanje od optimalnog rešenja	5,33%
Poboljšanje u odnosu na Klark-Rajt alg.	28/51 (54,90%)
Dobijeno isto rešenje kao Klark-Rajt alg.	22/51 (43,14%)
Pogoršanje u odnosu na Klark-Rajt alg.	1/51 (1,96%)

#### 4. ZAKLJUČAK

Predmet ovog rada je bila primena Karusel heuristike u rešavanju problema CVRP sa homogenim voznim parkom. Kako u dostupnoj literaturi nije pronađen rad u kome je Karusel heuristika korišćena za rešavanje ovog problema, cilj je bio ispitati da li je i koliko pogodna za njegovo rešavanje. Na osnovu dobijenih rezultata, može se zaključiti da osnovna varijanta Karusel heuristike ne daje značajna poboljšanja rešenja dobijenih Klark-Rajt algoritmom, odnosno poboljšava kvalitet rešenja u samo 28 od 51 test primera i samo u 2 slučaja dolazi do optimalnog rešenja.

Dalji rad bi mogao da bude usmeren u pravcu razvijanja hibridne metode kojom bi se omogućio sofisticiraniji izbor promenljivih kojima se dopunjuje degradirano rešenje, kao i njena primena na problemu CVRP sa ograničenom homogenom flotom.

#### REFERENCE

- [1] Augerat, P., Naddef, D., Belenguer, J. M., Benavent, E., Corberan, A., & Rinaldi, G. (1995). Computational results with a branch and cut code for the capacitated vehicle routing problem.
- [2] Capobianco, G., D'Ambrosio, C., Pavone, L., Raiconi, A., Vitale, G., & Sebastiano, F. (2022). A hybrid metaheuristic for the Knapsack Problem with Forfeits. *Soft Computing*, 26(2), 749-762.
- [3] Carrabs, F., Cerrone, C., Cerulli, R., & Golden, B. (2020). An adaptive heuristic approach to compute upper and lower bounds for the close-enough traveling salesman problem. *INFORMS Journal on Computing*, 32(4), 1030-1048.
- [4] Carrabs, F., Cerrone, C., D'Ambrosio, C., & Raiconi, A. (2017). Column generation embedding carousel greedy for the maximum network lifetime problem with interference constraints. In *International Conference on Optimization and Decision Science*. Springer, Cham. 151-159.
- [5] Cerrone, C., Cerulli, R., & Golden, B. (2017). Carousel greedy: a generalized greedy algorithm with applications in optimization. *Computers & Operations Research*, 85, 97-112.
- [6] Cerrone, C., D'Ambrosio, C., & Raiconi, A. (2019) Heuristics for the strong generalized minimum label spanning tree problem. *Networks*. <https://doi.org/10.1002/net.21882>
- [7] Cerrone, C., Gentili, M., D'Ambrosio, C., & Cerulli, R. (2018). An Efficient and Simple Approach to Solve a Distribution Problem. In *New Trends in Emerging Complex Real Life Problems*. Springer, Cham. 151-159.

- [8] Cerulli, R., D'Ambrosio, C., Iossa, A., & Palmieri, F. (2022). Maximum Network Lifetime Problem with Time Slots and coverage constraints: heuristic approaches. *The Journal of Supercomputing*, 78(1), 1330-1355.
- [9] Chen, P., Huang, H. K., & Dong, X. Y. (2010). Iterated variable neighborhood descent algorithm for the capacitated vehicle routing problem. *Expert Systems with Applications*, 37(2), 1620-1627.
- [10] Dantzig, G. B., & Ramser, J. H. (1959). The truck dispatching problem. *Management science*, 6(1), 80-91.
- [11] Duhamel, C., Lacomme, P., Prins, C., & Prodhon, C. (2010). A GRASP $\times$  ELS approach for the capacitated location-routing problem. *Computers & Operations Research*, 37(11), 1912-1923.
- [12] Escobar, J., Duque, J., & García-Cáceres, R. (2022). A granular tabu search for the refrigerated vehicle routing problem with homogeneous fleet. *International Journal of Industrial Engineering Computations*, 13(1), 135-150.
- [13] Hou, Y., Liu, B., Dang, L., He, W., & Gu, W. (2022). A Local Search-based Metaheuristic Algorithm Framework for the School Bus Routing Problem. *Engineering Letters*, 30(1).
- [14] <http://vrp.galagos.inf.puc-rio.br/index.php/en/>, poslednji pristup: jun 2022. (sajt predstavljen u Uchoa, E., Pecin, D., Pessoa, A., Poggi, M., Vidal, T., & Subramanian, A. (2017). New benchmark instances for the capacitated vehicle routing problem. *European Journal of Operational Research*, 257(3), 845-858.)
- [15] Kong, H., Kang, Q., Li, W., Liu, C., Kang, Y., & He, H. (2019). A hybrid iterated carousel greedy algorithm for community detection in complex networks. *Physica A: Statistical Mechanics and its Applications*, 536, 122124.
- [16] Makajić-Nikolić, D., Pavlović, P., Vujošević, M. (2019), Rešavanje problema minimalnog pogađanja skupova pomoću Karusel heuristike. Zbornik radova 46. Simpozijuma o operacionim istraživanjima, Sym-Op-Is 2019, Kladovo, 15. – 18. septembar 2019. godine, 193-198. ISBN: 978-86-7680-363-7
- [17] Nogueveu, S. U., Prins, C., & Calvo, R. W. (2010). An effective memetic algorithm for the cumulative capacitated vehicle routing problem. *Computers & Operations Research*, 37(11), 1877-1885.
- [18] Saha, B., Suthar, K., & Kumar, A. (2020). Optimizing Generalized Capacitated Vehicle Routing Problem Using Augmented Savings Algorithm. In *Computational Intelligence in Data Mining* (pp. 527-541). Springer, Singapore.
- [19] Singamsetty, P., & Thenepalle, J. (2021). Designing optimal route for the distribution chain of a rural LPG delivery system. *International journal of industrial engineering computations*, 12(2), 221-234.
- [20] Subramanian, A., Uchoa, E., & Ochi, L. S. (2013). A hybrid algorithm for a class of vehicle routing problems. *Computers & Operations Research*, 40(10), 2519-2531.

## THE MODIFIED U-GRASP METHOD FOR THE TWO-DIMENSIONAL VECTOR BIN PACKING PROBLEM

DORĐE STAKIĆ<sup>1</sup>, ANA ANOKIĆ<sup>2</sup>

<sup>1</sup> University of Belgrade – Faculty of Economics and Business, djordje.stakic@ekof.bg.ac.rs

<sup>2</sup> Academy of Technical and Art Applied Studies Belgrade – Department School of Information and Communication Technologies, ana.anokic@ict.edu.rs

**Abstract:** *This paper proposes an improved version of one of the three previously published variants of the Greedy randomized adaptive search procedure - GRASP for solving the two-dimensional heterogeneous vector bin packing problem (2DHet-VBPP). The method is improved by technical and algorithmic modifications. Initially developed for the 2DHet-VBPP, the proposed GRASP is adapted to solve the homogeneous vector bin packing problem (2DHom-VBPP). The study presents the results of parameter tuning tests and compares the obtained results with those produced by the previous methods. The improved version finds the optimal solutions of some new benchmark instances and provides better solutions in shorter time for some benchmark instances.*

**Keywords:** *GRASP method, bin packing problem, metaheuristics, optimization*

### 1. INTRODUCTION

Since the proposal of the GRASP method in [4], it has been applied independently or as a part of different hybrid solution approaches to numerous and various optimization problems: vehicle routing, pickup and delivery, location problems, packing problems etc. More about this metaheuristic method can be found in [5, 7, 12, 13], and more about its applications in [3, 6].

The two-dimensional heterogeneous vector bin packing problem (2DHet-VBPP) consists of finding the optimal packing of the given set of items (packages) into the set of different bins (containers), which are limited by the total mass and volume of its items. The objective is to minimize the total transportation cost. The 2DHet-VBPP is proved to be a NP-hard optimization problem, which makes it a challenge for designing an efficient solution approach. It appears in [9], as the problem of general dimension, whereas its two-dimensional variant can be found in the early work by [11]. Three solution methods - the simple greedy heuristics named First Fit by Ordered Deviation (FFOD), Simulated Annealing (SA) and Column Generation (CG) were applied in [11]. Unlike SA, FFOD was the fastest, but with a large deviation of the obtained solutions from the best one, while the exact CG did not provide solutions for instances with more than 50 items. In [8], the 34 simple heuristics were developed for the variant of the 2DHet-VBPP, with the main goal to determine whether the considered items can be placed into the set of bins or not. A more recently published work [14] proposed the three metaheuristics based on the GRASP method: a Basic, Uniform and a Reactive GRASP. The Uniform GRASP method (U-GRASP) provided better results for shorter running time compared to the other two. As the two-dimensional homogeneous Vector Bin Packing Problem (2DHom-VBPP) received more attention in literature, and considering the fact that there are standard benchmarks [2] for comparison with [1] and [8], the proposed U-GRASP in [14] was adapted for solving the 2DHom-VBPP, resulting in more optimal solutions.

In the present paper we attempt to improve the U-GRASP, the variant which was the most promising solution approach of the three variants in [14] for the 2DHet-VBPP and 2DHom-VBPP. The improved method proposal is named the modified U-GRASP. We introduce several modifications and analyze its performance compared to the previous version.

The paper is further organized in the following way: in Section 2 notations and the mathematical formulation of the problem are introduced; the proposed GRASP method with its modifications is described in Section 3, while Section 4 contains experimental results. Finally, the last section contains some concluding remarks pertaining to the research.

### 2. MATHEMATICAL FORMULATION

The 2DHet-VBPP is proved to be NP-hard in [10]. Several similar mathematical formulations of the considered problem appear in literature. In order to present one of them, the notations in Table 1 are introduced.

**Table 1:** Notation

Symbol	Description
$nt$	the number of types of bins
$np$	the number of items
$C_t$	price of bin of type $t$ , for $t \in \{1, \dots, nt\}$
$LV_t$	limit of volume of a bin of type $t$ , for $t \in \{1, \dots, nt\}$
$Lm_t$	limit of mass of a bin of type $t$ , for $t \in \{1, \dots, nt\}$
$V_i$	volume of item $i$ , for $i \in \{1, \dots, np\}$
$m_i$	mass of item $i$ , for $i \in \{1, \dots, np\}$
$Ln_t$	the upper bound on the number of bins of type $t$ , for $t \in \{1, \dots, nt\}$

Mathematical formulations for the considered problem can be found in [8] and [11]. The proposed mathematical model uses the following sets of variables:

- Binary variables  $p_{ijt}$ , for  $i \in \{1, \dots, np\}$ ,  $j \in \{1, \dots, Ln_t\}$ ,  $t \in \{1, \dots, nt\}$ , defined by  $p_{ijt} = 1$  if an item  $i$  is placed in a bin  $j$  of type  $t$ , otherwise  $p_{ijt} = 0$ .
- Binary variables  $k_{jt}$ , for  $j \in \{1, \dots, Ln_t\}$ ,  $t \in \{1, \dots, nt\}$  take the value  $k_{jt} = 1$  if a bin  $j$  of type  $t$  is used in transport, and  $k_{jt} = 0$  otherwise.

The considered 2DHet-VBPP can be described by the following Integer linear programming (ILP) model, as follows:

$$\min \quad \sum_{t=1}^{nt} \sum_{j=1}^{Ln_t} k_{jt} \cdot C_t \quad (1)$$

$$\sum_{t=1}^{nt} \sum_{j=1}^{Ln_t} p_{ijt} = 1, \quad i \in \{1, \dots, np\} \quad (2)$$

$$\sum_{i=1}^{np} p_{ijt} \cdot V_i \leq k_{jt} \cdot LV_t, \quad t \in \{1, \dots, nt\}, j \in \{1, \dots, Ln_t\} \quad (3)$$

$$\sum_{i=1}^{np} p_{ijt} \cdot m_i \leq k_{jt} \cdot Lm_t, \quad t \in \{1, \dots, nt\}, j \in \{1, \dots, Ln_t\} \quad (4)$$

$$p_{ijt} \in \{0, 1\}, \quad i \in \{1, \dots, np\}, t \in \{1, \dots, nt\}, j \in \{1, \dots, Ln_t\} \quad (5)$$

$$k_{jt} \in \{0, 1\}, \quad t \in \{1, \dots, nt\}, j \in \{1, \dots, Ln_t\} \quad (6)$$

The mathematical expression of the objective function (1), which should be minimized, is designed to reflect the total transport costs as the sum of costs over the set of used bins depending on their types. The purpose of constraints (2) is to ensure that each item is placed into exactly one bin. Constraints (3) and (4) refer to the limits of the total volume and mass of each bin, depending on its type, while constraints (5) and (6) define the type of variables used in the formulation.

The 2DHom-VBPP is a special case of the 2DHet-VBPP where all bins are of the same type. In other words, the number of different type of bins is set to  $nt = 1$  for this purpose.

### 3. THE MODIFIED U-GRASP METHOD

The GRASP method is an iterative procedure consisting of the two main phases: greedy randomized construction (GRC) and local search (LS), which alternate until the stopping criterion is met. The algorithm of the uniform GRASP method and its GRC phase are described in details in [14]. The upper bound on the running time ( $t_{max}$ ) is used as the termination criterion. Besides  $t_{max}$ , the U-GRASP method uses another real parameter  $k$ . The GRC phase of the U-GRASP in [14] uses the two candidate lists, which initially consist of all items sorted in a decreasing order according to their mass and volume, respectively. These lists are used depending on the type of a randomly chosen container. The parameter  $k$  defines the range of values for the real number  $\alpha$ , which is randomly chosen according to the uniform distribution from the interval  $[0, 1/k]$ . This value is used to define the

bounds of the restrictive candidate list. An item is randomly chosen from a restrictive candidate list and placed in the first bin if possible, depending on the limits of its mass and volume. If the attempt fails, GRC continues with the second bin, the third one, etc. Candidate lists are refreshed and the procedure of solution construction continues until all items are placed in bins, which is when the local search phase begins.

In accordance with [14], in our modified U-GRASP we used the pair  $\mathcal{S} = (S, a)$  for solution representation. The pair consists of an integer matrix  $S$  and an array  $a$ . Each row of  $S$  corresponds to one bin and contains indices of items that are placed in the considered bin. The dimension of matrix  $S$  in [14] was  $np \times np$ , which is chosen as the largest possible, having in mind that the number of bins is not limited and each item can be placed in each bin. On the other hand, unlike [14], in the present paper we have decreased the number of columns in matrix  $S$  so that  $S$  is no longer quadratic; its dimension is  $np \times n^*$ , where  $n^*$  is calculated according to the following formula:

$$n^* = \max_{1 \leq t \leq n_t} \min \left\{ \max_{1 \leq k \leq n} \sum_{j=1}^k m'_j \leq Lm_t, \max_{1 \leq k \leq n} \sum_{j=1}^k V'_j \leq LV_t \right\}, \quad (7)$$

For each bin type, the two maximal numbers of items that can be placed into the bin of the considered type are calculated, bearing in mind its limits in terms of mass and volume, respectively. The smaller of these two values is compared to the corresponding values of the remaining types of bins and,  $n^*$  is their maximum. The array  $a$  of length  $n$ , in a solution  $\mathcal{S} = (S, a)$ , contains the types of bins used in transport, as in [14].

Besides this modification, the LS phase is also improved. Its structure is presented in Algorithm 1. At the beginning of the LS phase, we have included the following modification. With the probability of 50%, which is controlled by the function *random*, the algorithm applies the function *Improve By Type*. This function is also included at the end of the LS phase, and consists of the attempts to change the bin type one by one in order to reduce the total transport cost. If all items from one bin can be placed into the bin with lower price, then this movement leads to a lower total cost. After this step, all items are sorted in increasing order according to their mass and the bins are sorted in increasing order as well according to the total mass of their items. Starting with the item of the largest mass from the bin of the smallest total mass, the algorithm of the LS phase tries to move the items of the considered bin into the next bin one by one. When it is not possible to move an item into any of the remaining bins, all the items in the considered bin are returned. When the sequence is repeated on all bins, and after applying the function *Improve By Type* one more time, the LS phase is finished.

---

**Algorithm 1** LS phase of the proposed modified U-GRASP method

---

```

procedure LOCALSEARCH( $S, a$ )
  if (random(0, 1) = 0) then
    ( $S, a$ )  $\leftarrow$  Improve By Type( $S, a$ );
  ( $S', a'$ )  $\leftarrow$  sorted( $S, a$ );
  for ( $j \leftarrow 1; j \leq nBin(S'); j++$ ) do
    for ( $i \leftarrow nItem(S', j); i \geq 1; i--$ ) do
      for ( $k \leftarrow nBin(S'); k \geq 1; k--$ ) do
        if (Transfer( $S', j, i, k$ )) then
          break;
      if (EmptyBin( $S', j$ )) then
        ( $S', a'$ )  $\leftarrow$  sorted( $S', a'$ );
        ( $S, a$ )  $\leftarrow$  ( $S', a'$ );
      else
        ( $S', a'$ )  $\leftarrow$  ( $S, a$ );
  ( $S', a'$ )  $\leftarrow$  Improve By Type( $S', a'$ );

```

---

If the LS phase provides a better solution, this solution replaces the current best one. At this stage, one iteration of the modified U-GRASP method is completed and the next one begins, unless the stopping criterion is satisfied.

In order to compare the proposed method with the relevant benchmarks from literature pertaining to the 2DHom-VBPP, our implementation is adapted to a homogeneous variant of the problem, in the same way as in [14]. All bins are considered to be of the same type and therefore, the coefficients of the objective function are equalized by 1. In this case, the purpose is to minimize the total number of used bins. As the bins are of the same type, at the beginning of each GRC phase, one of the two candidate lists (according to mass and volume) is chosen randomly with the same probability.

#### 4. EXPERIMENTAL RESULTS

All experimental results were obtained using the Intel Core i7-2600 CPU 3.40GHz processor with 12GB RAM under Linux operating system. The tests were conducted on the set of 10 large instances with: 50, 70, 100, 120, 150, 200, 350, 500, 750 and 1000 items, that are described in [14]. Running time is limited to 120 s, i.e. the parameter  $t_{max} = 120$  s. Considering the fact that the GRASP is generally a stochastic method, each instance was tested 30 times. The best obtained function value and the standard deviation of all solutions from the best one through 30 consecutive runs are reported. In order to estimate the quality of the obtained solutions, we used exact CPLEX solver and the mathematical model presented in Section 2 on the same computer platform with the running time limited to 1 h, which corresponds to the total running time of a modified U-GRASP method in its 30 runs.

To achieve better performances of the designed method, the value of parameter  $k$  was determined using parameter tuning tests. The following values were tested  $k \in \{1, 2, \dots, 12\}$  on the set of the considered instances with the time limit of 20 s and 10 consecutive runs. The results, presented in Table 2, consist of the best obtained function values. Instances are characterized by the number of items in the first column. The best objective function values for each instance are bolded.

**Table 2:** Parameter tuning tests for  $k$

Instance number of items	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
50	<b>32380</b>	32393	32393	<b>32380</b>	<b>32380</b>	<b>32380</b>
70	48885	48714	48859	48859	48714	48846
100	68342	67492	67637	67637	67466	67611
120	70952	70833	70794	70781	70102	70794
150	92397	92332	92358	92944	91363	<b>91350</b>
200	130904	129481	129507	129468	129494	129297
350	227077	225496	226214	225403	225719	225496
500	319474	317880	319461	319487	<b>317696</b>	317854
750	480150	478478	479577	479549	478893	479940
1000	648125	647446	645510	645733	644541	644963
Instance number of items	$k = 7$	$k = 8$	$k = 9$	$k = 10$	$k = 11$	$k = 12$
50	<b>32380</b>	<b>32380</b>	<b>32380</b>	<b>32380</b>	<b>32380</b>	<b>32380</b>
70	48714	48714	48022	<b>47996</b>	48833	48701
100	67598	66787	67479	<b>66774</b>	67492	67466
120	70807	70089	70781	70794	<b>70076</b>	70794
150	92319	92226	91534	91534	92332	92213
200	129468	<b>128802</b>	129455	129468	129323	129455
350	<b>223757</b>	225390	225470	225548	225364	224633
500	317893	318092	318717	317841	318533	317960
750	479940	479787	479142	<b>477587</b>	479157	479287
1000	<b>643263</b>	643336	644766	643540	645080	644126

The largest number of the best found solutions (4 of 10 tested instances) were obtained for  $k = 10$ . Therefore, this value was chosen for the final tests.

As the U-GRASP from [14] was tested using different computer platforms, the tests of U-GRASP were repeated on the platform used in the present work, for adequate comparison. The final results are presented in Table 3. Besides the best objective function value (*best sol.*) obtained within 30 consecutive runs, for each instance, the average running time until the best solution is reached ( $t_{best}$ ) for each instance and the percentage standard deviation of the obtained solutions from the best one ( $\sigma$ ) are presented.

For instances with 500 items or less, the best objective function value is provided by CPLEX solver. Note that the time limit for CPLEX is 1h, while the U-GRASP and the modified U-GRASP are limited to 120 s in each run. In the case of the last two larger instances (with 750 and 1000 items), the best solution is reached by our modified U-GRASP method. For the remaining 8 instances, it can be concluded that the U-GRASP is better for instances with up to 200 items, but the modified U-GRASP provides better quality of solutions for instances with more than 200 items. Based on the average values of  $t_{best}$  and  $\sigma$  on the considered set of instances, it can be concluded that the modified U-GRASP is slightly faster than the U-GRASP, as the average values in columns

**Table 3:** The proposed modified U-GRASP method for the 2DHet-VBPP compared to CPLEX and U-GRASP

Instance number of items	CPLEX	U-GRASP			modified U-GRASP		
		<i>best sol.</i>	$t_{best}(s)$	$\sigma(\%)$	<i>best sol.</i>	$t_{best}(s)$	$\sigma(\%)$
50	<b>31530</b>	31543	23.92	2.63	32367	49.65	0.05
70	<b>48688</b>	47996	59.52	1.42	47996	44.12	1.57
100	<b>65898</b>	66761	68.87	0.96	66774	55.91	1.13
120	<b>68376</b>	70089	63.86	1.43	70076	67.83	0.66
150	<b>88056</b>	90513	58.14	1.69	91482	65.11	0.95
200	<b>124394</b>	127965	65.63	1.45	128631	48.73	0.73
350	<b>219321</b>	225810	61.44	0.86	224475	62.56	0.46
500	<b>310801</b>	320863	64.06	0.66	317149	54.28	0.56
750	4910250	482636	59.87	0.76	<b>477587</b>	60.86	0.54
1000	6547000	650168	51.66	1.30	<b>641267</b>	50.82	0.58

$t_{best}$  are 55.99 s and 57.70 s, for these two methods, respectively. The difference in standard deviations is more significant, 0.72%, for the modified U-GRASP and 1.32% for the U-GRASP method, which shows a better stability of the modified U-GRASP.

The adaption of our modified U-GRASP method for the 2DHom-VBPP is tested on the benchmarks from literature. On the set of 400 standard instances ([2]) for the 2DHom-VBPP, the group of greedy heuristics from [8] solved 249 with 38 new optimal solutions among 70 instances unsolved in [1]. In the previous research [14], the U-GRASP reached 245 optimal solutions as well as 38 from the 70 unsolved in [1] during the time limit of 120 s. Extending the time limit from 120 s to 600 s, the 7 new optimal solutions were obtained, which gave a total of 252 solved instances of 400 and the 2 new of 70 unsolved, the total of 40 solved instances of these 70.

The modified U-GRASP method has shown better performances. With the time limit of only 1 s, the 235 optimal solutions on the set of 400 instances were obtained, while with the time limit increased to 10, 60 and 360 s, the number of reached optimal solutions were: 245, 250 and 254, respectively. The number of obtained optimal solutions by the modified U-GRASP method with the time limit of 360 s per each run are shown, as shown in Table 4. Instances are classified into classes, as described in [14], which contains the same table format referring to the results of U-GRASP. Considering the number of optimal solutions, the obtained results which are better than the ones from the remaining benchmarks are bolded.

**Table 4:** The number of optimal solutions of the 2DHom-VBPP, obtained by the modified U-GRASP method with  $t_{max} = 360$  s

class	25/24	50/51	100/99	200/201
1	10	10	2	0
2	10	10	9	5
3	10	10	10	5
4	9	<b>10</b>	<b>10</b>	0
5	10	<b>10</b>	<b>10</b>	0
6	10	9	0	0
7	9	6	3	0
8	10	10	10	10
9	10	10	0	0
10	7	0	0	0
Total	95	85	54	20

With the number of 254 solved instances, our modified U-GRASP method is better than the methods proposed in [8] and [14], which solved 249 and 252 instances, respectively, on the same data set. From 70 instances that remained unsolved in [1], the modified U-GRASP method reaches optimal solutions for 40 instances with the time limit of 360 s, while in [14] 38 instances were solved for 120 s and 40 for the time limit of 600 s, per each run.

## 5. CONCLUSION

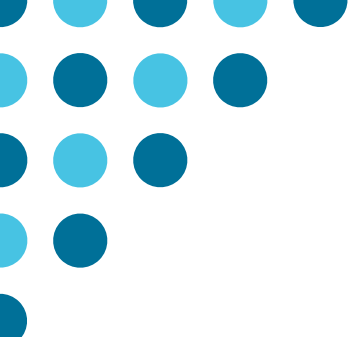
The implementation of the uniform GRASP method for the two-dimensional heterogeneous vector bin packing problem, published in [14] has been improved through several modifications. These modifications pertain to reducing the number of columns in the solution representation matrix (as the result of a better estimation of

the number of necessary bins) and the improvement of the local search phase. The parameter tuning tests were performed to determinate the parameter value which resulted in better quality solutions. The obtained results show improved efficiency of the method, especially for the instances with the large number of items. The method was further adapted for solving the homogenous problem and was able to produce two more optimal solutions for a shorter running time on the standard set of instances. In the future, we plan to extend the set of instances, develop different metaheuristic methods for the considered problems and adapt the developed method for use with other similar problems.

## REFERENCES

- [1] Filipe Brandao and João Pedro Pedroso. Bin packing and related problems: general arc-flow formulation with graph compression. *Computers & Operations Research*, 69:56–67, 2016.
- [2] Alberto Caprara and Paolo Toth. Lower bounds and algorithms for the 2-dimensional vector packing problem. *Discrete Applied Mathematics*, 111(3):231–262, 2001.
- [3] Xavier Delorme, Xavier Gandibleux, and Joaquin Rodriguez. GRASP for set packing problems. *European Journal of Operational Research*, 153(3):564–580, 2004.
- [4] S. Feo and M.G.C. Resende. A probabilistic heuristic for a computationally difficult set covering problem. *Oper. Res. Lett.*, 8:67–71, 1989.
- [5] S. Feo and M.G.C. Resende. Greedy randomized adaptive search procedures. *J. Global Optim.*, 6:109–133, 1995.
- [6] P. Festa and M. G.C. Resende. An annotated bibliography of GRASP–part II: Applications. *Int. T. Oper. Res.*, 16(2):131–172, 2009.
- [7] P. Festa and M.G.C. Resende. An annotated bibliography of GRASP–part I: Algorithms. *Int. T. Oper. Res.*, 16(1):1–24, 2009.
- [8] Michaël Gabay and Sofia Zaourar. Vector bin packing with heterogeneous bins: application to the machine reassignment problem. *Annals of Operations Research*, 242(1):161–194, 2016.
- [9] Michael R Garey, Ronald L Graham, David S Johnson, and Andrew Chi-Chih Yao. Resource constrained scheduling as generalized bin packing. *Journal of Combinatorial Theory, Series A*, 21(3):257–298, 1976.
- [10] Michael R Garey and David S Johnson. Computers and intractability: a guide to np-completeness, 1979.
- [11] Bernard T Han, George Diehr, and Jack S Cook. Multiple-type, two-dimensional bin packing problems: Applications and algorithms. *Annals of Operations Research*, 50(1):239–261, 1994.
- [12] M.G.C. Resende and C.C. Ribeiro. Greedy randomized adaptive search procedures: Advances, hybridizations, and applications. In M. Gendreau and Potvin J.Y., editors, *Handbook of metaheuristics*, pages 283–319. Springer, 2010.
- [13] M.G.C. Resende and C.C. Ribeiro. Grasp: Greedy randomized adaptive search procedures. In E.K. Burke and G. Kendall, editors, *Search Methodologies - Introductory tutorials in optimization and decision support systems*, pages 287–312. Springer, 2014.
- [14] Đorđe Stakić, Ana Anokić, and Raka Jovanović. Comparison of different grasp algorithms for the heterogeneous vector bin packing problem. In *2019 China-Qatar International Workshop on Artificial Intelligence and Applications to Intelligent Manufacturing (AIAIM)*, pages 63–70. IEEE, 2019.





# INFORMACIONI SISTEMI I TEHNOLOGIJE

---

## INFORMATION SYSTEMS AND TECHNOLOGIES



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business



## GENERATOR PROGRAMA – LUČA

### PROGRAM GENERATOR - LUČA

SINIŠA VLAJIĆ<sup>1</sup>, VOJISLAV STANOJEVIĆ<sup>1</sup>, MILOŠ MILIĆ<sup>1</sup>, ILIJA ANTOVIĆ<sup>1</sup>, DUŠAN SAVIĆ<sup>1</sup>

<sup>1</sup> Fakultet organizacionih nauka, Katedra za softversko inženjerstvo, Beograd, sinisa.vlajic@fon.bg.ac.rs, vojislav.stanojevic@fon.bg.ac.rs, milos.milic@fon.bg.ac.rs, ilija.antovic@fon.bg.ac.rs, dusan.savic@fon.bg.ac.rs

**Rezime:** U ovom radu je razmatran sopstveni generator programa – Luča, koji generiše programske komponente na osnovu Luča šablona i Luča struktura. Luča šablona sadrže generički programski kod i Luča izraze, dok Luča struktura definiše domen problema i ona sadrži specifične delove programskog koda. Proces generisanja komponente pomoću generatora programa Luča je detaljno objašnjen uz konkretan primer u kome se generiše sql skript datoteka (ConterGlumac.sql) na osnovu Luča šablona (MySQLCounterTableTemplate.lctm) i Luča strukture (Glumac.java). Na kraju je izvršena analiza generatora programa Luča i navedeni su pravci daljeg istraživanja.

**Ključne reči:** Generatori programa, Domenski specifični jezici, Generator programa - Luča

**Abstract:** In this paper, we present our own program generator - Luca, which generates program components based on Luca templates and Luca structures. Luca templates contain generic program code and Luca expressions, while Luca structure define a specific domain of the problem and thus contains program code that is specific to a particular domain. The process of generating a program component using the program generator - Luca is explained in detail through a case study that includes an sql script file generation (CounterGlumac.sql) based on the Luca template (MySQLCounterTableTemplate.lctm) and the Luca structure (Glumac.java). Finally we present the results of analysis of the program generator - Luca and the further research directions.

**Keywords:** Program generators, Domain specific languages, Program generator - Luca

## 1. UVOD

Postoje različiti pristupi koji razmatraju generisanje programa na osnovu modela [3,7]. Istraživanje koje su sprovedeli [5] razmatra pristupe u razvoju softvera vođenog modelima i kroz sopstveni CSTL jezik definišu iskaze koji sadrže konkretnu sintaksu u izvornom i odredišnom programskom jeziku. Pored toga, jezik omogućava definisanje dodatnih fragmenata u vidu izraza, naredbi i pravila koja treba primeniti [5].

U istraživanju [9] definisan je Metaedit+ jezik koji omogućava domenski specifično modelovanje. Specifičnost ovog pristupa ogleda se u definisanju domenski specifičnog jezika kao metamodela koji sadrži domenske koncepte, attribute, pravila, simbole i generatore [9]. Na osnovu definisanog jezika za modelovanje mogu se kreirati konkretni modeli i može se izvršiti generisanje programskog koda.

Istraživanje sprovedeno od strane Kroiss, Koch, i Knapp podrazumeva primenu razvoja softvera vođenog modelima za generisanje veb aplikacija [4]. U tom smislu razvijen je UWE4JSF dodatak zasnovan na Eclipse EMF platformi i razvojnom okruženju koji omogućava generisanje veb aplikacija na osnovu prethodno kreiranih UWE meta-modela [4]. Kao rezultat procesa generisanja dobijaju se komponente veb aplikacije zasnovane na JSF platformi koja omogućava izradu bogatog korisničkog interfejsa za interakciju sa korisnikom.

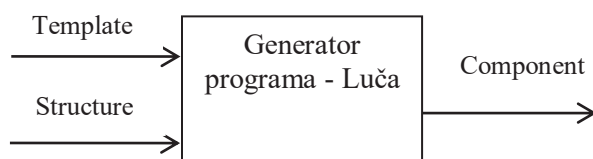
Posmatrani pristupi koriste šablone i domenski specifične jezike za definisanje modela i njegove transformacije. Komplementarni pristup može podrazumevati izradu generatora koda primenom metode supstitucije regularnih izraza [2]. U tom smislu se razmatra konstrukcija radnog okvira (engl. framework) baziranog na regularnim izrazima. Nakon kreiranja skeleta projekta, kao ulaz u generator prihvata se referentni programski kod iz prethodnog projekta i vrši parametrizacija njegovih komponenti [2]. Pored toga, kao ulaz se prihvataju programski kod i konfiguracione datoteke tekućeg projekta čija se parametrizacija takođe vrši [2]. Na osnovu definisanih regularnih izraza generiše se modifikovani programski kod tekućeg projekta čiju je promenu moguće izvršiti od strane programera [2].

Na Fakultetu organizacionih nauka u okviru Laboratorije za softversko inženjerstvo razvijen je sopstveni generator programa – Luča, koji je napisan u Java programskom jeziku. Generator programa – Luča opisuje domen problema preko Luča strukture, koja predstavlja hijerarhijski model. Generički delovi programa su opisani pomoću Luča šablona, koji se koriste za različite domene problema. Luča šabloni sadrže tekst i Luča izraze. U procesu generisanja komponenti Luča šabloni, odnosno Luča izrazi se parametrizuju sa Luča strukturama.

## 2. GENERATOR PROGRAMA – LUČA

Rezultat procesa razvoja sofvera je softverski sistem koji se sastoji iz skupa softverskih komponenti. Softverske komponente se sastoje iz generičkih i specifičnih delova programskog koda. Specifični delovi programskog koda se koriste za konkretan domen problema, dok se generički delovi programskog koda mogu koristiti za više različitih domena problema.

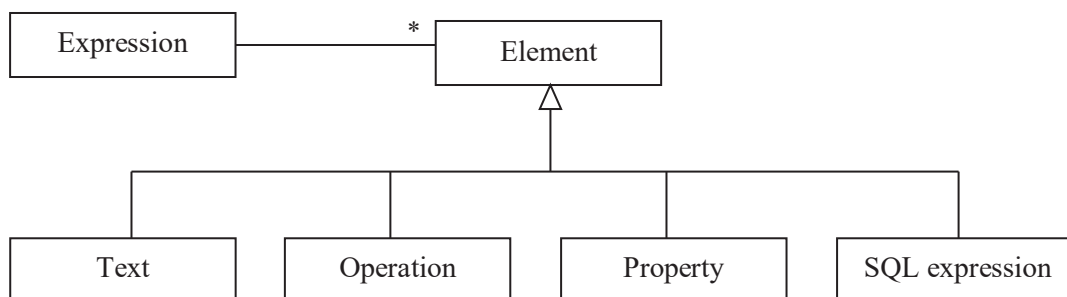
Generator programa – Luča je programski alat (Slika 1) koji omogućava generisanje softverskih komponenti (Component) na osnovu Luča šablona (Template) i Luča struktura (Structure).



Slika 1: Generator programa - Luča

**Luča šablon** je datoteka koja se sastoji iz: a) generičkog programskog koda, koji je nepromenljiv i b) Luča izraza. **Luča izraz** (Expression) se sastoji iz više Luča elemenata (Element). Postoje sledeći Luča elementi: a) Tekst (Text) b) Luča operacije (Operation) c) Svojstva strukture (Property) i d) SQL izraz (SQL expression)

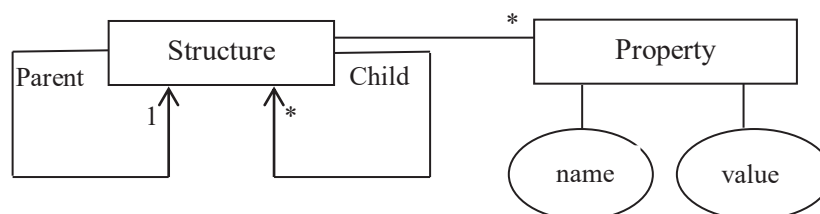
Luča izraz je predstavljen na slici 2.



Slika 2: Luča izraz

**Luča struktura** (Slika 3) je datoteka koja sadrži specifičan programski kod koji je vezan za konkretan domen problema. Luča struktura se sastoji iz: a) drugih Luča struktura i b) svojstava (property).

Luča struktura predstavlja hijerarhijski model, pri čemu Luča struktura može imati jednog roditelja i više dece. **Svojstvo Luča strukture** sastoji se iz para: naziv (name) i vrednost (value).



Slika 3: Luča struktura

### 3. PROCES GENERISANJA KOMPONENTE

Proces generisanja komponente sastoji se iz sledećih koraka:

- a) kreira se Luča struktura
- b) kreira se tabela Luča strukture u bazi podataka
- c) pune se svojstva Luča strukture sa vrednostima
- d) generiše se komponenta na osnovu Luča šablona i Luča strukture

**Kreiranje Luča strukture** obrazuje hijerarhijski model, koji se sastoji iz **osnovne strukture** i svih struktura izvedenih iz nje, što ujedno čini i oblast važenje Luča strukture. Osnovna struktura je na vrhu Luča strukture. Naziv Luča strukture i osnovne strukture je isti. Oblast važenje Luča strukture i osnovne strukture je takođe ista. Nakon kreiranja Luče strukture pamti se lista naziva struktura (**ls**) koje je sačinjavaju. Zatim se dodaju svojstva do struktura (pripadajuća svojstva). Naziv svojstva ne može biti isti kao nazivi struktura Luča strukture. Takođe naziv svojstva ne može da bude "structureName". Nazivi svojstava se dodaju do liste različitih imena svojstava (**lp**) Luča strukture. Lista lp definiše sva različita imena svojstava Luča strukture, koja može biti uključena u Luča izraze, koji su definisani u oblasti važenja Luča strukture. Nakon dodavanja svojstava do struktura dodaju se nasledena svojstva njihovih nadređenih struktura. Strukture ne formiraju posebne objekte od nasledenih svojstava već dobijaju reference na svojstva koja nasleduju. Struktura koja nasleduje svojstva, od nadređenih struktura koje imaju ista svojstva, nasleduje svojstva od njihovih najbližih nadređenih struktura. Nasledena svojstva ne mogu da prekriju ista pripadajuća svojstva neke strukture. Neka struktura može imati najviše jednu nadređenu strukturu (roditelj struktura). Osnovna struktura je korenska struktura i ona nema roditelja. Neka struktura može imati više podređenih struktura (deca strukture). Za svaku strukturu se kreira lista različitih imena svojstava osim za osnovnu strukturu, jer je ona već bila kreirana. Kreirana lista definiše sva različita imena svojstava neke strukture, koja mogu biti uključena u Luča izraze, koji su definisani u oblasti važenja te strukture. Oblast važenja neke strukture obuhvata tu strukturu i strukture izvedene iz nje.

Proces **kreiranja Luča tabele u bazi podataka** se sastoji iz nekoliko koraka:

- a) Kreira se Luča tabela, koja ima isti naziv kao i Luča struktura. Luča tabeli se dodaju kao atributi:
  - Naziv strukture (structureName).
  - Lista naziva struktura (**ls**) koje sačinjavaju Luča strukturu.
  - Lista različitih imena svojstava (**lp**) Luča strukture.

Ukoliko se lista **ls** sastoji iz  $m$  struktura, nazive struktura ćemo predstaviti na sledeći način:  $strName_1, strName_2, \dots, strName_m$ .

Ukoliko se lista **lp** sastoji iz  $n$  svojstava, nazive svojstava ćemo predstaviti na sledeći način:  $propName_1, propName_2, \dots, propName_n$ .

Luča tabela ima sledeću strukturu:  $structureName, propName_1, propName_2, \dots, propName_n, \dots, strName_1, strName_2, \dots, strName_m$

- b) Kreiraju se i ubacuju u tabelu slogovi. Broj slogova je isti kao i broj struktura koji sačinjavaju Luča strukturu. Svaki slog je vezan za jednu od struktura. Vrednosti svojstava svih slogova tabele se na početku postavljaju na null. Vrednosti atributa  $strName_1, strName_2, strName_3, \dots, strName_m$  dobijaju logičku vrednost (true ili false), u zavisnosti od oblasti važenja tih struktura. Npr: ukoliko  $strName_1$  nasleduje  $strName_2$  a  $strName_2$  nasleduje  $strName_3$  tada se ubacuje slog:

$(structureName = "strName_1", strName1 = true, strName2 = true, strName3 = true)$

$(structureName = "strName_2", strName1 = false, strName2 = true, strName3 = true)$

$(structureName = "strName_3", strName1 = false, strName2 = false, strName3 = true)$

Nakon kreiranja i punjenja Luča tabele **pune se svojstva Luča strukture sa vrednostima**. Između svojstava iste ili različitih struktura mogu postojati zavisnosti, koje imaju uticaj na redosled punjenja svojstava. Svakom svojstvu se dodeljuje nivo punjenja (load level), koji određuje kada će svojstvo biti napunjeno. Svojstva koja ima loadLevel = 1 (prvi nivo punjenja), prva će biti napunjena. Vrednost svojstva koje ima loadLevel=1 je neka String vrednost. Svaki put kada se neko svojstvo strukture napuni, tada se ažurira odgovarajući slog Luča tabele koji je povezan sa tom strukturom. Na primer, ako je svojstvo propName1 strukture StrName2 napunjeno sa vrednošću  $v1$  tada se ažurira sledeći slog u Luča tabeli:

$(structureName = "strName_2", propName1 = v1)$

Takođe i kod svih svojstava, koja imaju referencu na svojstvo koje je napunjeno, ažuriraju se u Luča tabeli odgovarajući slogovi, koji su povezani sa strukturama tih svojstava.

Kod svojstava koja imaju `loadLevel = 2` (drugi nivo punjenja), vrednost svojstva je Luča izraz. Luča generator, na osnovu Luča izraza, generiše String vrednost i dodeljuje je do odgovarajućeg svojstva koje se puni. Proces generisanja vrednosti iz Luča izraza sastoji se iz nekoliko koraka:

- Luča interpreter kontroliše da li je Luča izraz regularan izraz.
- Luča interpreter prepoznaje iz Luča izraza Luča elemente: tekst, operacije, svojstva i SQL izraze.
- Luča generator pravi SQL upit na osnovu Luča elemenata i izvršava upit nad Luča tabelom. Rezultat upita je String vrednost koja se dodeljuje do svojstva koje se puni.

Proces punjenja svojstava koja imaju `loadLevel = 3,4,5, ...` nastavlja se na sličan način kao proces punjenja svojstava koja imaju `loadLevel = 2`.

Nakon punjenja svojstava strukture prelazi se na **generisanje komponente na osnovu Luča šablona i Luča strukture**. Proces generisanja komponente sastoji se iz nekoliko koraka. Sledećih 5 koraka obavlja Luča interpreter:

- čita Luča šablon datoteku liniju po liniju,
- nalazi Luča izraze u pročitanoj liniji (u jednoj liniji može biti više Luča izraza),
- procesira Luča izraze, jedan po jedan,
- kontroliše da li je Luča izraz regularan izraz i
- prepoznaje iz Luča izraza Luča elemente: tekst, operacije, svojstva i SQL izraze.

Na kraju procesa generisanja komponente Luča generator pravi SQL upit na osnovu Luča elemenata i izvršava upit nad Luča tabelom. Rezultat upita je String vrednost koja se dodeljuje komponenti koja se generiše, na poziciji koja odgovara poziciji Luča izraza koji se procesira.

#### 4. PRIMER PROCESA GENERISANJA KOMPONENTE

U okviru projekta `MySQLCounterTable` napisali smo program koji generiše komponentu (`CounterGlumac.sql`) na osnovu Luča strukture (`Glumac`) i Luča šablona (`MySQLCounterTableTemplate.lctm`).

```
public class MySQLCounterTable {
    public static void main(String[] args) throws IOException, FileNotFoundException, ClassNotFoundException, InterruptedException {
        String projectName = "MySQLCounterTable";
        String componentFilePath = rootPath + "src/Generated/" + projectName + "/";
        String templateFilePath = rootPath + "src/Templates/";
        Project pr = new Project(projectName);
        Structure s = new Glumac(pr);

        pr.addLucaFile(s, "MySQLCounterTableTemplate.lctm", "CounterGlumac.sql", templateFilePath, componentFilePath);
        LucaGenerator.generateComponents(pr);
    }
}
```

Luča šablon `MySQLCounterTableTemplate.lctm` ima sledeći izgled:

```
#1 CREATE DATABASE /*!32312 IF NOT EXISTS*/~! dbname` /*!40100 DEFAULT CHARACTER SET utf8 */;
#1 USE ~! dbname`;

#1 CREATE TABLE IF NOT EXISTS ~! tl` (
#* ~a `at` tpMySQL sizeSQL Null defaultMySQL,
#1 PRIMARY KEY (~a-l `aPK`,~! `aPK`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8;

DELIMITER $$
DROP PROCEDURE IF EXISTS `upgrade_database` $$
CREATE PROCEDURE upgrade_database()
BEGIN
    DECLARE _count INT;
#* ~a addColumn
#1 ALTER TABLE ~! tl DROP PRIMARY KEY, ADD PRIMARY KEY (~a-l `aPK`,~! `aPK`);

END $$
DELIMITER ;

CALL upgrade_database();
```

Luča struktura **Glumac.java** je definisana na sledeći način:

```
public class Glumac extends Structure {

    @Override
    public void create () throws FileNotFoundException, IOException {
        String structureName = getStructureName();

        s.addProperty("dbname","videoklub");
        s.addProperty("tl",structureName);

        Structure at = new Structure("Sifra glumca");
        at.addProperty("at", "sifraGlumca"); at.addProperty("aPK", "sifraGlumca");
        at.addProperty("tpMySQL", "int"); at.addProperty("sizeSQL", "(11)"); at.addProperty("defaultMySQL", "");
        Template.addColumn(at);
        addStructure(at);

        at = new Structure("Ime i prezime");
        at.addProperty("at", "imePrezime"); at.addProperty("tpMySQL", "varchar"); at.addProperty("sizeSQL", "(30)");
        at.addProperty("defaultMySQL", "DEFAULT NULL");
        Template.addColumn(at);
        addStructure(at);

        at = new Structure("Broj filmova");
        at.addProperty("at", "brojFilmova"); at.addProperty("tpMySQL", "int"); at.addProperty("sizeSQL", "(11)");
        at.addProperty("defaultMySQL", "");
        Template.addColumn(at);
        addStructure(at);

        at = new Structure("Datum rođenja");
        at.addProperty("at", "datumRodjenja"); at.addProperty("tpMySQL", "date");
        at.addProperty("sizeSQL", ""); at.addProperty("defaultMySQL", "DEFAULT NULL");
        Template.addColumn(at);
        addStructure(at);

    }
}
```

Generisana komponenta **CounterGlumac.sql** ima sledeći izgled:

```
CREATE DATABASE /*!32312 IF NOT EXISTS*/ `videoklub` /*!40100 DEFAULT CHARACTER SET utf8 */;

USE `videoklub`;

CREATE TABLE IF NOT EXISTS `Glumac` (
  `sifraGlumca` int(11) NOT NULL ,
  `imePrezime` varchar(30) DEFAULT NULL,
  `brojFilmova` int(11) ,
  `datumRodjenja` date DEFAULT NULL,
  PRIMARY KEY (`sifraGlumca`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8;

DELIMITER $$
DROP PROCEDURE IF EXISTS `upgrade_database` $$
CREATE PROCEDURE upgrade_database()
BEGIN
  DECLARE _count INT;
  SET _count = ( SELECT COUNT(*) FROM INFORMATION_SCHEMA.COLUMNS
  WHERE TABLE_SCHEMA = 'videoklub' AND TABLE_NAME = 'Glumac' AND COLUMN_NAME = 'sifraGlumca');
  IF _count = 0 THEN ALTER TABLE Glumac ADD COLUMN `sifraGlumca` int(11) NOT NULL ; END IF;
  SET _count = ( SELECT COUNT(*) FROM INFORMATION_SCHEMA.COLUMNS
  WHERE TABLE_SCHEMA = 'videoklub' AND TABLE_NAME = 'Glumac' AND COLUMN_NAME = 'imePrezime');
  IF _count = 0 THEN ALTER TABLE Glumac ADD COLUMN `imePrezime` varchar(30) DEFAULT NULL; END IF;
  SET _count = ( SELECT COUNT(*) FROM INFORMATION_SCHEMA.COLUMNS
  WHERE TABLE_SCHEMA = 'videoklub' AND TABLE_NAME = 'Glumac' AND COLUMN_NAME = 'brojFilmova');
  IF _count = 0 THEN ALTER TABLE Glumac ADD COLUMN `brojFilmova` int(11) ; END IF;
  SET _count = ( SELECT COUNT(*) FROM INFORMATION_SCHEMA.COLUMNS
  WHERE TABLE_SCHEMA = 'videoklub' AND TABLE_NAME = 'Glumac' AND COLUMN_NAME = 'datumRodjenja');
  IF _count = 0 THEN ALTER TABLE Glumac ADD COLUMN `datumRodjenja` date DEFAULT NULL; END IF;
  ALTER TABLE Glumac DROP PRIMARY KEY, ADD PRIMARY KEY (`sifraGlumca`);

END $$
DELIMITER ;

CALL upgrade_database();
```

## 5. ZAKLJUČAK I PRAVCI DALJEG ISTRAŽIVANJA

Generator programa – Luča je programski alat koji generiše programske komponente na osnovu Luča šablona i Luča struktura. Luča šabloni sadrže generički programski kod i Luča izraze, koji se koriste za različite domene problema. Luča struktura definiše konkretan domen problema i ona sadrži specifične delove programskog koda. U procesu generisanja komponenti Luča šabloni, odnosno Luča izrazi se parametrizuju sa Luča strukturama. Proces generisanja komponente pomoću generatora programa Luča je detaljno objašnjen uz konkretan primer u kome se generiše sql skript datoteka (CounterGlumac.sql) na osnovu Luča šablona (MySQLCounterTableTemplate.lctm) i Luča strukture (Glumac.java). Generator programa Luče je razvijen u Java programskom jeziku. Trenutna verzija generatora programa Luča je 1.3. Tokom ove godine planiramo da razvijemo verziju 2.0 u kojoj će logika generatora programa i softverska arhitektura generatora programa biti redizajnirani, kako bi se omogućilo: a) krajnjim korisnicima olakšano korišćenje generatora i b) olakšan dalji razvoj generatora. Generator programa – Luča je korišćen u nastavi na predmetu **Softverski proces** u poslednje tri školske godine, kada su se objašnjavali generički mehanizmi u razvoju softverskih sistema.

Jedan od nedostataka rada sa Luča generatorom bilo je kreiranje strukture (sa svojstvima) u Java programskom jeziku. Iz ovoga smo izveli zaključak da bi Luča generator mogli da koriste samo korisnici koji imaju znanje Java programskog jezika što smo smatrali kao nepotrebno ograničenje. Da bismo uklonili ovaj nedostatak, odlučili smo se da razvijemo domensko specifičan jezik za kreiranje strukture koja predstavlja ulaz u Luča generator.

U našem radu [8] predstavili smo SilabBiz domensko specifični jezik za specifikaciju poslovne logike softverskog sistema. **SilabBiz DSL** omogućava kreiranje kompletne hijerarhijske strukture i generisanje Java datoteka koje predstavljaju ulaz u Luča generator. Kao što je u radu opisano jezik je kreiran u **xText** alatu na osnovu čega je kreiran **Eclipse editor** koji omogućava unos hijerarhijske strukture u editoru koji podržava standardni set funkcionalnosti (opciju kompletiranja koda (eng. *intellisense*), prikaz sintaksnih grešaka, naglašavanje bojama (eng. *color highlighting*), brzu ispravku (eng. *quickfixies*)). Alat dalje omogućava automatsku transformaciju unete hijerarhijske strukture u Java programski kod koji bi predstavljao ulaz u Luča generator.

U širem kontekstu Luča generator će dalje biti iskorišćen za generisanje poslovne logike softverskog sistema. Ovo će omogućiti da se uporede mogućnosti i performanse Luča generatora sa **Xpend** jezikom (podrazumevanim jezikom za kreiranje obrazaca za transformaciju (generisanje) programskog koda). U doktorskoj disertaciji mr. Vojislava Stanojevića, koja se trenutno radi, biće predstavljeni detaljniji rezultati upotrebe Luča generatora u generisanju poslovne logike softverskog sistema. Dalji pravci razvoja se odnose na proširenje **SilabBiz DSL** u pogledu mogućnosti kreiranja **Luča šablona** u editoru. Pored toga pokušaćemo da rezultate koje smo predstavili u radovima [6, 1] proširimo korišćenjem Luča generatora.

## LITERATURA

- [1] Antović, I., Vlajić, S., Milić, M., Savić D., & Stanojević V. (2012). Model and software tool for automatic generation of user interface based on use case and data model. *IET Software*, ISSN 1751-8806, 6(6), 559-573.
- [2] Franky, M. C., & Pavlich-Mariscal, J. A. (2012). Improving implementation of code generators: A regular-expression approach. In *2012 XXXVIII Conferencia Latinoamericana En Informatica (CLEI), IEEE*, 1-10.
- [3] Jörges, S. (Ed.) (2013). Construction and evolution of code generators: A model-driven and service-oriented approach, (Vol. 7747), Springer.
- [4] Kroiss, C., Koch, N., & Knapp, A. (2009). Uwe4jsf: A model-driven generation approach for web applications. In *International Conference on Web Engineering, Springer, Berlin, Heidelberg*, 493-496.
- [5] Lano, K., Xue, Q., & Kolahdouz-Rahimi, S. (2020). Agile specification of code generators for model-driven engineering. *Proceedings of ICSEA 2020*, 9-15.
- [6] Savić, D., Vlajić, S., Lazarević S., Stanojević, V., Antović, I., Milić, M., & Rodrigues da Silva, A. (2015). Use Case Specification Using the SILABREQ Domain Specific Language. *Computing and Informatics Journal*, ISSN 1335-915, 34(4), 877-910.
- [7] Sebastián, G., Gallud, J. A., & Tesoriero, R. (2020). Code generation using model driven architecture: A systematic mapping study. *Journal of Computer Languages*, 56, 100935.
- [8] Stanojević, V., Vlajić S., & Savić, D. (2022). A SilabBiz dsl for business logic specification. *Symorg 2022, Belgrade*, 196-198.
- [9] Tolvanen, J. P., & Rossi, M. (2003). Metaedit+ defining and using domain-specific modeling languages and code generators. In *Companion of the 18th annual ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications*, 92-93.



## POSLEDICE PRIMENE DIGITALNE KOLABORACIJE

### CONSEQUENCES OF THE DIGITAL COLLABORATION USE

RAJKO IVANIŠEVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Novom Sadu – Ekonomski fakultet u Subotici, rajko.ivanisevic@ef.uns.ac.rs

**Rezime:** Razvojem tehnologije dolazi do promene načina kolaboracije. Kolaboracija suštinski ne menja svoju namenu, već uz pomoć tehnologije doprinosi novim mogućnostima u oblasti njene primene i poboljšanja rezultata, tj. ispunjenja ciljeva zbog koji se i stupilo u kolaboraciju. Digitalna kolaboracije predstavlja kolaboraciju koja se odvija putem digitalnih platformi i drugih digitalnih kolaboracionih alata. Ona ima za cilj povećanje efektivnosti i efikasnosti saradnje što na poslovnom, što na svakodnevnom društvenom nivou. Svedoci smo naglog skoka u primeni ovih platformi i alata u prethodne dve godine. One imaju svoje vrline, ali isto tako i mane. Rad se bavi posledicama korišćenja digitalnih kolaborativnih alata. Izvršen je pregled literature na datu temu.

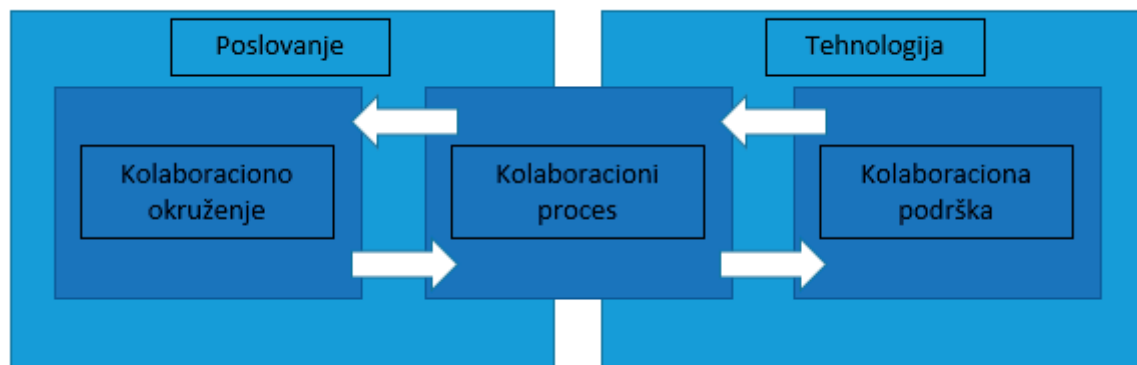
**Ključne reči:** Digitalna kolaboracija, Kolaboracija, Kolaboracioni alati

**Abstract:** With the development of technology, there is a change in the way of collaboration. Collaboration essentially does not change its purpose, but with the help of technology contributes to new possibilities in the field of its useage and improvement of results, ie. fulfillment of the goals for which the collaboration was started. Digital collaboration is collaboration that takes place through digital platforms and other digital collaboration tools. It aims to increase the effectiveness and efficiency of cooperation both at the business and at the everyday social level. We are witnessing a sharp jump in the application of these platforms and tools in the past two years. They have their virtues, but also their flaws. The paper deals with the consequences of using digital collaborative tools. A review of the literature on the given topic was performed.

**Keywords:** Digital collaboration, Collaboration, Collaboration tools

#### 1. UVOD

Kolaboracija se dešava u trenutku kada dve ili više osoba komunicira i ima saobraćanje da bi postigli cilj. Ovo je svakodnevica u današnjem poslovanju organizacija. Kako bi se omogućilo stvaranje novih vrednosti i pospešilo postizanje organizacionih ciljeva, ključno je razumeti i poboljšati način na koji ljudi saraduju (vrše kolaboraciju). E-kolaboracija se definiše kao kolaboracija kod koje izostaje interakcija lice u lice, gde se u zamenu koriste komunikacione i informacione tehnologije. Radni okvir (eng. *framework*) za e-kolaboraciju se sastoji od okruženja, procesa i podrške. Kolaboraciono okruženje se sastoji od prirode zadatka i organizacionih postavki poslovanja, trgovanja, aktera, organizacione strukture... (Weiseth et al., 2006)



**Slika 1:** Kolaboracioni radni okvir (framework)  
Slika kreirana na osnovu Weiseth i sar. (2006)

Da bi kolaboracija bila uspešna sva tri segmenta sa prikazane slike moraju biti upravljana na pravi način. Ipak, najveći naglasak je na usklađivanju poslovne i tehnološke perspektive. (Weiseth et al., 2006)

Andone i Frydenberg (2011) navode istorijku liniju razvoja kolaboracije. U početku su to bili blogovi, wiki-ji, a zatim su se sa razvojem tehnologije pojavili podkasti, veb bazirani (eng. *web-based*) audio i video alati za deljenje i uživo prenos, alati za omogućavanje kolaboracije (kolaborativnog kreiranja sadržaja).

Kada je reč o upotrebi veb alata danas se najčešće koriste Web 2.0 alati. Rad se bavi analiziranjem kolaboracije i softvera koji se primenjuju radi sprovođenja iste, sa naglaskom na onlajn kolaboraciju.

## 2. METODOLOGIJA I PROCES SPROVOĐENJA PREGLEDA LITERATURE

Jasno definisanje istraživačkog pitanja je ključno po Xiao i Watson (2019). Oni navode određene korako po kojima treba sprovesti istraživanje, a dati koraci su grupisani u kategorije planiranje i sprovođenje pregleda i izveštavanje o rezultatima.

### 2.1. Planiranje pregleda literature

Planiranje pregleda literature se sastoji iz dva dela. Prvi deo je formulisanja istraživačkog problema. Drugi deo je razvijanje i validacija procesa pregleda.

Prilikom definisanja istraživačkog problema, kada je reč o kolaboraciji, tražena je korelacija između kolaboracije i određenih veb baziranih (eng. *web-based*) platformi koje omogućavaju kolaboraciju. Tako je definisano istraživačko pitanje – posledice primene digitalne kolaboracije.

Radovi nad kojima je vršen pregled su u vremenskom intervalu od 2006. do 2021. godine.

### 2.2. Sprovođenje pregleda literature

Pet korako kroz koje je vršeno sprovođenje literature su:

- pretraga literature,
- pregled za uključivanje u rad,
- ocenjivanje kvaliteta,
- izdvajanje podataka i
- analiziranje i sintetizovanje podataka.

Pretraga literature vršena je preko elektronske baze podataka – „Google Scholar“. Za pronalazak radova korišćena je fraza „digital collaborative tools“. Pristupilo se pregledu prvih sedam stranica sa rezultatima. Nakon čitanja naslova i apstrakta izdvojeno je sedamnaest radova. Nakon pronalaženja celokupnih radova broj se smanjio na jedanaest. Jedan rad je isključen jer je bio knjiga, jedan nije bilo dovršeno istraživanje, te se broj radova nad kojima je vršen pregled smanjio na devet. Svi radovi su kvalitativno pogodovali uključivanju u pregled, a za izdvajanje podataka korišćena je tekstualna narativna sinteza i narativni pregled. Za svaki rad vođena je evidencija o informacijama značajnim za pregled literature. Kako bi se dobila celokupna slika na kraju je izvršeno sintetizovanje svih radova i izvlačenje suštine iz celokupne sadržine.

## 3. REZULTATI PREGLEDA LITERATURE – IZVEŠTAVANJE O NALAZIMA

Rad koji je najstariji (2006. godina), a nalazi se u odabranim radovima za pregled literature jeste rad autora Weiseth i sar. (2006). Oni prepoznaju termin WCT - *Wheel of Collaboration Tools*, ili točak kolaboracionih alata. Ovaj termin predstavlja tipologiju mogućnosti alata za kolaboraciju. WCT se sastoji od kolaboracionog interfejsa, kolaboracionih funkcija, menadžment sadržaja i integracija procesa. U kolaboracionom interfejsu postoje tri sloja, i to:

- radno mesto (fizičko radno mesto),
- uređaji i
- pogledi (portals).

Kolaboracione funkcije se sastoje od koordinacije, produkcije i donošenja odluka sa svim svojim podkategorijama. Ono što značajno za rad jesu uređaji i pogledi. Oni predstavljaju konkretne uređaje sa kojih se vrši kolaboracija, kao i način prikaza sadržaja (nije isti prikaz na računaru, telefonu, itd.). Tako je omogućeno korisnicima da sa različitih radnih mesta (čak i u toku puta, npr. voz) pristupe potrebnim informacijama ili doprinesu poslovnom procesu.

Wink (2009) u svom radu navodi termine VoIP (eng. *voice over internet protocol*) i web-based document development tools. VoIP omogućava transmisiju poruka, poziva, video poziva, glasovnih poruka. Veb bazirani alati za razvoj dokumenata dozvoljavaju kreiranje dokumenata koristeći reči, listove, prezentacije, itd. Zbog

pristupačnosti jer se dokumenta nalaze na internetu moguće je pristupiti im i menjati ih bilo kada sa bilo kog mesta. Rad zaključuje da verbalna komunikacija ima bolji rezultat od razmene poruka. Rad govori da je najpopularniji VoIP Skype (reč je o tadašnjem vremenskom periodu). Veb bazirani alati svoju primenu nalaze kao i danas. Primer veb baziranog kolaborativnog radnog prostora je googledocs, zoho, etherpad. Korišćenje veb baziranih kolaborativnih alata ima za cilj da unapredi izvršavanje zadataka, a ne da bi se tehnologija koristila samo radi tehnologije. Korisnici moraju znati da koriste veb bazirane kolaboracione alate. Saradnici moraju biti obavešteni o informacijama koje se tiču njihovih računara, zbog različitosti u kvalitetu hardvera pre svega (danas je to uglavnom nepotrebno, jer svi računari podržavaju osnovne alate) i moraju znati kako da barataju alatima. Rad naglašava bitnost sigurnosti, ne sme se dozvoliti neovlašćeni pristup, potrebno je obezbediti privatnost, itd.

Po Andone i Frydenberg (2011) postoji velika rasprostranjenost korišćenja veb baziranih alata kako za kompanije, tako i za individualnu komunikaciju i kolaboraciju. Posmatraju onlajn kolaboraciju kroz prizmu učenja. Ono zahteva od današnjih studenata poznavanje određenih digitalnih veština i umeće baratanja njima. Rad pruža empirijske podatke istraživanja koje je sprovedeno u Rumuniji od strane autora i pokazuje listu veb baziranih alata za komunikaciju i kolaboraciju koju najčešće studenti koriste i koju će najpre koristiti. Zanimljivo je da što bi pre koristili neki alat, to je i taj alat na lestvici važnosti bio na višoj poziciji. Najvažniji alat je bio pregledač (Google), a odmah zatim mobilni telefon. Najmanje važan alat za njih je bio Second Life (virtuelna realnost). Rezultati su pokazali da ispitani studenti najčešće koriste e-mail, potom IM (eng. *instant messaging*), a na trećem mestu našlo se deljenje dokumenata.

Collaborative networks (CNs) - jeste moderna naučna disciplina organizacija i pojedinaca koji omogućavaju zajedničku interakciju preko onlajn alata da bi se postigli zajednički ciljevi, nezavisno od geografske, kulturološke i ostalih razlika. Rad Michaelides i sar. (2013) posmatra ovaj pojam kroz prizmu malih i srednjih preduzeća. Posedovanje napredne tehnološke platforme sa modernom komunikacijom i kolaboracionim alatima nije samo po sebi dovoljno za poboljša učešće saradnika (kolaboratora). Web 2.0 alati omogućavaju nove mogućnosti u kolaborativnim mrežama u smislu da se kontrola, kreiranje i razvoj znanja se posmatra kao da teče iznutra ka spolja. Naglašava se značaj poverenja kao vitalne stavke u ohrabrivanju učesnika u virtuelnoj kolaboracionoj mreži. Kod malih i srednjih preduzeća ovde može doći do problema ukoliko zaposleni smatraju da su njihove veštine i znanja nedovoljno dobra da pruže relevantan doprinos organizaciji i poslovanju.

Polako dolaskom digitalizacije dolazi do produbljanja kolaboracije, tj. digitalne kolaboracije. Orellana, (2017) deli proces digitalizacije kolaboracije u dve faze:

- identifikovanje trenutne upotrebe i novih trendova i
- istraživanje kako digitalni alati oblikuju kolaboraciju.

Na stanovištu je da digitalizacija individualnih elemenata kolaboracionog procesa ne proizvodi dovolju efektivnost i efikasnost kada je reč o digitalizaciji kolaboracije. Uvodi se termin virtuelne realnosti, koji je danas sve prisutniji u svakodnevnom poslovanju. Novonastala virtuelna realnost, po njoj, će možda moći da pruži dovoljan opseg i dubinu podataka koji su potrebni za pravu digitalizaciju. Autor pominje Second Life i Active Worlds kao dva virtuelna okruženja koji mogu obuhvatiti neverovatnu količinu podataka. Najveća prepreka za usvajanje ovih načina kolaboracije mogu biti privatnost i sigurnost. Organizacije moraju da zaštite i učesnike i podatke. Takođe, prisutan je i ljudski otpor ovim tehnologijama (kulturološka promena). Rad navodi da smo spremni za digitalizaciju kolaboracije, samo da treba osmisliti kako to izvesti.

Hernández-Sellés i sar. (2019) daju značaj pojmu CSCL - Computer Supported Collaborative Learning (kompjuterski podržano kolaborativno učenje). Posmatrani domen jesu studenti, nastavnici i učenje. Rad pokazuje da alati za kolaboraciju imaju pozitivan i značajan uticaj, sa niskom i srednjom veličinom efekta, uz poštovanje interakcije učenika u radnim timovima i emocionalnom podrškom unutar grupe. U CSCL kontekstu da bi interakcija, u ovom slučaju studenta i nastavnika, bila efikasna, potrebna je kvalitetna povratna informacija, podrška, motivacija i dobar menadžmet. Onlajn kolaborativni alati, po ovom radu, su osnovni za omogućavanje interakcije između studenata. U prethodne dve godine bili smo svedoci sveopšte primene takvih alata poput Teams-a, Moodle-a, itd.

Pregled literature se dotiče i informatičkih uloga CIO i CDO. Borissova i sar. (2020) naglašavaju ulogu CIO-a u promovisanju novih razvojnih metoda, novih softvera, korišćenje novih IT uređaja. Prvo je potrebno da CIO da odgovor da li su nove tehnologija prihvatljive i potrebne za organizaciju. Ukoliko jesu, treba izvršiti procenu da li nadograđivati postojeću infrastrukturu ili uvoditi novu. Treba ustanoviti i kratkoročne i dugoročne posledice (promene). Ako saradnja na daljinu nije omogućena CIO treba da iznađe rešenje u kratkom roku. Moguće je uvesti i novi tim koji će se baviti rešavanjem kompleksnih i izazovnih problema digitalne transformacije. Za „udaljenu kolaboraciju“ govore da je novo normalno na radnom mestu. Potrebno

je evaluirati razne radne okvire (eng. *framework*) i njihov uticaj na ljude, posao i IT. Potrebno je iznaći odgovarajući softver koji će podržati kolaboraciju iz daljine. Naglašavaju se tri kategorije promatranja:

- video konferencije,
- e-učenje i
- alati za upravljanje projektima.

Nalaze svoju primenu kod poslovnih, univerzitetskih i istraživačkih organizacija. Softveri koji omogućavaju video konferencije se razlikuju po broju učesnika koje je moguće istovremeno opsluživati i vremenskom ograničenju poziva. Istraživanje je pokazalo da prednost u broju korisnika odnose platforme Zoom i Webex, dok je prednost u trajanju sastanka na strani platformi Skype i Google Hangouts. LMS – Learning Management System Software je značajan za polje edukacije i za polje treninga zaposlenih (biznis trening). Njima se želi postići povećanje motivacije, bolje razumevanje domena koji se obrađuje, povećanje efektivnosti, itd. Projektni menadžment je veoma značajan za tekuće aktivnosti na projektima, stoga je i softver koji PM koristi potrebno da bude fleksibilno rešenje koje će kombinovati različite alate, mogućnosti i funkcionalnosti. Za kolaboraciju je bitno praćenje potencijalnih problema, obaveštenja, dokumenta i mogućnost podele znanja. Za sam softver je bitno da je veb baziran kako bi omogućio saradnju u realnom vremenu.

Na posletku rad čiji su autori Marion i Fixson (2021) bavi se kolaboracijom u procesu razvoja novog proizvoda (eng. *NPD - New Product Development*). Rad navodi četiri IT kategorije za alate i softver i platforme:

- IT alati za komunikaciju (email, video konferencije...),
- IT alati za dizajniranje proizvoda (CAD...),
- IT alati za projekt menadžment i podatke o proizvodu i
- IT alati menadžmenta znanja.

Sve ovo pruža korist organizaciji, pre svega poboljšanje performansi i automatizacije procesa u budućnosti.

## 5. ZAKLJUČAK

Kolaboracija kao međusobna saradnja minimalno dve osobe je preduslov za bilo kakvo poslovanje. Da li kao kolaboracije među ljudima, organizacijama (van okvira organizacije), ili kao kolaboracija unutar organizacija. Napretkom tehnologije, primenom digitalizacije u poslednjim godinama, došlo je i do napretka kolaboracije. Uživno sastanci su sada unapređeni sa onlajn sastancima, što predstavlja prednost u vidu troškova, ali ujedno i nedostatak u vidu direktnog ljudskog kontakta. Saradnja unutar i van organizacije danas se obavlja kako uživo, tako i preko određenih platformi. Te platforme omogućavaju digitalnu kolaboraciju i predstavljaju alate digitalne kolaboracije (ili osnovu za implementaciju tih alata). Posledice korišćenja platformi i alata za digitalnu kolaboraciju jesu u ekonomskoj sferi pozitivne, ali nije tako jednostavna slika i po pitanju ljudi. Kako je u radu naglašavano glavni problemi su privatnost i sigurnost i pojedincu i organizaciji. Jeffrey, (2012) se u svom radu bavio problemom novog digitalnog mračnog doba. Njegovo osnovno pitanje je bilo kako izbeći drugo digitalno mračno doba? Po autoru migriranje naše komunikacije, saradnje i kreiranje sadržaja i upravljanje novim uslugama, je, u smislu dugoročne održivosti, posao koji je jednako krhak, verovatno krhkiji od rada na tradicionalnim IT infrastrukturama.

Kako je u usponu razvoj digitalnih platformi i alata, ova tema ima dosta prostora i prosperiteta po pitanju istraživanja. Posledice primene digitalne kolaboracije mogu se posmatrati u realnom vremenu što je još jedan od benefita teme rada.

## LITERATURA

- [1] Andone, D., & Frydenberg, M. (2011). Across Continents: Using Web Based Collaboration Tools for Learning. 2011 IEEE 11th International Conference on Advanced Learning Technologies, 100–102. <https://doi.org/10.1109/ICALT.2011.35>
- [2] Borissova, D., Dimitrova, Z., & Dimitrov, V. (2020). How to Support Teams to be Remote and Productive: Group Decision-Making for Distance Collaboration Software Tools. *Information & Security: An International Journal*, 46(1), 36–52. <https://doi.org/10.11610/isij.4603>
- [3] Hernández-Sellés, N., Pablo-César Muñoz-Carril, & González-Sanmamed, M. (2019). Computer-supported collaborative learning: An analysis of the relationship between interaction, emotional support and online collaborative tools. *Computers & Education*, 138, 1–12. <https://doi.org/10.1016/j.compedu.2019.04.012>

- 
- [4] Jeffrey, S. (2012). A new Digital Dark Age? Collaborative web tools, social media and long-term preservation. *World Archaeology*, 44(4), 553–570. <https://doi.org/10.1080/00438243.2012.737579>
- [5] Marion, T. J., & Fixson, S. K. (2021). The Transformation of the Innovation Process: How Digital Tools are Changing Work, Collaboration, and Organizations in New Product Development\*. *Journal of Product Innovation Management*, 38(1), 192–215. <https://doi.org/10.1111/jpim.12547>
- [6] Michaelides, R., Morton, S. C., Michaelides, Z., Lyons, A. C., & Liu, W. (2013). Collaboration networks and collaboration tools: a match for SMEs? *International Journal of Production Research*, 51(7), 2034–2048. <https://doi.org/10.1080/00207543.2012.701778>
- [7] Orellana, S. (2017). Digitalizing Collaboration. *Research-Technology Management*, 60(5), 12–14. <https://doi.org/10.1080/08956308.2017.1348125>
- [8] Weiseth, P. E., Munkvold, B. E., Tvedte, B., & Larsen, S. (2006). The wheel of collaboration tools. *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work - CSCW '06*, 239. <https://doi.org/10.1145/1180875.1180913>
- [9] Wink, D. M. (2009). Web-Based Collaboration Tools. *Nurse Educator*, 34(6), 235–237. <https://doi.org/10.1097/NNE.0b013e3181bc7348>
- [10] Xiao, Y., & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, 39(1), 93–112. <https://doi.org/10.1177/0739456X17723971>



## PRIMENA BENCHMARKDOTNET BIBLIOTEKE U MERENJU PERFORMANSI ALATA ZA OBJEKTNO-RELACIONO PRESLIKAVANJE .NET OKVIRA

### APPLICATION OF BENCHMARKDOTNET LIBRARY IN MEASURING PERFORMANCE TOOLS FOR OBJECT-RELATIONAL MAPPING IN .NET FRAMEWORK

TIJANA MILOŠEVIĆ<sup>1</sup>, SAŠA D. LAZAREVIĆ<sup>1</sup>, TATJANA STOJANOVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, milosevic.tijana111@gmail.com, {sasa.lazarevic, tatjana.stojanovic}@fon.bg.ac.rs

**Rezime:** Problem kojim se bavi ovaj rad je izbor dobrog ORM alata u pogledu performansi, dok je cilj rada izračunavanje i analiza performansi alata za objektno-relaciono preslikavanje u .NET okruženju. Metrike koje su uočene kao dobri pokazatelji performansi ORM alata su vreme i alocirana memorija. U radu su izmerene performanse Entity Core okvira, Entity Framework i NHibernate okvira i u tu svrhu je korišćena BenchmarkDotNet biblioteka. Ispitane su performanse osnovnih, CRUD, operacija nad bazom podataka - učitavanje, upisivanje, izmena i brisanje rekorda iz baze podataka. Prikazani su dobijeni rezultati i data je njihova analiza.

**Ključne reči:** BenchmarkDotNet, objektno-relaciono preslikavanje, performanse, metrike, ORM alat

**Abstract:** The problem of this paper is the choice of a good ORM tool in terms of performance, while the aim of the paper is to calculate and analyze the performance of tools for object-relational mapping in .NET environment. Metrics that have been observed as good indicators of ORM tool performance are time and allocated memory. The paper measured the performance of the Entity Core framework, the Entity Framework and the NHibernate framework, and used the BenchmarkDotNet library for this purpose. The performance of basic, CRUD, database operations - loading, writing, modifying and deleting records from the database were examined. The obtained results are presented and their analysis is given.

**Keywords:** BenchmarkDotNet, object-relational mapping, performance, metrics, ORM tool

## 1. UVOD

Objektno-relaciono preslikavanje (ORM) je tehnika programiranja za preslikavanje podataka između nekompatibilnih sistema u objektno orijentisanim programskim jezicima. Drugim rečima, to je koncept preslikavanja poslovnih objekata aplikacije u relacione tabele baze podataka, tako da se podacima može lako pristupiti i u potpunosti ih ažurirati kroz objektni model aplikacije.[8]

Kada se govori o objektno-relacionom preslikavanju, treba spomenuti i pojam perzistentnosti (postojanosti) objekta. Za neki proces ili objekat se kaže da je perzistentan ako nastavi da postoji i onda kada je njegov roditelj/glavni proces prestao da postoji.[9] Relacioni sistemi baze podataka čuvaju podatke u vidu rekorda i tabela, gde se može reći da su podaci perzistentni, tj. postojani. To bi značilo da ukoliko nas zanimaju informacije o nekom objektu, morali bismo svaki put da pozivamo bazu podataka. Stalno pozivanje baze podataka je skupa operacija, koja oduzima dosta vremena i memorijskog prostora. [7] Zbog toga se javila potreba za mehanizmom koji će jednom učitani rekord iz baze u vidu objekta čuvati i pratiti kroz njegova stanja. Upravo su ORM alati ti koji ovo omogućavaju.

Međutim, između objektnog i relacionog modela postoji neusklađenost. Ona se ogleda u nekoliko problema prilikom učitavanja/upisivanja podataka iz baze podataka i njihovog predstavljanja u objektnom/relacionom modelu: nasleđivanje, identitet, veze između entiteta, kretanje kroz podatke.

ORM alat bi trebalo da usaglasi objektni i relacioni model i obezbedi preslikavanje iz jednog modela u drugi. Na svetu postoji na stotine ORM alata, a neki od najpoznatijih su Hibernate (za Javu), Entity Framework, Entity Framework Core i NHibernate(za C#), Sequelize (za TypeScript i Node.js).

### 3. MERENJE PERFORMANSI ORM ALATA

Jedan od aspekata dobrog ORM alata su njegove performanse. Performanse predstavljaju meru koliko dobro osoba, mašina, sistem, itd. dobro obavljaju neku aktivnost, operaciju. To može biti vreme za koje se obavlja ta operacija, koliko resursa (novca, ljudstva, memorije...) je potrošeno, snaga motora za vozila itd. S obzirom da je alat za objektno-relaciono preslikavanje direktno vezan za bazu podataka, jasno je da je glavni pokazatelj performansi vreme upisivanja i učitavanja podataka. Međutim on nije i jedini pokazatelj. Memorija koja je potrebna za učitavanje/upisivanje podataka u bazu takođe je bitan pokazatelj performansi nekog ORM alata. [7]

Metrike koje su odabrane u ovom radu, kao pokazatelji performansi ORM alata su vreme izvršenja operacije i alocirana memorija prilikom izvršenja operacije. Jasno je da je vreme odgovora kako neke aplikacije tako i kod ORM alata bitno iz ugla korisnika i da se očekuje odgovor što pre. Iz tog razloga je vreme uzeto u razmatranje kada je reč o merenju performansi ORM alata. Pored vremena izvršenja kao metrika je odabrana i alocirana memorija jer se može desiti da alat bude brz, ali da za izvršenje svojih operacija troši više memorije. Zbog toga se u ovom radu performanse ORM alata sagledaju i sa aspekta memorije. [7]

### 4. BENCHMARKDOTNET BIBLIOTEKA

BenchmarkDotNet je biblioteka koja pomaže da se metode transformišu u repere (*engl. benchmarks*) i prate njihove performanse. Reper (benčmark) se može definisati na više načina, s obzirom da ima široku primenu u različitim oblastima. U oblasti softverskog inženjerstva benčmark se definiše kao program koji meri performanse drugog programa ili parčeta koda. On bi trebalo da obezbedi rezultate koji omogućavaju pristup novim informacijama o programu, izvršnom okruženju, operativnom sistemu. [1] BenchmarkDotNet biblioteka napisana je u C# programskom jeziku i njen izvorni kod se može naći na GitHub-u.

BenchmarkDotNet pruža pouzdane i precizne rezultate zahvaljujući *Perfolizer* statističkom mehanizmu.[5] *Perfolizer* je biblioteka koja sadrži kolekciju korisnih algoritama za analizu performansi. BenchmarkDotNet podržava C#, F# i Visual Basic programske jezike i izvršiv je na Windows, Linux i macOS operativnim sistemima.

BenchmarkDotNet daje slobodu korisniku da odabere način pokretanja procesa (*Through, ColdStart, Monitoring*), broj iteracija kroz koje se vrši merenje, statističke vrednosti, koje će se nalaziti u rezultatima merenja, poput srednje vrednosti, maksimuma, minimuma, standardne devijacije, medijane i drugih. Pored toga, ova biblioteka daje mogućnost korišćenja *EtwProfiler-a*, *EventPipeProfiler-a* i disemblera. Glavni aspekti BenchmarkDotNet biblioteke su jednostavnost, automatizacija, pouzdanost i prilagođenost korisniku.[5]

Rezultati su prikazani jasno, nedvosmisleno i u različitim formatima (csv, log, html, md) što pruža korisniku bolji pregled i olakšanu analizu rezultata.

BenchmarkDotNet vrši merenja u nekoliko koraka:[5]

- *BenchmarkRunner* generiše zasebni projekat po runtime-u, i gradi ga u *Release* režimu.
- Za svaku kombinaciju metoda i parametara pokušava da izračuna performanse pokretanjem benčmark procesa nekoliko puta (*LaunchCount*).
- Gomila operacija je iteracija. BenchmarkDotNet razlikuje nekoliko vrsta iteracija: pilot, biće odabran najbolji broj iteracija, *OverheadWarmup*, *ActualWarmup*, *ActualWorkLoad* koje predstavlja stvarna merenja, rezultat je razlika *ActualWorkLoad* i medijane.
- Nakon svih merenja kreira se *Summary* klasa koja sadrži sve informacije o pokretanjima benčmarka, datoteke koje sadrže rezime koji je čitljiv korisnicima i grafikone.

### 5. PRIMENA BENCHMARKDOTNET BIBLIOTEKE

Za potrebe rada kreirana je konzolna aplikacija u Visual Studio razvojnom okruženju. Korišćen je .NET 5 okvir. ORM alati čije su performanse merene su *Entity Framework Core*, *Entuty Framework* i *NHibernate*. Kao sistem za upravljanje bazama podataka korišćen je SqlServer Management Studio.

Performanse su merene na računaru WIndows 10 operativnog sistema, karakteristika Intel Core i7-6500 CPU 2.50Ghz, 16GB RAM.

#### 5.1. Konfiguracija

U ovom radu korišćena je BenchmarkDotNet biblioteka verzije 0.13.1. Metode čije se performanse mere označene su atributom *Benchmark*, i prilikom pokretanja programa one će biti uzete u obzir za analizu. Klasa u kojoj se nalaze *Benchmark* metode označena je kao *Job* preko atributa *SimpleJob*. Kao strategija za



pokretanje merenja odabrana je strategija ‘hladni start’, gde u vreme i memoriju, koju alokira metoda koja se meri, ne ulazi vreme i memorija koja je potrebna za pokretanje i pripremu metode. Definisan je i broj iteracija u kojima će se meriti metoda, u ovom radu je to deset iteracija. Odabrane su kolone koje će se pojaviti u rezultatu merenja: najmanje vreme koje je potrebno za izvršenje operacije, maksimalno vreme, srednja vrednost i alocirana memorija. Prikaz rezultata je sortiran tako da se prvo prikazuju najbrže metode, a redosled izvršenja metoda je po redosledu u kom su napisane u klasi.

S obzirom da je aplikacija za potrebe ovog rada konzolna, u okviru Program klase se navodi `BenchmarkRunner.Run<Klasa>()` čime se pokreće merenje određene klase. Dobra praksa je da se benčmarking pokreće u *Release* modu, što je i primenjeno u ovom radu.

```
[SimpleJob(RunStrategy.ColdStart, targetCount: 10)]
[MinColumn, MaxColumn, MeanColumn, MemoryDiagnoser]
[Orderer(SummaryOrderPolicy.FastestToSlowest, MethodOrderPolicy.Declared)]
public class ORMToolPerformance
{
    [Benchmark]
    public void Method(){
}
}

class Program
{
    static void Main(string[] args)
    {
        BenchmarkRunner.Run<ORMToolPerformance>();
    }
}
```

Slika 1: Konfiguracija BenchmarkDotNet biblioteke

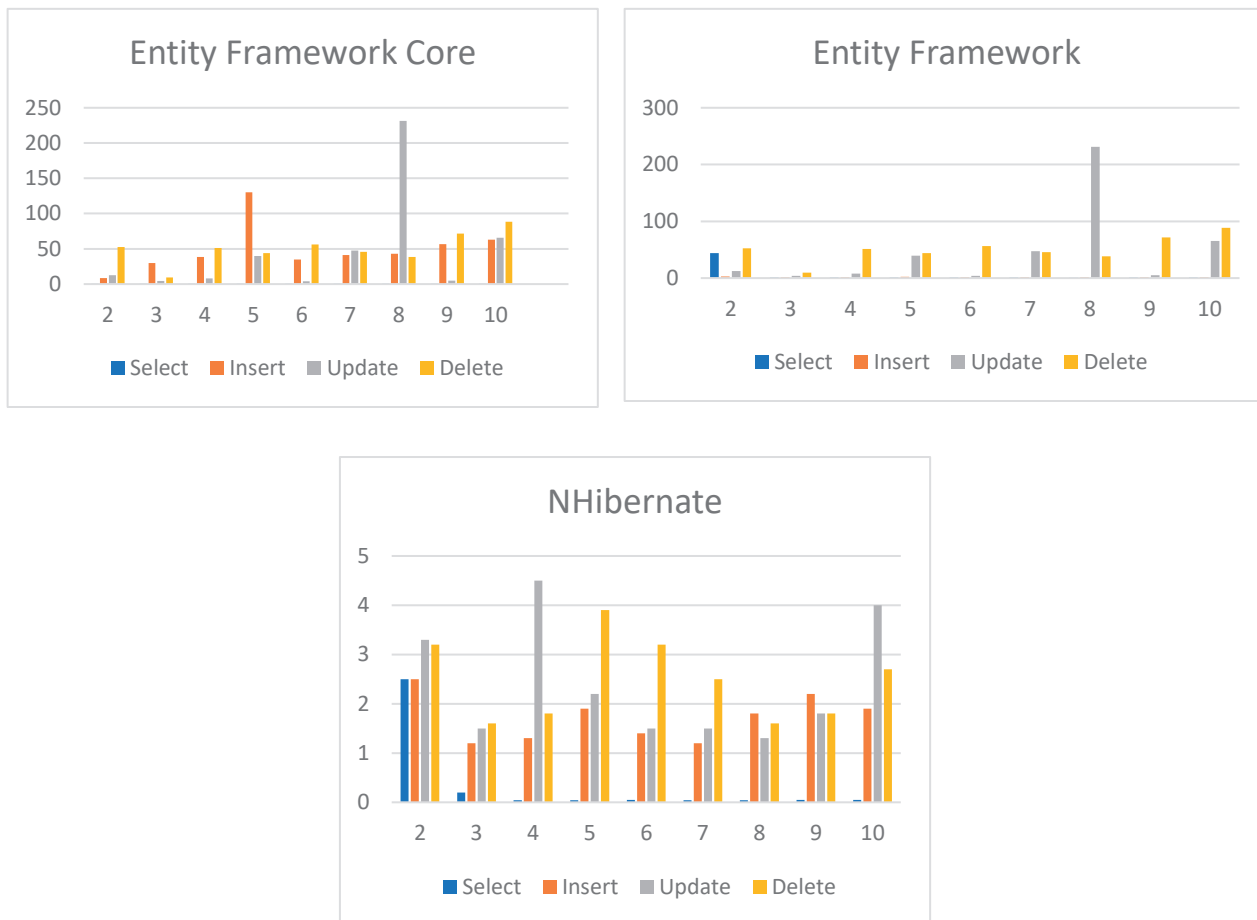
## 5.2. Merenje performansi

Primenom BenchmarkDotNet biblioteke, merene su performanse osnovnih, CRUD, operacija nad tabelom *City* za *Entity Framework Core*, *Entity Framework* i *NHibernate* okvire. Tabela *City* sadrži 26569 rekorda. Od kolona ima *Id*, koji nije automatski generisan, i *Naziv*. *Select* upit je rađen po *Id*-ju, pa je očekivan broj rekorda koji upit treba da vrati jedan. Prilikom pozivanja metode za dodavanje novog rekorda, kao parametri prosleđeni su *Id* i *Naziv*, prilikom izmene rekorda izmenjen je *Naziv* određenog rekorda, dok se iz baze briše rekord sa prosleđenim *Id*-jem.

U tabeli 1 dati su rezultati merenja čitanja, kreiranja, izmene i brisanja jednog rekorda iz table *City* za ove okvire. Merenja su rađena u 10 iteracija, iz razloga što je primećena oscilacija u vremenima izvršavanja metode. U prvoj iteraciji, izvršavanje metode je uvek najsporije, dok je u svakoj sledećoj iteraciji brže izvršavanje. Razlog tome je uspostavljanje konekcije sa bazom podataka. Međutim, pad u vremenu izvršenja metode nije linearan, što se može videti na slici 2.

Tabela 1: Rezultati merenja u 10 iteracija

Iteracija	Select			Insert			Update			Delete		
	EC	EF	NH	EC	EF	NH	EC	EF	NH	EC	EF	NH
1	5489.2	3140.3	155.3	4045.9	2589.5	129.6	4034.4	2689.1	140.1	3957.8	2696	134.6
2	0.5	43.8	2.5	8.3	2.6	2.5	12.6	3.1	3.3	52.5	4.5	3.2
3	0.2	1.1	0.2	29.7	1.6	1.2	4.2	2.8	1.5	9.6	2.8	1.6
4	0.03	1.1	0.04	38.5	1.6	1.3	8.1	3.4	4.5	51.1	1.6	1.8
5	0.04	0.9	0.04	130.04	2.5	1.9	39.6	1.9	2.2	43.8	3.6	3.9
6	0.03	0.9	0.05	35.02	1.6	1.4	4.1	1.9	1.5	56.3	1.9	3.2
7	0.03	1.1	0.04	41.1	1.9	1.2	47.6	2.5	1.5	45.8	2.4	2.5
8	0.04	0.8	0.04	43.2	1.8	1.8	231.3	2.2	1.3	38.6	1.7	1.6
9	0.04	1.3	0.05	56.5	1.4	2.2	4.9	2.6	1.8	71.7	2.1	1.8
10	0.04	1.2	0.05	62.9	1.5	1.9	65.5	3.2	4.0	88.4	1.6	2.7
Prosečno vreme(ms)	549	319.3	15.8	449.1	260.6	14.5	445.2	271.3	16.2	441.6	271.8	15.7
Alocirana memorija (KB)	0.6	58	0.5	70.8	67	19.3	48.5	145	25.7	68.3	61	24.1

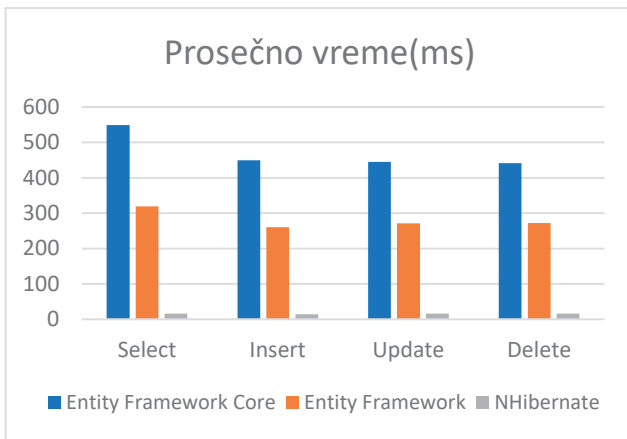


**Slika 2:** Vremena izvršavanja operacija po iteracijama za Entity Core, Entity i NHibernate okvire

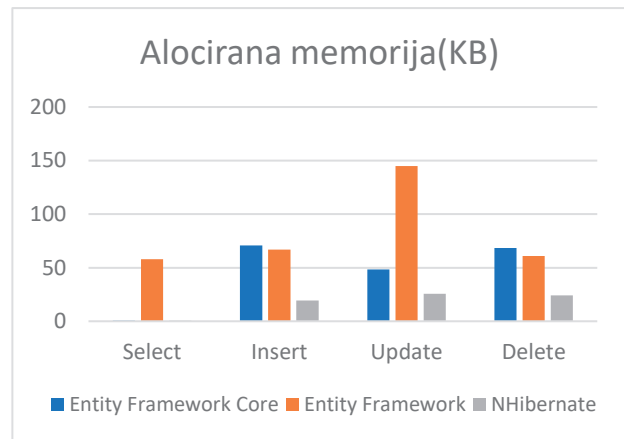
### 5.3. Analiza izmerenih rezultata

Kada su u pitanju performanse ORM alata, kao što je već rečeno, cilj je bio da se ispitaju brzina izvršavanja operacije i alocirana memorija. BenchmarkDotNet daje rezultate koji su lako obradivi i iz kojih se mogu izvesti određeni zaključci, po pitanju ovih parametara. Na slici 3 prikazan je grafikon sa prosečnim vremenima izvršavanja svake operacije za Entity Core, Entity i NHibernate okvire, za deset iteracija. Sa grafika se vidi da je, po pitanju vremena izvršenja operacije, bolji NHibernate okvir, u odnosu na ostale okvire. NHibernate okviru je najmanje vremena potrebno da izvrši operaciju dodavanja, a najviše operaciju izmene rekorda iz baze podataka. Entity Core okvir najbrže izvrši operaciju brisanja, a najsporije operaciju učitavanja rekorda iz baze podataka. Entity Framework najbrže izvrši operaciju dodavanja, a najsporije operaciju učitavanja rekorda iz baze podataka.

Po pitanju alociranje memorije, bolje rezultate ima NHibernate okvir, koji alocira manje memorije za izvršenje operacija, u odnosu na ostale okvire. Ovaj okvir je alocirao najmanje memorije tokom izvršavanja operacije učitavanja, a najviše prilikom izmene rekorda. Entity Core okvir je alocirao najmanje memorije prilikom izvršavanja operacije učitavanja, kao i Entity Framework, a najviše tokom dodavanja novog rekorda, dok je Entity Framework alocirao najviše memorije prilikom izmene rekorda.



Slika 3: Prosečno vreme



Slika 4: Alocirana memorija

U tabeli 2 dato je procentualno poboljšanje NHibernate okvira u odnosu na druga dva okvira, za sve operacije. Iz rezultata se vidi da bi NHibernate okvir doneo poboljšanje u svim slučajevima. Ovo bi značilo da ukoliko umesto Entity Core okvira odaberemo NHibernate, za operaciju učitavanja rekorda, NHibernate će utrošiti 97.1% manje vremena, a 16.7% manje memorije.

**Tabela 2:** Procentualno poboljšanje NHibernate okvira za sve operacije u odnosu na Entity Core okvir i Entity Framework

Operacija/procentualno poboljšanje	Entity Framework Core		Entity Framework	
	Vreme	Memorija	Vreme	Memorija
Select	97,1%	16,7%	95,1%	99,1%
Insert	96,8%	72,7%	94,4%	71,2%
Update	96,4%	47,0%	94,0%	82,3%
Delete	96,4%	64,7%	94,2%	60,5%

## 6. ZAKLJUČAK

Kako danas postoji dosta ORM alata postavlja se pitanje koji alat upotrebiti. U ovom radu je predstavljena analiza ORM alata upotrebom BenchmarkDotNet biblioteke. Analizom performansi ovih alata, može se utvrditi koji od alata pruža najbolje performanse za dati primer. Ova biblioteka pruža da se metode koje se ispituju označe kao benchmark i na taj način rade statističke analize. U ovom radu su kao pokazatelji performansi ispitivani vreme izvršenja operacije i alocirana memorija. BenchmarkDotNet pruža podršku za merenje ovih pokazatelja. Odabirom MeanColumn atributa dobićemo prosečno vreme izvršenja metode, dok za količinu alocirane memorije tokom izvršenja metode biramo atribut MemoryDiagnoser. Ova biblioteka može biti vrlo korisna u ispitivanju performansi samog koda aplikacije i uočavanju 'uskog grla' aplikacije.

Rezultati ovog istraživanja pokazali su da NHibernate ima najbolje performanse (najmanje vreme izvršenja operacije i najmanje alocirana memorija), u poređenju sa Entity Framework i Entity Core okvirom, za sve tipove operacija nad tabelom baze podataka.

## LITERATURA

- [1] Akinshin, A. (2019). *Pro .NET Benchmarking*. Apress.
- [2] BenchmarkDotNet: <https://benchmarkdotnet.org/articles/overview.html> (posećeno 11.05.2022.)
- [3] Boake, A., Zyl, P. V., & Kourie, D. G. (2015). *Comparing the Performance of Object Databases and ORM Tools*.
- [4] Cvetković, S., & Janković, S. D. (2019). *A Comparative Study of the Features and Performance of ORM Tools in a .NET Environment*. University of Nis.
- [5] Github BenchmarkDotNet. <https://github.com/dotnet/BenchmarkDotNet> (posećeno 11.05.2022.)
- [6] Microsoft: <https://docs.microsoft.com/en-us/ef/core/performance/performance-diagnosis?tabs=simple-logging%2Cload-entities> (posećeno 11.05.2022.)
- [7] Milošević, T., & Lazarević, S. D. (2021). *Merenje i analiza performansi Entity Core okvira i NHibernate okvira*.

- [8] Tutorials Point: [https://www.tutorialspoint.com/nhibernate/nhibernate\\_orm.htm](https://www.tutorialspoint.com/nhibernate/nhibernate_orm.htm) (posećeno 10.05.2022.)
- [9] Vignesh S. Study: <https://study.com/academy/lesson/object-persistence-definition-overview.html> (posećeno 10.05.2022.)
- [10] Zmaranda, D., Györödi, C., Gyorodi, R., & Lucian-Laurentiu Pop-Fele. (2020). *Performance Comparison of CRUD Methods using NET Object Relational Mappers: A Case Study*. University of Oradea.

## UTICAJ INTERNET PLATFORMI ZA KOMUNIKACIJU NA ISHODE UČENJA I OBRAZOVANJA

### THE INFLUENCE OF INTERNET PLATFORMS FOR COMMUNICATION ON LEARNING AND EDUCATION OUTCOMES

ALEKSANDRA ZEČEVIĆ<sup>1</sup>, ĐORĐE STAKIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, [aleksandra.zecevic@ekof.bg.ac.rs](mailto:aleksandra.zecevic@ekof.bg.ac.rs)

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, [djordje.stacic@ekof.bg.ac.rs](mailto:djordje.stacic@ekof.bg.ac.rs)

---

**Rezime:** U radu su predstavljene internet platforme koje se koriste kao alati u procesu nastave i učenja. Grupisane su na platforme za video komunikaciju, platforme za elektronsko učenje i platforme za trenutnu razmenu informacija. Iz svake grupe predstavljeni su najviše korišćeni predstavnici platformi. Kroz anketu ispitano je mišljenje studenata o korišćenju ovih platformi u nastavi i učenju. Predstavljeni su dobijeni rezultati.

**Ključne reči:** internet platforme za učenje, Moodle, Zoom, Viber, WhatsApp, Facebook

**Abstract:** The paper presents internet platforms that are used as tools in the process of teaching and learning. They are grouped into video communication platforms, e-learning platforms and instant information exchange platforms. The most used platform representatives from each group were presented. The survey examined students opinions on the use of these platforms in teaching and learning. The obtained results are presented.

**Keywords:** internet learning platforms, Moodle, Zoom, Viber, WhatsApp, Facebook

## 1. UVOD

Tokom pandemije virusa korona 2020-2022 obrazovni procesi su bili u izazovnoj situaciji koja je zahtevala povećanu primenu internet platformi u komunikaciji između nastavnika i studenata, kao i studenata međusobno. Nemogućnost fizičkog održavanja nastave u učionici zahtevala je prilagođavanje svih učesnika nastavnog procesa.

Promene nastale tokom pandemije uzrokovale su situaciju gde su se nastavnici suočili sa izazovima u pripremi i pružanju kvalitetnog predavanja na daljinu. Nastavnici su bili primorani da razmotre nekoliko aspekata tokom ovih dešavanja: potrebu za korišćenjem Internet platformi, način prilagođavanja korišćenju platformi koje odgovaraju potrebama procesu učenja, prilagođavanje učenju u realnom vremenu, prilagođavanje sistemu upravljanja učenju [1].

U ovom radu predstavljaju se platforme koje su najčešće korišćene u te svrhe na Ekonomskom fakultetu Univerziteta u Beogradu. Prikazane platforme su podeljene u tri grupe: za video komunikaciju, elektronsko učenje i trenutnu razmenu informacija. Najveća pažnja posvećena je sledećim platformama: Zoom, Moodle, Viber, WhatsApp, Facebook, veb sajtu predmeta.

Sprovedena je anketa kojom su ispitani stavovi studenata i njihovo iskustvo u korišćenju ovih platformi. Anketa ima više pitanja i na nju je odgovor dalo 84 studenta. Nakon prikaza korišćenih platformi prikazani su rezultati ankete i zaključci koji su na osnovu toga dobijeni.

## 2. INTERNET PLATFORME ZA KOMUNIKACIJU

Internet platforme za komunikaciju postoje i razvijaju se dugi niz godina ali njihova upotreba u nastavi nije bila previše zastupljena pre pandemije, nego je bila ostavljena na volju nastavnicima i studentima da ih koriste po izboru i po potrebi. Ovakvih platformi ima mnogo, međutim u nastavku su probrane one koje su najčešće korišćene tokom pandemije, kada su neke od njih doživele pravu ekspanziju posebno kada su stavljene u tešku ulogu da zamene nastavni proces uživo u klasičnoj učionici.

Neke od ovih platformi su korišćene i pre pandemije u meri u kojoj su bile potrebne kao adekvatna dopuna nastavnim procesima tokom održavanja nastave uživo. Tu se pre svega misli na sajt predmeta kao platformu za učenje i sve pomenute platforme za trenutnu razmenu informacija (Viber, WhatsApp, Facebook), koje su dodatno dobile na značaju tokom pandemije. Tokom pandemije su najveću ekspanziju doživele platforme Moodle i Zoom. Moodle je u prvim mesecima pandemije bio dominantna platforma za elektronsko učenje naročito za predmete koji nisu imali do tada svoj sajt i morali su u veoma kratkom roku da se transformišu. Već od jeseni 2020. godine Zoom kao platforma za video komunikaciju postaje pretežni način držanja nastave na daljinu, stavljajući platforme za elektronsko učenje u drugi plan koje postaju repozitorijumi za skladištenje nastavnog sadržaja. Osim toga, platforme kao što su Moodle korišćene su za testiranje znanja studenata, kao i za njihovo samovrednovanje rešavanjem testova i kvizova.

Osim pomenutih platformi ponegde je korišćen i Skype, zatim imejl, međutim udeo njihove upotrebe nije bio tako veliki niti značajno promenjen tokom pandemije usled postojanja većeg izbora drugih platformi.

## 2.1. PLATFORME ZA VIDEO KOMUNIKACIJU

Platforme za video komunikaciju su platforme koje su se koristile i pre pandemije, međutim situacija koja je nastala pojavom pandemije omogućila je ekspanziju korišćenja ovih vrsta platformi. Ova vrsta platformi uopšteno omogućava fleksibilan rad, prisustvovanje času ili sastanku (meeting) sa bilo kog mesta, vizuelnu saradnju među učesnicima, razmenu ideja, kao i održavanje bilo kojih virtuelnih događaja.

Platforme za video komunikaciju ne mogu u potpunosti da zamene direktnu komunikaciju, ali u mnogim situacijama mogućnost da se vide i čuju učesnici sastanka na daljinu poboljšava komunikaciju. Tehnologija ovih platformi zapravo obuhvata niz različitih tehnologija koje se koriste u širokom spektru situacija, gde komunikacija često nije samo video i audio interakcija, već se prenose i podaci, omogućavajući kolaborativni rad kroz zajedničke aplikacije. Platforme mogu da obezbeđuju:

- sastanke jedan-prema-jedan, koji obično uključuju potpunu dvosmernu audio i video interakciju.
- sastanke jedan-prema-više koji uključuju potpunu audio i video interakciju pokrenutu sa host lokacije, na koju ostale lokacije mogu da šalju zvuk ili sliku. Na primer, u situaciji predavanja, studenti mogu postavljati pitanja ili diskutovati, uz mogućnost uključivanja i video interakcije.
- sastanke više-prema-više, koji predstavljaju komunikaciju sa više lokacija, gde se obezbeđuje audio i video interakcija između više od dve lokacije.

Poslednjih godina na tržištu su se pojavili različiti tipovi softvera za video komunikaciju. Neki od najčešće korišćenih softvera za ovu vrstu komunikacije su Zoom, MS Teams i Google Meet.

### Zoom

Zoom platforma je jedna od najčešće korišćenih platformi ovog tipa u svetu i pruža veliki broj funkcionalnosti za interakciju između učesnika. Takođe, platforma nudi veliki broj alata u vidu razgovora, pisanja poruka, interaktivnog pisanja po "tabli", deljenja ekrana, omogućavanje više "soba i prostora" za rad, snimanja sastanaka, razmenu fajlova, itd. Zoom platforma nudi i svoja posebna rešenja za primenu u sferi edukacije, finansija i zdravstva.

### MS Teams

MS Teams je platforma za video komunikaciju, razvijena od strane Microsoft-a. Ova platforma, takođe pruža veliki broj funkcionalnosti, primarno radno okruženje za razgovor, video konferenciju i integraciju više aplikacija. MS Teams nudi različite proizvode za sastanke i konferencije, video pozive, razgovor i kolaboraciju, razmenu fajlova. Takođe, obezbeđuje korišćenje veb verzije programa kao što su Word, Excel i PowerPoint, kao i mogućnost skladištenja i deljenja fajlova.

### Google Meet

Google Meet je platforma koja je integrisani deo Google G-Suite softvera i koja omogućava video-konferencije, gde se na primer u edukaciji i poslovanju mogu koristiti različiti alati poput okruženja za interakciju, postavljanja anketa i pitanja, ali se mogu uključiti i integrisani paketi Google aplikacija za bolju saradnju i produktivnost, kao i alati koji su projektovani za pomoć u edukaciji (nastavnike i učenike).

## 2.2. PLATFORME ZA ELEKTRONSKO UČENJE

Elektronsko učenje je pojam novijeg datuma i predstavlja proces prenošenja znanja i vještina elektronskim putem. E-učenje je prvi put definisano 1999. godine kao “upotreba mrežne tehnologije za dizajniranje, isporuku, odabir, administriranje i proširenje učenja”[2]. Godinu dana kasnije je predstavljena i prva verzija SCORM (Shareable Content Object Reference Model) standarda, koji se odnosi na međunarodne standarde definisane za elektronske kurseve.

Da bi elektronsko učenje uopšte bilo moguće, potrebno je da nastavni materijal postane dostupan putem Interneta ili intraneta. To je jednostavno postići izradom veb sajta na kome bi bili dostupni svi sadržaji koji su potrebni za nastavu i kome svaki korisnik može da pristupi i preuzme sadržaj. Za razliku od veb sajtova koji predstavljaju jednostavne sisteme za pomoć učenju, postoje složeni sistemi elektronskog učenja koji se odnose na platforme za elektronsko učenje.

Platforme za elektronsko učenje ili učenje na daljinu, nazivaju se još i LSM (Learning Management System) platforme. Ova vrsta platforme predstavlja softversku aplikaciju koja se koristi za kreiranje okruženja za obuku ili nastavni proces, administraciju, dokumentovanje i praćenje procesa.

Jedna od najčešće korišćenih LSM platformi jeste Moodle platforma. Moodle platforma je edukaciona tehnologija otvorenog koda. Moodle platforma omogućava prostor za postavljanje nastavnog materijala, kome se pristupa dodeljivanjem prava korisnika za pristup, postavljanje informacija, testova za proveru znanja, itd. Za razliku od preuzimanja nastavnog materijala sa veb sajta, ova vrsta platformi obezbeđuje administraciju materijala, ali i upisivanje korisnika za pristup materijalu, kao i ostale pogodnosti koje korisnici materijala mogu prihvatiti.

Kursevi učenja na daljinu koji su razvijeni na Moodle platformi, sastoje se od dve vrste elektronskih resursa:

- oni koji su namenjeni organizovanju obrazovnog procesa, uključuju obrazovne materijale koji se koriste u redovnoj obuci: predavanja sa prezentacijama i tekstom u elektronskom obliku; praktični materijali za obuku i relevantne prezentacije, beleške, komentari, konsultacije;
- provera rada studenata, kompletiranje zadataka, savladavanje materijala, proveravanje nivoa znanja i razumevanja materijala. Oblici kontrole koji se koriste su testovi, pisani zadaci, prezentacije na temu iz nastavnog materijala [4].

Pomoću Moodle platforme sem postavljanja raznih informacija, moguće je obezbediti kolaborativno učenje, kao i vođenje diskusija na forumima ili razgovora u chat grupama. Korisnici uvek mogu da preuzmu sadržaj, završavaju određene aktivnosti, a sve promene koje se beleže se sinhronizuju dok postoji Internet konekcija.

## 2.3. PLATFORME ZA TRENUTNU RAZMENU INFORMACIJA

Tempo savremenog doba nameće potrebu za dobijanjem informacija u što kraćem roku, bez obzira na prirodu informacije. Trenutna razmena informacija podrazumeva skraćenje perioda potrebnog da se dobije validna informacija, bilo da se odnosi na tekst, zvuk, sliku ili video zapis.

Skype platforma je bila početak takvog koncepta za slanje poruka i ostvarivanje poziva putem Interneta. Kao suparnik Skype-u pojavile su se platforme poput Viber-a i WhatsApp-a, koje su u početku bile fokusirane na pozive i poruke, da bi kasnije razvijale funkcionalnosti značajne za krajnjeg korisnika.

Jedna od prednosti aplikacija za trenutnu razmenu informacija jeste njihova dostupnost. Svako ko poseduje pametni telefon i ima pristup Internetu u prilici je da koristi ovu vrstu aplikacije.

### Viber

Viber platforma je besplatna višepatformska aplikacija koja omogućava komunikaciju putem poruka, glasovnu i video komunikaciju putem Interneta. Kod ovih vrsta platformi moguće je napraviti tzv. grupe, kako bi se omogućilo grupno slanje poruka ili grupni pozivi, što je veoma značajno za kolaborativan rad sa grupom ljudi istovremeno. Istovremeni rad sa grupom studenata na primer, omogućava značajnu brzu razmenu informacija između velikog broja učesnika grupe, što obezbeđuje značajno umanjeње prenosa netačnih informacija i pravovremeno dobijanje tačnih informacija u vezi nastavnog procesa.

Ova vrsta platformi je prvenstveno zamišljena kao aplikacija za pametne telefone, ali u međuvremenu su se pojavile i verzije aplikacije za desktop računare. Desktop verzije aplikacija su dostupne za većinu platformi ovog tipa, uključujući i WhatsApp.

## WhatsApp

WhatsApp platforma je nešto ranije startovala sa radom od Viber platforme. Kao i Viber platforma, WhatsApp je besplatna aplikacija koja može da se koristi na različitim operativnim sistemima mobilnih uređaja, ali i desktop računara. Ova platforma takođe omogućava razmenu tekstualnih poruka, glasovnih poruka, slika ili video zapisa, ali i govorne pozive.

WhatsApp platforma je od 2014. godine u vlasništvu kompanije Facebook (sada Meta Platforms). Interesantno je da Facebook takođe poseduje sličnu aplikaciju koja se zove Facebook Messenger i koja je prvenstveno stvorena za komunikaciju korisnika Facebook-a. Međutim, ova vrsta aplikacije nije ni blizu popularna kao WhatsApp. Zapravo, prema statistici [3] iz oktobra 2021. značajno je veći broj mesečnih korisnika WhatsApp platforme nego Facebook Messenger-a.

## Facebook

Kao društvena mreža, platforma Facebook se po prirodi ne može u potpunosti porediti sa aplikacijama kao što su Viber ili WhatsApp, međutim ukoliko korisnik ima nalog na ovoj platformi, primenom Facebook aplikacije putem mobilnog uređaja ili pristupom sa desktop računara, može se stvoriti utisak korišćenja platforme za trenutnu razmenu informacija. Obezbeđenim pristupom nekom Facebook nalogu (na primer, otvorenom nalogu informatičkih predmeta) mogu se u trenutku primiti sve notifikacije objavljenih postova ili poslanih slika na tom nalogu, a koje je postavio nastavnik s ciljem slanja obaveštenja studentima.

### 3. REZULTATI ISTRAŽIVANJA O STAVOVIMA STUDENATA I NJIHOVE PERCEPCIJE UTICAJA NA PROCES UČENJA

Ideja ovog rada, a samim tim i istraživanja, bila je da se ispituju stavovi studenata prema korišćenju Internet platformi koje koriste ili su ih koristili u procesu praćenja nastave i spremanja ispitne materije. Pitanja u anketi su podeljena po grupama prema vrstama Internet platformi, odnosno prilagođena su na način gde odgovori oslikavaju specifičnost primene neke od opisanih platformi.

Prvenstveno je bilo od značaja da se dobije informacija koliko njih je koristilo Internet platforme u toku procesa učenja, a koje su značajne za istraživanje. Rezultati ankete su pokazali da je Moodle platformu u procesu spremanja ispitne materije koristilo 72.6% studenata, Zoom platformu 84.5%, Facebook platformu 32.1%, a Viber ili WhatsApp platformu čak 89.3%. Broj korisnika platformi za trenutnu razmenu informacija je nadmašio i broj korisnika Zoom platforme, a koja je u prethodnom periodu bila vodeća platforma za održavanje nastave.

Prvom grupom pitanja ankete pokušano je da se proceni stav studenata prema korišćenju Zoom platforme. Rezultati odgovora na ovo pitanje mogu se videti u sledećoj tabeli:

**Tabela 1:** Pitanje: Po vašem mišljenju da li je Zoom platforma uspeła da zameni predavanja i/ili vežbe koje bi inače bile održavane u prostorijama fakulteta?

Odgovori	Učešće prema broju odgovora
Bez obzira na okolnosti, bolje je što su predavanja i/ili vežbe bile održavane putem ovih vrsta platformi	32.1%
Bilo bi bolje da su predavanja i/ili vežbe održavane u potpunosti uživo	42.9%
S obzirom na okolnosti, potpuno je bilo svejedno da li su predavanja održavana uživo ili preko ovih vrsta platformi	23.8%
Nemam stav o ovom pitanju	1.2%

Iz rezultata odgovora može se zaključiti da studenti u većem broju (42.9%) preferiraju nastavu koja bi se održavala u prostorijama fakulteta. Ipak, ne tako mali broj studenata (32.1%) bez obzira na specifične okolnosti smatra da je u svakom slučaju bolje održavati nastavu putem platformi za video komunikaciju.

Na pitanje da ocene kvalitet savladavanja gradiva korišćenjem platformi za video komunikaciju, 69% studenata je ocenilo korišćenje Zoom platforme najvišim ocenama (4 ili 5).



Druga grupa pitanja se odnosila na korišćenje platforme za elektronsko učenje, konkretno Moodle platforme. Stav studenata u vezi sa korišćenjem ovih platformi za dostupnost materijala se najbolje može percipirati u odgovorima, čiji se rezultati vide u sledećoj tabeli:

**Tabela 2** Pitanje: Po vašem mišljenju da li je ova vrsta platforme u domenu dostupnosti online materijala i davanja informacija uspeła da zameni sajt predmeta (ukoliko postoji)?

Odgovori	Učešće prema broju odgovora
Platforma za elektronsko učenje je preglednija i jednostavnija za preuzimanje nastavnih/ispitnih materijala i dobijanja informacija od sajta	28.6%
Ukoliko postoji predmetni sajt, jednostavnije je preuzimanje materijala i dobijanje informacija od korišćenja platforme za elektronsko učenje	39.3%
Potpuno je svejedno šta će se koristiti za svrhe preuzimanja materijala i dobijanja informacija	27.4%
Nemam stav o ovom pitanju	4.8%

Iz pregleda odgovora došlo se do zaključka da kada je u pitanju dostupnost nastavnih i ispitnih materijala i dobijanje informacija, jednostavnije je koristiti predmetni sajt, ukoliko postoji (39.3%). Ovo pitanje je upravo postavljeno iz razloga što studenti Moodle naloge predmeta u najvećem broju slučajeva koriste za preuzimanje materijala i interakciju sličnog karaktera.

Treća grupacija pitanja se odnosila na korišćenje platformi za trenutnu razmenu informacija. Na pitanje da li je mogućnost korišćenja ovih vrsta platformi omogućila lakši proces učenja, rezultati odgovora se mogu videti u sledećoj tabeli:

**Tabela 3:** Pitanje: Da li je korišćenje grupe putem platformi za trenutnu razmenu informacija (Viber, WhatsApp, Facebook) olakšalo proces učenja?

Odgovori	Učešće prema broju odgovora
Da, mogao/la sam da dobijem trenutnu informaciju u vezi nastavnog materijala ili ispita	92.9%
Ne, čak me nije zanimalo da učestvujem u radu grupe	2.4%
Potpuno je svejedno, informacije su mogle da se dobiju i na predavanjima	2.4%
Nema stav o ovom pitanju	2.4%

Veoma je zanimljivo da su u velikom broju (92.9%) studenti prepoznali ovu vrstu aplikacija kao veoma korisnu u procesu savladavanja nastavne materije, što potvrđuju i ocene kvaliteta informacija koje se dobijaju putem korišćenja ovih platformi, gde je najviše ocene (4 ili 5), odnosno izuzetan način da se brzo dođe do validnih informacija putem ovih platformi dalo 91.7% studenata.

#### 4. ZAKLJUČAK

Internet platforme za učenje poslednjih godina doživele su značajan razvoj i povećale svoje prisustvo u procesu nastave i učenja, što se naročito pojačalo u vreme pandemije virusa korona kada je to često bio i jedini način da se nastava realizuje. Ovde je predstavljeno više Internet platformi koje su korišćene u te svrhe pri čemu su detaljnije opisane one koje su najviše bile prisutne na Ekonomskom fakultetu Univerziteta u Beogradu. Od platformi za video komunikaciju to je Zoom, od platformi za elektronsko učenje to je Moodle i od platformi za trenutnu razmenu informacija to su Viber, WhatsApp i Facebook. Anketom je ispitano mišljenje studenata o korišćenju ovih platformi. Na osnovu toga je utvrđeno da su studenti u najvećoj meri koristili neke od ovih platformi, da su ih pozitivno ocenili, gde su najveći procenat zastupljenosti i visokih ocena imale platforme za trenutnu razmenu informacija. Jedan od zaključaka je da studenti za preuzimanje materijala više preferiraju sajt predmeta nego Moodle, kao i to da smatraju da bi bilo bolje da se nastava u potpunosti odvijala uživo.

Iako je pandemija u prvi plan stavila ove platforme opšti je utisak da će mnoge od njih naći i dalje svoje mesto u nastavnom procesu.

## LITERATURA

- [1] Cardullo, V., Wang, C. H., Burton, M., & Dong, J. (2021). K-12 teachers' remote teaching self-efficacy during the pandemic. *Journal of Research in Innovative Teaching & Learning*, 14(1), 32-45.
- [2] What Is eLearning? Definition, Examples, and QuickStart Guide: <https://www.ispringsolutions.com/blog/what-is-elearning>, (pristupano: 04.06.2022.)
- [3] WhatsApp, WeChat and Meta Messenger Apps – Global usage of Messaging Apps, Penetration and Statistics: <https://www.messengerpeople.com/global-messenger-usage-statistics/>, (pristupano: 07.06.2022.)
- [4] Ushanov, A., Morgunova, N., & Petunina, I. (2021). Internet Technologies in Distance Education. *International Journal of Emerging Technologies in Learning (iJET)*, 16(10), 85-95.

## TRANSFORMACIJA LOGIČKE SCHEME BAZE PODATAKA: OD RELACIONOG (T-SQL) MODELA DO DOKUMENTACIONO-ORIJENTISANOG MONGODB)

## TRANSFORMATION OF LOGICAL SCHEME OF DATABASE: FROM RELATIONAL (T-SQL) TO DOCUMENTATION-ORIENTED (MONGODB) DATABASE

NEMANJA RADULOVIĆ<sup>1</sup>, TATJANA STOJANOVIĆ<sup>2</sup>, SAŠA D. LAZAREVIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka fakultet, nr20203725@student.fon.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka fakultet, tatjana.stojanovic@fon.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka fakultet, sasa.lazarevic@fon.bg.ac.rs

**Režime:** *Kako ne postoji jedinstven način da se deo podataka predstavi korisniku, tako ne postoji jedan način za modeliranje baza podataka. Potrebno je da se razume problem koji rešava aplikacija i kako će ona proizvoditi, konzimirati i obraditi podatke. U ovom radu biće predstavljena transformacija logičkih shema iz relacionih baza podataka u dokument-orijentisane baze podataka bez sheme. Primenom nekih od smernica iz ovog rada, može se modelirati baze podataka koji odgovaraju potrebama aplikacije i rešavanjem korisničkih problema, kao i poboljšanju performansi. Kada aplikacija treba da se promeni, može se iskoristiti fleksibilnost koje pružaju dokument-orijentisane baze podataka bez shema da bi se ta promena primenila lako razvila u novi model podataka.*

**Ključne reči:** *MongoDB, Json, dokumenti, baze podataka, SQL*

**Abstract:** *As there is no single way to present a piece of data to the user, there is no single way to model databases. It is necessary to understand the problem that the application solves and how it will produce, consume and process data. This paper will present the transformation of logical schemas from relational databases to document-oriented databases without a schema. By applying some of the guidelines from this paper, one can model databases that meet application needs and solve user problems. When an application needs to change, the flexibility provided by document-oriented databases without schemas can be used to easily develop that change into a new data model.*

**Keywords:** *MongoDB, Json, documents, database, SQL*

### 1. UVOD

Ključni izazov u modeliranju podataka je balansiranje potreba aplikacije i karakteristika poput performansi *baze podataka* i obrazaca preuzimanja podataka. Kada se dizajnira model *baze podataka*, uvek se razmatra upotreba podataka u aplikaciji, upiti, ažuriranja i obrada, kao i veze između samih podataka. Za razliku od *SQL* baza podataka, gde se mora odrediti i deklarirati shema tabela pre umetanja podataka, *dokument-orijentisane* kolekcije *baze podataka* poput *MongoDB*, ne zahtevaju da njihovi *dokumenti* imaju istu shemu. To znači da *dokumenti* u jednoj kolekciji ne moraju da imaju isti skup polja, a tip podataka za polje može se razlikovati među dokumentima unutar kolekcija. Da bi se promenila struktura dokumenata u kolekciji, kao što je dodavanje novih polja, uklanjanje postojećih ili promena vrednosti polja u novi tip, *dokumenti* se ažuriraju u novu strukturu. Ova fleksibilnost olakšava mapiranje dokumenata u entitet ili objekat. Svaki dokument može da odgovara poljima podataka predstavljenog entiteta, čak i ako dokument ima značajne varijacije od drugih dokumenata u kolekciji. U praksi, međutim, *dokumenti* u kolekciji dele sličnu strukturu i može se primeniti pravila validacije dokumenata za kolekciju tokom operacija ažuriranja i upisivanja.

### 2. ŠABLON UGRAĐIVANJE PODATAKA

Kada se započne modelovanje entiteta u *MongoDB* bazi podataka, oni se tretiraju kao samostalne stavke predstavljene kao *Json* dokumenti. Na Slici 1 prikazan je primer čuvanja podataka o osobi i njenim kontakt podacima, dato je poređenje modelovanja ovog entiteta za relacionu i dokument-orijentisanu bazu.

Osoba(ID, Ime, Prezime)  
 Adresa(ID, UlicaBroj, Grad, Drzava, PttBroj, OsobaID)  
 Kontakt(ID, Opis, OsobaID, TipKontaktID)  
 TipKontakta(ID, Opis)

Osoba			Kontakt			
ID	Ime	Prezime	ID	OsobaID	Opis	TipKontaktID
1	Pera	Perić	1	1	+381 64 123 4567	1
			2	1	pera@ peric.com	2

Adresa						TipKontakta	
ID	UlicaBroj	Grad	Drzava	PttBroj	OsobaID	ID	Opis
1	Neka Ulica 1/1	Beograd	Srbija	11000	1	1	Telefon
						2	Email

```
Osoba dokument:
{
  "id": "1",
  "ime": "Pera",
  "prezime": "Perić",
  "adrese": [
    {
      "ulicaIBroj": "Neka Ulica1/1",
      "grad": "Beograd",
      "drzava": "Republika Srbija",
      "pttBroj": "11000"
    }
  ],
  "kontakti": [
    {
      "email": "pera@peric.com"
    },
    {
      "telefon": "+381 64 123 4567",
      "vrsta": "mobilni"
    }
  ]
}
```

Slika 1 Relacioni model(levo) i dokument-orjentisan(desno) model baze podataka za entitete Osoba, Adresa, Kontakt.

Kada se modeluje relacionim bazama podataka, strategija je da se normalizuju podaci. Normalizacija podataka obično uključuje uzimanje entiteta, kao što je osoba, i razbijanje na diskretne komponente. U primeru sa Slike 1, osoba može imati više zapisa kontakta, kao i više zapisa adresa. Kontakt se može dalje rasčlaniti daljim izdvajanjem uobičajenih polja kao što je tip, poslovni ili privatni broj telefon, email i slično. Vodeća premisa pri normalizaciji podataka je izbegavanje skladištenja suvišnih podataka u svakom zapisu i upućivanje preko reference na te podatke. U ovom primeru, da bi se pročitala osoba, sa svim njenim kontaktima i adresama, potrebno je da se koristi *JOIN* klauzula da bi ce efikasno sastavili (ili denormalizovali) podaci u vremenu izvršavanja.

```
SELECT o.Ime, o.Prezime, a.Grad, k.Email
FROM Osoba o
JOIN Kontakt k ON k.OsobaID = o.ID
JOIN TipKontakt tk ON tk.Id = k.KontaktTipID
JOIN Adresa a ON a.OsobaID = o.ID
```

Slika 2 JOIN klauzula za prikaz relacionog modela osobe

Ažuriranje jedne osobe sa njenim kontaktima i adresama zahteva operacije upisivanja u više pojedinačnih tabela. Koristeći gornji pristup, denormalizovala se evidencija o osobi, tako što se sve informacije u vezi sa ovom osobom, kao što su njeni kontakti i adrese, nalaze u jedan *Json* dokumentu. Pored toga, pošto ne postoje ograničenja za fiksnu shemu postoji mogućnost fleksibilnost, da kontakti mogu biti različitog oblika. Preuzimanje kompletnog zapisa o osobi iz baze podataka je sada jedna operacija čitanja kao jedan dokument i za jednu stavku. Ažuriranje zapisa o osobi, sa njihovim kontaktima i adresama, takođe je jedna operacija upisivanja za jednu stavku. Denormalizacijom podataka, aplikacija će možda morati da izda manje upita i upisivanja da bi izvršila uobičajene operacije.

### 3. KADA UGRAĐIVATI

Ugrađeni šablon modeliranja podataka treba koristiti u sledećim slučajevima: 1) Kada postoje sadržani odnosi između entiteta. 2) Između entiteta postoje odnosi jedan prema nekoliko. 3) Postoje ugrađeni podaci koje se ne menjaju često. 4) Postoje ugrađeni podaci čiji broj neće rasti bez ograničenja. 5) Postoje ugrađeni podaci koji se često ispituju zajedno.

#### Kada ne treba ugrađivati

Kada ne treba ugrađivati, iako je pravilo u *MongoDB bazi podataka* da se sve denormalizuje i svi podaci ubace u jedan dokument, postoje situacije kada ovo treba izbegavati, jedna od takvih modela je prikazan na sledećem primeru *Json* dokumenta:

Članak (ID, Naslov)  
Komentar (ID, Autor, Tekst)

Članak		Komentar		
ID	Naslov	ID	Autor	Tekst
1	Neki naslov	1	aca	komentar 1
		2	pera	komentar 2
		...	...	...
		∞ + 1	beskonacno	komentar beskonacno

```

Članak dokument:
{
  "id": "1",
  "naslov": "Naslov članka",
  "komentar": [
    {
      "id": 1,
      "autor": "aca",
      "komentar": "Komentar 1"
    },
    {
      "id": 2,
      "autor": "pera",
      "komentar": "Komentar 2"
    },
    ...
    {
      "id": ∞ + 1,
      "autor": "beskonacno",
      "komentar": "Komentar beskonacno"
    }
  ]
}

```

Slika 3 Levo je prikazan relacioni model baze podataka za primer članka i komentara, a desno dokument-orientisan model

Ovako bi mogao izgledati entitet članak sa ugrađenim komentarima kada bi se modelirao tipičan blog. Problem sa ovim primerom je što je niz komentara neograničen, što znači da ne postoji ograničenje broja komentara koji svaki pojedinačni članak može imati. Ovo može postati problem jer bi veličina članka mogla narasti beskonačno, pa je to i model koji bi trebalo izbegavati. Kako veličina dokumenta stavke raste, to će uticati na mogućnost prenosa podataka, kao i na čitanje i ažuriranje u velikoj meri.

#### 4. ŠABLON REFERENCIRANJE PODATAKA

Relacione baze podataka nisu jedino mesto gde se može kreirati odnose između entiteta. U dokument-orientisanim bazama podataka, možda se posedovati informacije u jednom dokumentu koje se odnose na podatke u drugim dokumentima. U ovakvim slučajevima ugrađivanje podataka nije dobra ideja, jer kada se ugrađeni podaci često koriste u svim dokumentima i često će se menjati. Kada postoji slučaj kao sa Slike 3, bolji model bi bio:

```

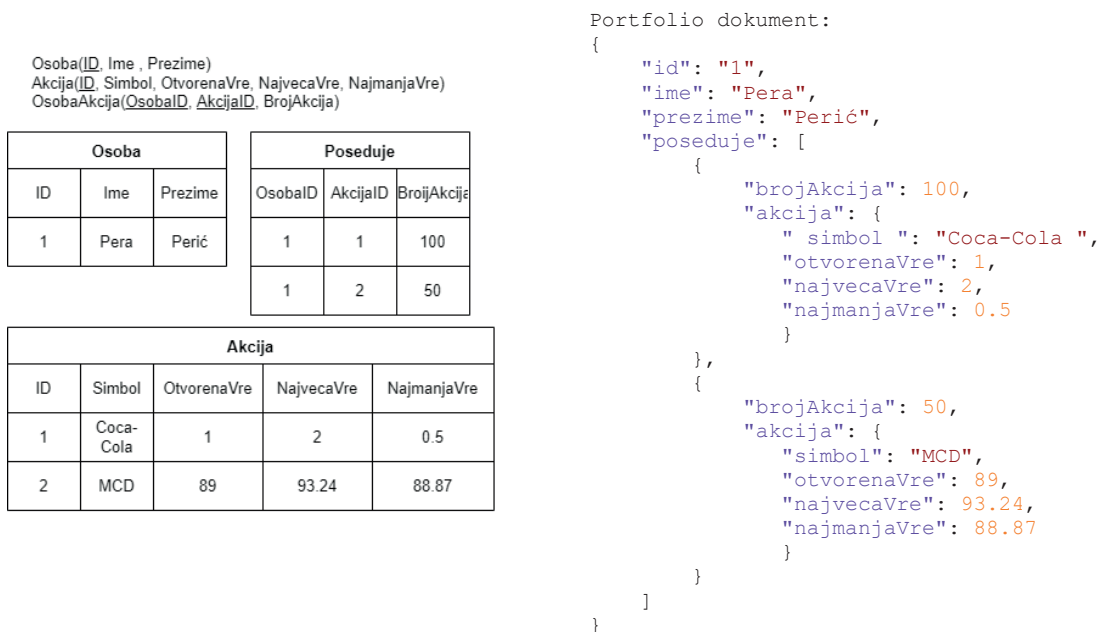
Članak dokument:
{
  "id": "1",
  "naslov": "Naslov članka",
  "recentComments": [
    {"id": 1, "autor": "aca", "komentar": "Komentar 1"},
    {"id": 2, "autor": "pera", "komentar": "Komentar 2"},
    {"id": 3, "autor": "mika", "komentar": "..."}
  ]
}
Komentari dokumenta:
[
  {"id": 1, "clanakId": "1", "autor": "aca", "komentar": "Komentar 1"},
  {"id": 2, "clanakId": "1", "autor": "pera", "komentar": "Komentar 2"},
  ...
  {"id": 199, "clanakId": "2", "autor": "ivana", "komentar": "Komentar 199"}
  ...
]

```

Slika 4 Bolji primer modelovanja dokument-orientisan bazom podataka članak i komentari

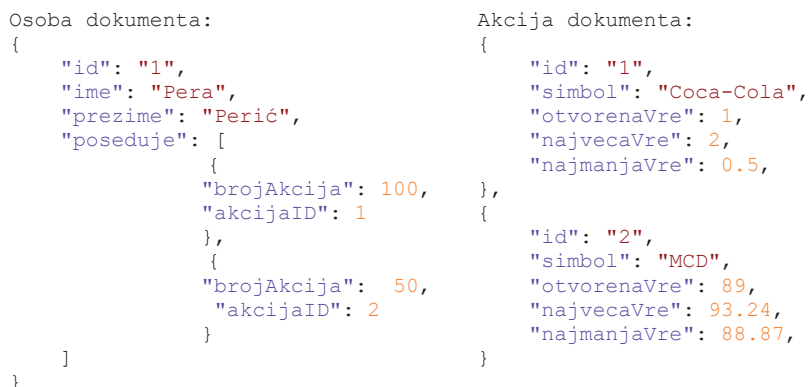
Primenom šablona referenciranja podataka ovaj model ima dokument za svaki komentar sa svojstvom koje sadrži ID članka. Ovo omogućava da članak sadrži bilo koji broj komentara čiji broj može efikasno da raste. Korisnici koji žele da vide više od najnovijih komentara, poslali bi upit prosledivši članakId koji bi trebalo da bude particioni ključ za dokument komentari. Drugi slučaj u kojem ugrađivanje podataka nije dobra ideja je kada se ugrađeni podaci često koriste u svim dokumentima i često će se menjati. Ovaj slučaj se može predstaviti na primeru modela portfolija akcija. Ako je odlučeno da se ugrade informacije o akcijama u svaki dokument portfolija, u okruženju u kojem se povezani podaci često menjaju, kao što je aplikacija za trgovanje akcijama, to će značiti da se stalno ažurira svaki dokument portfolija svaki put kada se trguje akcijama. "Coca-Cola" akcijama se može trgovati stotinama puta u jednom danu i hiljade korisnika bi moglo da ima te akcije u svom portfoliju. Sa modelom podataka kao što je prikazan na Slici 5, morao bi se ažurirati hiljade dokumenata portfolija mnogo puta svakog dana, što bi dovelo do sistema koji neće dobro skalirati.

Ugrađivanje podataka dobro funkcioniše u mnogim slučajevima, ali postoje scenariji kada će denormalizacija podataka izazvati više problema nego što vredi.



Slika 5 Primer modelovanja portfolija akcija, levo relacioni model, desno dokument-orjentisani model

Ne preporučuje se pravljenje sistema koji bi bolje odgovarali relacionoj bazi, ali jednostavni odnosi su u redu i mogu biti korisni. U primeru na *Slici 6* se koristi primer portfolija akcija, ali ovog puta se poziva na stavku akcija u portfoliju umesto da se ugrađuje. Na ovaj način, kada se stavka akcije često menja tokom dana, jedini dokument koji treba ažurirati je jedan dokument te akcije.



Slika 6 Bolji primer modelovanja modela portfolija akcija, osoba i akcije dokumenta

Neposredna mana ovog pristupa je ako se od aplikacije zahteva da prikaže informacije o svakoj akciji koja se drži prilikom prikazivanja portfolija osobe, u ovom slučaju bi trebalo da se izvrši upit više puta do *baze podataka* da bi se učitale informacije za svaki dokument o akcijama. Ovde je doneta odluka da se poboljša efikasnost operacija upisanja, koje se dešavaju često tokom dana, ali zauzvrat ugrožava operacije čitanja koje potencijalno imaju manji uticaj na performanse ovog konkretnog sistema. Pošto trenutno ne postoji koncept ograničenja, stranog ključa ili drugog, bilo koji odnosi između dokumenata koje postoje u dokumentima su zapravo "slabe veze" i neće biti verifikovane od strane same *baze podataka*. Kako bi se proverilo da podaci na koje se dokument poziva zaista postoje, onda to treba da se uradi u preko aplikacije.

## 5. KADA REFERENCIRATI PODATKE

Referenciranje se generalno koristiti kada normalizovani modeli podataka: 1) Predstavljanje odnosa jedan-prema-više. 2) Predstavlja odnose više-prema-više. 3) Povezani podaci se često menjaju. 4) Referentni podaci mogu biti neograničeni.

## 6. GDE STAVITI VEZU

Rast veze će pomoći da se odredi u kom dokumentu da se čuva referenca. Ako se pogleda primer na *Slici 7* koji modelira izdavače i knjige:

Knjige			Izdavac	
ID	Naslov	IzdavacID	ID	Ime
1	Learning C#	1	1	O'Reilly Media
2	Seven NoSQL Databases in a Week	1		
...	....	....		
100	MongoDB: The Definitive Guide	1		

Izdavač dokument

```
{
  "id": "oreilly",
  "ime": "O'Reilly Media"
  "books": [ 1, 2, ..., 1000 ]
}
```

Knjige dokumenta:

```
{ "id": "1", "naslov": "Learning C#" }
{ "id": "2", "naslov": "Seven NoSQL Databases in a Week" }
...
{ "id": "100", "naslov": "MongoDB: The Definitive Guide" }
```

*Slika 7 Primer modelovanja entiteta izdavač i knjiga, relacioni (levo) i dokument-orjentisan(desno) model*

Kako je broj knjiga po izdavaču mali sa ograničenim rastom, čuvanje reference knjige u dokumentu izdavača može biti korisno. Međutim, ako je broj knjiga po izdavaču neograničen, onda bi ovaj model podataka doveo do promenljivih, rastućih nizova, kao u primeru na *Slici 8* dokumenta izdavača. Malo menjanje stvari bi rezultiralo modelom koji i dalje predstavlja iste podatke, ali sada se izbegava velike promenljive kolekcije.

Izdavač dokument

```
{
  "id": "oreilly",
  "ime": "O'Reilly Media"
}
```

Knjige dokumenta:

```
{ "id": "1", "naslov": "Learning C#", "izdavacID": "oreilly" }
{ "id": "2", "naslov": "Seven NoSQL Databases in a Week", "izdavacID": "oreilly" }
...
{ "id": "100", "naslov": "MongoDB: The Definitive Guide", "izdavacID": "oreilly" }
```

*Slika 8 Bolj primer modelovanja entita izdavač I knjiga*

U gornjem primeru, izbačen je neograničena kolekcija na dokumentu izdavača. Umesto toga, postoji samo referenca na izdavača na svakom dokumentu knjige.

## 7. MODELIRANJE VEZE VIŠE-PREMA-VIŠE

Veza više-prema-više u relacionoj bazi podataka se često modeluju sa spojenim tabelama, koje samo spajaju zapise iz drugih tabela zajedno. Učitavanje autora sa njegovim knjigama, ili učitavanje knjige sa njegovim autorom, uvek bi zahtevalo najmanje dva dodatna upita ka bazi podataka. Jedan upit za dokument koji se pridružuje, a zatim drugi upit za preuzimanje dokumenta koji se spaja.

AutorKnjiga		Autor			Knjiga	
AutorID	KnjigaID	ID	Ime	Prezime	ID	Naslov
1	1	1	Kristina	Chodorow	1	MongoDB: The Definitive Guide
2	1	2	Michael	Dirolf	2	50 Tips and Tricks for MongoDB Developers
1	2					

Autor dokumenta:

```
{ "id": "a1", "name": "Kristina Chodorow" }
{ "id": "a2", "name": "Michael Dirolf" }
```

Knjige dokumenta:

```
{ "id": "b1", "naslov": "MongoDB: The Definitive Guide" }
{ "id": "b2", "naslov": "50 Tips and Tricks for MongoDB Developers" }
```

Spajanje dokumenta:

```
{ "autorId": "a1", "knjigaId": "b1" }
{ "autorId": "a2", "knjigaId": "b1" }
{ "autorId": "a1", "bookId": "b2" }
```

*Slika 9 Autor i knjiga relacioni model(levo), dokument-orjentisan(desno) model*

Ako ovo spajanje samo spaja dva podatka, zašto ga onda ne odbaciti u potpunosti, kao što je prikazano na *Slici 12*, Sada, ako autor postoji, odmah se zna koje knjige je napisao, i obrnuto, da se učita dokument knjige,

znao bi se ID autora. Ovo šteti taj posrednički upit prema tabeli spajanja i smanjuje broj povratnih putovanja servera koje aplikacija mora da napravi.

```
Autor dokumenta:
{"id": "a1", "name": "Kristina Chodorow", "books": ["b1", "b2"]}
{"id": "a2", "name": "Michael Dirolf", "books": ["b1"]}

Knjige dokumenta:
{"id": "b1", "naslov": "MongoDB: The Definitive Guide", "authors": ["a1", "a2"]}
{"id": "b2", "naslov": "50 Tips and Tricks for MongoDB Developers", "authors": ["a1"]}
```

*Slika 10 Bolji primer veze više-prema-više u dokument-orjentisanim bazama podataka*

## 8. HIBRIDNI MODELI

Postoje i hibridni modeli baza podataka koji koriste i ugrađivanje (ili denormalizaciju) i referenciranje (ili normalizaciju) podataka. Svaki pristup ima prednosti i kompromise, ne mora uvek da bude ili-ili. Na osnovu specifičnih obrazaca korišćenja i radnih opterećenja aplikacije, mogu postojati slučajevi u kojima mešanje ugrađenih i referentnih podataka ima smisla i može dovesti do jednostavnije logike aplikacije sa manjim brojem povratnih poziva ka serveru uz održavanje dobrog nivoa performansi.

```
Autor dokumenta:
{
  "id": "a1",
  "ime": "Kristina",
  "prezime": "Chodorow",
  "brojKnjiga": 2,
  "knjige": ["b1", "b2", "b3"],
  "slike": [
    {"thumbnail": "https://....png"}
    {"profile": "https://....png"}
    {"large": "https://....png"}
  ]
},
{
  "id": "a2",
  "ime": "Michael",
  "prezime": "Dirolf",
  "brojKnjiga": 1,
  "knjige": ["b1"],
  "slike": [
    {"thumbnail": "https://....png"}
  ]
}

Knjiga dokumenta:
{
  "id": "b1",
  "naslov": "MongoDB: The Definitive Guide",
  "autori": [
    {"id": "a1", "name": "Kristina Chodorow",
     "thumbnailUrl": "https://....png"},
    {"id": "a2", "name": "Michael Dirolf",
     "thumbnailUrl": "https://....png"}
  ],
  {
    "id": "b2",
    "name": "50 Tips and Tricks for MongoDB
    Developers",
    "authors": [
      {"id": "a1", "name": "Kristina Chodorow",
       "thumbnailUrl": "https://....png"},
    ]
  }
}
```

*Slika 11 Primer hibridnog modela*

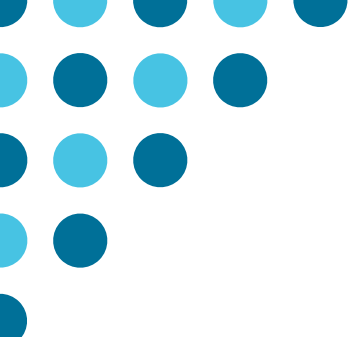
## 9. ZAKLJUČAK

Kako ne postoji jedinstven način da se deo podataka predstavi korisniku, tako ne postoji jedan način za modeliranje baza podataka. Potrebno je da se razume problem koji rešava aplikacija i kako će ona proizvoditi, konzumirati i obraditi podatke. Primenom nekih od smernica iz ovog rada, može se modelirati baze podataka koji odgovaraju potrebama aplikacije i rešavanjem korisničkih problema, i time poboljšati performance same aplikacije. Takođe kada aplikacija treba da se promeni, može se iskoristiti fleksibilnost koje pružaju dokument-orjentisane baze podataka bez shema da bi se ta promena primenila lako razvila u novi model podataka.

## 10. LITERATURA

- [1] Ted Hills (2016). NoSQL and SQL Data Modeling: Bringing Together Data, Semantics, and Software
- [2] George Tillmann (2017). Usage-Driven Database Design
- [3] Shannon Bradshaw, Eoin Brazil & Kristina Chodorow (2019). MongoDB: The Definitive Guide: Powerful and scalable data storage
- [4] <https://www.mongodb.com/docs/> - MongoDB Documentation





# ISTRAŽIVANJE I RAZVOJ

## RESEARCH AND DEVELOPMENT



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business



## GAPS AND NEEDS ANALYSIS OF RESEARCH INFRASTRUCTURES IN THE WESTERN BALKAN REGION

LAZAR ŽIVKOVIĆ<sup>1</sup>, ĐURO KUTLAČA<sup>2</sup>

<sup>1</sup> University of Belgrade – Institute Mihajlo Pupin, lazar.zivkovic@pupin.rs

<sup>2</sup> University of Belgrade – Institute Mihajlo Pupin, djuro.kutlaca@pupin.rs

---

**Abstract:** *This paper aims to identify the current state and development of research infrastructures in the Western Balkan region in order to increase the general understanding of the Western Balkan's current research capabilities to conduct state of the art research projects. Considering that global research infrastructures are becoming increasingly complex and expensive, participation in large pan-European research infrastructures is considered as important condition to improve national research capabilities and to foster research mobility. When it comes to the participation in pan-European research infrastructures, WB economies are lagging significantly behind EU member states, participating only in a few top research infrastructures. To successfully integrate into the European Research Area, Western Balkan economies need to invest more efforts in recognising research infrastructures as strategically important for future economic development.*

**Keywords:** *Research infrastructures, Western Balkans, pan-European RIs, ESFRI*

### 1. INTRODUCTION

Research Infrastructures (RIs) represent the basic tools for conducting excellent research. Taking into account that in today's world RIs are becoming increasingly complex and expensive, it is important to ensure that investments in RIs are efficient as possible. Research infrastructures include major research facilities, laboratory environments, complex digital research systems and databases, etc. According to the EU Regulation, RIs are defined as follows (EU Regulation, 2013): “RIs are facilities, resources and services that are used by the research communities to conduct research and foster innovation in their fields. They include: major scientific equipment (or sets of instruments), knowledge-based resources such as collections, archives and scientific data, e-infrastructure, such as data and computing systems and communication networks and any other tools that are essential to achieve excellence in research and innovation”.

The concept of RIs has a political origin and apparent political usefulness, and it was most likely invented by EU policymakers for political purposes (Olaf, 2020). Within the EU's innovation growth policy, RIs have received significant attention. In 2002 a special EU counsellor body was formed, the European Strategy Forum for Research Infrastructures (ESFRI), and in 2008 the European Commission also established a new organizational form – the European Research Infrastructure Consortium (ERIC). In line with the EU's policy and guidelines, most EU member states developed their national roadmaps for research infrastructures. These roadmaps are vital blueprints which enable countries to set national priorities and to earmark funds for both national and pan-European RIs including ESFRI ones (National Roadmaps, 2022).

RI Roadmaps recently created by all WB economies, set the principles for the future development of RIs and showcases the existing research potential of WB economies. It assists ministries in charge of research how to better leverage investments in research infrastructures to ensure their national and international relevance, as well as to ensure their availability to the entire research and business community in the region.

This paper presents the key results of research commissioned by Regional Cooperation Council (RCC) and carried out within the consultancy services titled „Support to RCC Secretariat for the Creation of Western Balkans' Research and Innovation Infrastructure Roadmap“. Research methodology for identifying RIs in the WB region encompassed the following research methods: desktop research including document analysis as a systematic procedure for reviewing and evaluating documents - both printed and electronic material and online interviews conducted with the line ministries in charge of research and innovation in all Western Balkans economies to obtain the existing roadmaps of research infrastructures and overviews of key RIs.

Next paragraph provides the general overview of research and development sector in the WB region. Paragraph 3 identifies key RIs in the region and shows the current state of development in each out of six

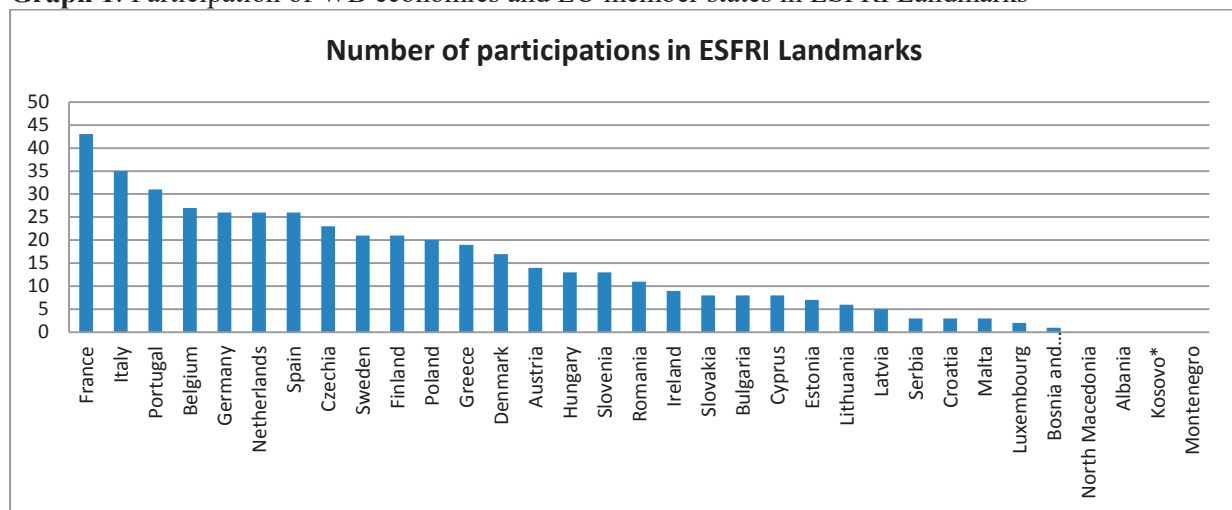
thematic areas. Finally, last section concludes the research and includes the key recommendations to improve the situation in the region.

## 2. GENERAL OVERVIEW OF R&D IN THE WESTERN BALKAN REGION

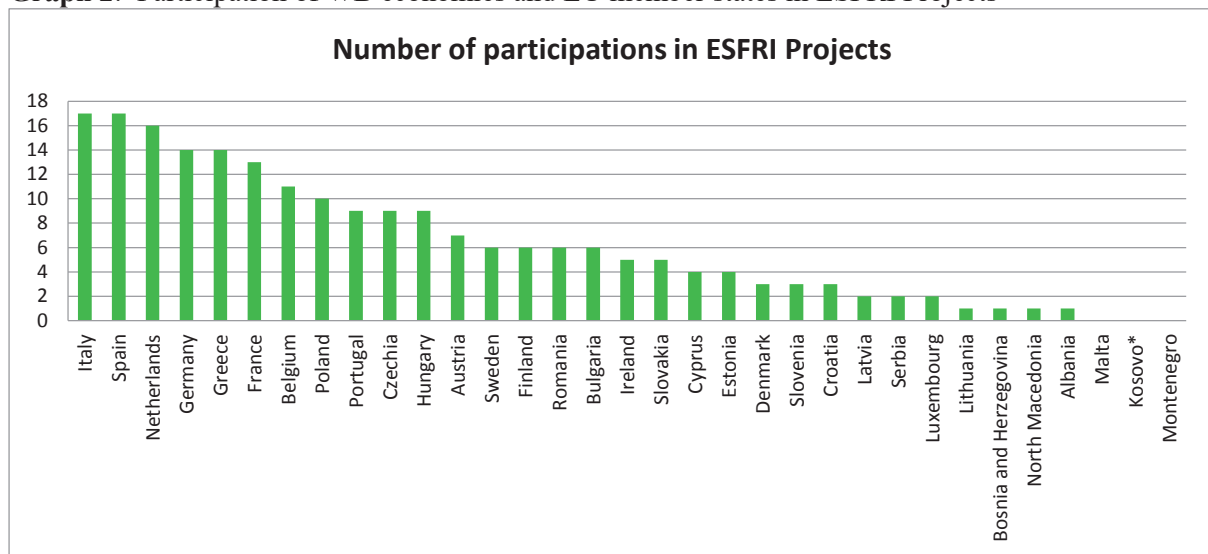
While developed EU member states strive to achieve the gross expenditures on research and development (GERD) of 3% of GDP, in the context of WB economies, these figures have been consistently low and some economies are experiencing a declining trend. According to the Eurostat data (2021), GERD in North Macedonia was 0.38 percent of GDP, in Montenegro 0.49 percent of GDP, and in Serbia 0.9 percent of GDP, while official data for other economies are not available. The investments in R&D consist essentially of funds allocated by the public sector (government and higher education sector), with private sector participation in R&D of around 0.10% of GDP (Eurostat, 2021).

WB region lacks large RIs and the current integration into pan-European RIs is insufficient. Graph 1 shows that WB economies are considerably lagging behind EU member states when it comes to the number of participations in ESFRI Landmarks. The ESFRI Landmarks are RIs that were implemented or reached an advanced implementation phase under the Roadmap and that represent major elements of competitiveness of the ERA. Currently, Serbia participates in 3 landmarks, BiH in 1 Landmark while other economies do not participate in any of the selected top ESFRI Landmarks.

**Graph 1:** Participation of WB economies and EU member states in ESFRI Landmarks



Regarding ESFRI projects, Serbia participates in 1 project, Albania, North Macedonia and BiH participates in 1 project, while other 2 economies do not participate in ESFRI projects (Graph 2). The ESFRI Projects are RIs in their preparation phases which have been selected for the excellence of their scientific case and their maturity, according to a sound expectation that the project will enter the Implementation Phase within the ten-year term.

**Graph 2:** Participation of WB economies and EU member states in ESFRI Projects

RIs in the WB region are mainly based on research equipment and research facilities located at public research institutes and universities. Aside from a few well equipped research institutions, most research equipment is outdated and need considerable investments in order to be fully operational. Degree of utilization of existing research equipment is at a low level due to the lack of an open access policy to RIs.

Most research institutions/laboratories do not have an official access policy. Internal procedures for accessing research equipment within the higher education institutions do not differ greatly between WB economies. Most institutions have open access for research staff, while external users are required a special permit, which includes: the official submission of requests for access to research infrastructure and formal approval by the management.

### 3. OVERVIEW OF THE CURRENT STATE OF RIS IN THE WB ECONOMIES GROUPED THEMATICALLY

The table below shows identified large pan-European RIs that WB economies have access to. WB economies have been achieving the highest level of participation in pan-European RIs within the areas of social and cultural innovation and e-infrastructures.

**Table 1:** Large pan-European RIs available to WB economies

Research Infrastructure	Participating WB Economies
<b>ENERGY</b>	
European Organization for Nuclear Research (CERN)	<b>Serbia</b> is a full member; <b>North Macedonia, Albania and BiH</b> has signed an International Cooperation Agreement <b>Montenegro</b> signed an Agreement on scientific and technical cooperation in 2007
<b>ENVIRONMENT</b>	
eLTER – Integrated European Long-Term Ecosystem	<b>Serbia</b> is a full member
<b>HEALTH &amp; FOOD</b>	
METROFOOD-RI Infrastructure for Promoting Metrology in Food and Nutrition	<b>North Macedonia</b> is a full member
International Institute for Sustainable Technologies in South East Europe (SEEIIST)	<b>All WB economies</b> are the partners in the project
Population Health Information Research Infrastructure	<b>Serbia and BiH</b> are the project members
<b>PHYSICAL SCIENCES AND ENGINEERING</b>	
HL-LHC – High-Luminosity Large Hadron	<b>Serbia</b> is a full member
Virtual Atomic and Molecular Data Centre (VAMDC)	<b>Serbia</b> is a member of this RI
<b>SOCIAL AND CULTURAL INNOVATION</b>	
Consortium of European Social Science Data Archives (CESSDA)	<b>Serbia and North Macedonia</b> are the full members of the CESSDA; BiH, Albania, Montenegro and Kosovo* are the potential partners currently outside the consortium.

DARIAH ERIC – Digital Research Infrastructure for the Arts and Humanities	<b>Serbia and BiH</b> are the members
European Social Survey (ESS)	All WB economies participated in at least one round of the ESS
OPERAS	<b>Serbia</b> is a member
RESILIENCE	<b>BiH and Albania</b> are the members
<b>E-INFRASTRUCTURES</b>	
GEANT	<b>Serbia, North Macedonia, Albania, Montenegro and Kosovo*</b> are the members
GEANT Project GN4-3	Serbia, Albania, North Macedonia and Montenegro are the members
EOSC Association – European Open Science Cloud	Serbia, BiH North Macedonia are the members
EGI Federation	<b>Serbia</b> is the members
EUROCC - National Competence Centres in the framework of EuroHPC	<b>North Macedonia and Montenegro</b> are the members of the EuroCC project
EGI-ACE: Advanced computing for research	<b>North Macedonia</b> is the project member of the EGI-ACE project
NI4OS-Europe: National Initiatives for Open Science in Europe	<b>Serbia, Albania, North Macedonia, BiH and Montenegro</b> are the members

The following text provide an overview of the current state of RIs in the WB economies grouped thematically within the following 6 thematic areas, according to ESFRI classification:

- Energy
- Environment
- Health and Food Sciences
- Physical Sciences and Engineering
- Social and Cultural Innovation
- E-infrastructures

It is worth noting that research infrastructures are grouped into 6 broad thematic areas, although many of them are multidisciplinary in nature, covering more than one area.

### 3.1. Energy

Considering the high gap between WB economies and EU member states when it comes to the use of renewable energy sources, it is expected from the WB economies to speed up the decarbonisation process. It is clear that research projects and research infrastructures in the area of energy are needed more than ever. However, the mapping exercise has shown that the participation of WB economies in high-level RIs and research projects in this area is extremely limited.

With regards to the participation in large pan-European RIs, Serbia is the full member of the European Organization for Nuclear Research (CERN), participating in projects and experiments conducted in CERN. Other economies are mainly allowed to access educational contents and trainings based on International cooperation Agreement signed with CERN.

With regards to the participation in the Horizon Europe (H2020) programme's thematic priority "Secure, Clean and Efficient Energy", more than 65% of the total EC contribution (14,31 million out of 21,71 million) to the projects from the WB region was implemented by research institutions from Serbia, indicating very low research performance of other WB economies in this area.

Serbia is the central node in the regional network while the project cooperation between the research institutions in the WB region is rather low.

### 3.2. Environment

WB economies are not integrated into the most important European RIs in the area of Environment. Regarding RIs recognised by ESFRI, Serbia is a member of one distributed RI: eLTER (Long-Term Ecosystem Research in Europe), while other economies are not participating in any of existing ESFRI RIs.

The participation of the WB region in Horizon 2020 programme's thematic priority: "Climate Action, environment, resource efficiency and raw materials" is modest, getting only 0.34% of the total EC contribution within this thematic priority.

Serbia is the most advanced economy with the participation in 36 out of a total of 65 projects. Research cooperation between WB economies is weak, while the most important research partners are coming from Germany, Spain and Italy.

### 3.3. Health & Food

Speaking of the current strategic framework, it follows that Health and Food Sciences are one of the most prioritised research fields in all WB economies. As a result of the strong potential in the WB region, it is worth noting that the first regional RI is currently under development: the South East European International Institute for Sustainable Technologies. It is currently in the design study development phase. The SEEIST will support the hadron cancer therapy and biomedical research with protons and heavy ions as the main research area for the Institute.

North Macedonia is a member of Infrastructure for Promoting Metrology in Food and Nutrition (METROFOOD-RI). This RI consists of physical infrastructure and an electronic infrastructure to coordinate and integrate existing networks of plants, laboratories, experimental fields/farms for crop production/animal breeding, small-scale plants for food processing and storage, and kitchen labs for food preparation.

Serbia has remarkable research results in the fields of health and food even at the European level. BioSense Institute from Serbia is a national centre of excellence and the most successful institution in the H2020 programme in Serbia. Besides the number of relevant research projects, one current H2020 project of particular importance is Antares.

### 3.4. Physical Sciences and Engineering

The level of integration of WB economies into pan-European RIs in the area of Physical sciences is weak. In addition, the WB economies have participated in only 3 of 1.451 H2020 projects within the thematic area of Physical sciences. Research activities within this area are mainly oriented towards research papers in national and international publications, with a very limited applied research.

Serbia is the only WB economy that participates in the CERN's project of high importance: High-Luminosity Large Hadron Collider (HL-LHC), which is the highest-energy particle collider in the world.

### 3.5. Social and Cultural Innovation

Although social sciences are not among the most prioritised research fields in the WB region, WB economies are the members of a certain number of important pan-European RIs. Serbia is participating in the Consortium of European Social Science Data Archives (CESSDA), European Social Survey (ESS-ERIC), The Digital Research Infrastructure for the Arts and Humanities (DARIAH) and the Open Scholarly Communication in the Social Sciences and Humanities (OPERAS). BiH has recently become a member of DARIAH and a Research Infrastructure for Religious Studies (RESILIENCE). Albania is the member of RESILIENCE while other WB economies are not members of pan-European RIs. However, it is important to note that most WB economies conducted at least one Social survey coordinated by ESS-ERIC.

Serbia North Macedonia and BiH have made significant progress in developing their national infrastructures/data centres providing long-term preservation and distribution of research data in the social sciences.

### 3.6. E-Infrastructures

While Information and Communication Technologies (ICT) represent the absolute priority in all WB economies, they are lagging behind the EU in terms of using High-Performance Computing (HPC). However, the diversity among WB economies is evident. While Serbia and North Macedonia have HPC infrastructures and distinguished research organizations appointed to represent them in European HPC-related projects and infrastructures, other WB economies do not have HPC infrastructures in their economies.

In the last decade, there were several South-East European e-infrastructure initiatives (such as: VI-SEEM, NI4OS-Europe, EGI-ACE, EUROCC, GN4-3 and others) aimed at creating conditions for equal participation of the less-resourced economies of the WB region in European networking and Grid computing trends by providing e-infrastructure resources, application support and training. The above initiatives have raised awareness of line ministries on the necessity of local programmes and financial support for e-Infrastructures.

Research cooperation between institutions from WB region is on much higher level than in other thematic areas. Most H2020 projects in the area of e-infrastructures include several research institutions from more than 2 WB economies which makes this cooperation better than in other fields.

#### 4. CONCLUSION

If WB economies want to integrate into European Research Area, investments in R&D need to be increased significantly. Additionally, RIs need to be recognised as an important asset on that path. Analysis of the current state of R&D shows a very low level of investments in current research equipment which is mainly outdated in most WB economies. The greater budget allocation for R&D is a necessary condition for the future development of RIs in the Region. Policy measures should be directed in two directions: direct investments in research equipment and facilities and stimulations of the business sector to invest in RIs.

Research collaboration between research institutions within the region is low. Since some WB economies are too small to enhance internal research capacities, fostering research cooperation in the region would result in increased research excellence and relevance in the WB region.

Research results of this paper have shown extremely low integration of WB economies into large pan-European RIs. National governments in the WBs need to consider this issue since the benefits of accessing large RI for researchers and research institutions are multiple. Due to the lack of appropriate research equipment, enhanced integration into pan-European RIs would allow the implementation of high-level research projects that cannot be implemented within the economy and within the WB region.

Research has also shown that all WB economies lack open access to RIs. To improve this situation, line ministries in all WB economies should encourage research institutions to adopt their Open Access Policy documents in order to increase university-industry research collaboration as well as collaboration between research institutions.

#### Acknowledgement

The research described in this paper was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia

#### REFERENCES

- [1] ESFRI (2021). ESFRI Roadmap 2021, retrieved from: <https://www.esfri.eu/esfri-roadmap-2021>
- [2] EU Regulation (2013) Regulation (EU) No 1291/2013 of 11 December 2013 - "Establishing Horizon 2020 – the Framework Programme for Research and Innovation 2014-2020"
- [3] Eurostat (2021). Eurostat Database, retrieved from: <https://ec.europa.eu/eurostat/data/database>
- [4] National Roadmaps (2022). Retrieved 08. Jun 2022 from: [www.esfri.eu/national-roadmaps](http://www.esfri.eu/national-roadmaps).
- [5] Olof, H. (2020). Research infrastructures in Europe: The hype and the field, *European Review*. 28(4):617-635



## MERENJE EFIKASNOSTI NACIONALNIH SISTEMA ZASNOVANO NA TRIPLE HELIX MODELU

### MEASURING EFFICIENCY OF NATIONAL SYSTEMS BASED ON THE TRIPLE HELIX MODEL

MILICA JOVANOVIĆ<sup>1</sup>, GORDANA SAVIĆ<sup>2</sup>, MAJA LEVI-JAKŠIĆ<sup>3</sup>, MILAN MARTIĆ<sup>4</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, [milica.jovanovic@fon.bg.ac.rs](mailto:milica.jovanovic@fon.bg.ac.rs)

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, [gordana.savic@fon.bg.ac.rs](mailto:gordana.savic@fon.bg.ac.rs)

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, [maja-levi.jaksic@fon.bg.ac.rs](mailto:maja-levi.jaksic@fon.bg.ac.rs)

<sup>4</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, [milan.martic@fon.bg.ac.rs](mailto:milan.martic@fon.bg.ac.rs)

**Rezime:** Eksponencijalna stopa tehnološkog razvoja značajno obeležava poslovno, nacionalno i globalno okruženje. Zbog toga, nacionalne vlade definišu strategije i politike naučnog, tehnološkog i inovacionog razvoja, koje, zbog velikog broja učesnika i aktivnosti, zahtevaju sistemski pristup i višedimenzionalno sagledavanje performansi. Triple Helix model saradnje privrede, univerziteta i države pruža adekvatan okvir za sistemsko praćenje upravljanja tehnologijom na različitim nivoima ekonomije i društva. U ovom radu predlaže se kompozitni indeks efikasnosti nacionalnih sistema koji se zasniva na dvofaznom modelu analize obavljanja podataka (DEA) primenom: 1) regiona sigurnosti i DEA modela superefikasnosti; 2) DEA modela zasnovanom na Benefit-of-Doubt pristupu bez eksplicitnih ulaza. Model je primenjen nad bazom podataka o indikatorima performansi koja uključuje 34 OECD zemlje i Srbiju. Rezultati su omogućili identifikovanje kritičnih oblasti za unapređenje nacionalnih politika i strategija, kao i nadležnosti za unapređenje i usmeravanje strateških odluka.

**Gljučne reči:** Analiza obavljanja podataka, Triple Helix, tehnološki razvoj, nacionalna strategija, kompozitni indeks.

**Abstract:** Exponential rate of technological development significantly determines business, national, and global environment. Thus, national governments define strategies and policies of scientific, technological, and innovation development, which require systemic approach and multidimensional performance evaluation due to the considerable number of participants and activities. The Triple Helix model of business, academic, and government collaboration provides an adequate framework for systemic monitoring of technology management at different economic and societal levels. This paper suggests composite index for national systems based on the two-phase Data Envelopment Analysis (DEA) model by applying: 1) assurance regions and DEA superefficiency model, 2) DEA based Benefit-of-Doubt approach without explicit inputs. The model is applied on database of performance indicators for 34 OECD countries and Serbia. The results provide identification of critical areas for improving national policies and strategies, as well as responsibilities for improving strategic decisions.

**Keywords:** Data Envelopment Analysis, Triple Helix, technological development, national strategy, composite index.



## PRILOG METODOLOGIJI PRAĆENJA REALIZACIJE 4S – POUKE PRAKSE EU

### CONTRIBUTION TO THE METHODOLOGY FOR MONITORING OF REALISATION OF THE 4S – LESSONS OF EU PRACTICE

ĐURO KUTLAČA<sup>1</sup>, DUŠICA SEMENČENKO<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu, Institut „Mihajlo Pupin“, Beograd, [djuro.kutlaca@pupin.rs](mailto:djuro.kutlaca@pupin.rs)

<sup>2</sup> Univerzitet u Beogradu, Institut „Mihajlo Pupin“, Beograd, [dusica.semencenko@pupin.rs](mailto:dusica.semencenko@pupin.rs)

**Rezime:** 4S (Strategija pametne specijalizacije Srbije) usvojena je 2020, a Akcioni plan za njenu realizaciju usvojen je godinu dana kasnije, 2021. U radu se analiziraju metodološki aspekti praćenja realizacije 4S, poređenjem procedura predloženih u postavci „smart specialisation“ pristupa Evropske unije sa procedurama definisanim u usvojenoj 4S u Srbiji, uz primer pristupa primenjenog u Hrvatskoj, koji je više primeren uslovima realizacije strategije pametne specijalizacije u regionu Zapadnog Balkana.

**Ključne reči:** Strategija pametne specijalizacije, Praćenje i ocenjivanje realizacije S3, EU, Srbija

**Abstract:** 4S (Smart Specialization Strategy of Serbia) was adopted in 2020, and the Action Plan for its implementation was adopted a year later, in 2021. The paper analyzes the methodological aspects of monitoring the implementation of 4S, comparing the procedures proposed in the "smart specialization" approach of the European Union with the adopted 4S in Serbia, supported with the example of the approach applied in Croatia, which is more appropriate to the conditions of the implementation of the smart specialization strategy in the Western Balkans.

**Keywords:** Smart Specialization Strategy, Monitoring and Evaluation of S3 Implementation, EU, Serbia

#### 1. UVOD

Koncept „pametne specijalizacije“ definisala je ekspertska grupa *Knowledge for Growth* [6] kao deo pripreme nove Kohezione politike Evropske unije (EU) za period 2014–2020. Pored metodološke podrške, Evropska komisija je izgradila S3 platformu u Sevilji kao ekspertsku podršku u proceduri izrade nacionalnih i/ili regionalnih istraživačkih i inovativnih strategija za pametnu specijalizaciju (*Research and Innovation Smart Specialisation Strategy– RIS3*), [15]. Tu proceduru čine sledeći koraci [12]:

- Korak 1: Analiza državnog ili regionalnog konteksta i potencijala, u odnosu na druge države i regione;
- Korak 2: Upravljanje: obezbeđivanje učešća i vlasništva – uspostavljanje inkluzivne strukture i podsticaja za široko učešće zainteresovanih strana;
- Korak 3: Vizija – stvoriti zajedničku viziju među zainteresovanim stranama – razrada ukupne vizije za region/državu;
- Korak 4: Određivanje prioriteta – Identifikacija prioriteta – izbor ograničenog broja prioriteta za razvoj region/države;
- Korak 5: Kombinacija politika – Definisane koherentne kombinacije politika, mapa puta i akcionih planova – kombinacija mešovite mere politike i njihova podrška mapama puta ili akcionim planovima za obezbeđivanje sprovođenja;
- Korak 6: Evaluacija i praćenje – Integracija mehanizama za praćenje i evaluaciju – razvoj sistema za kontinuirano praćenje procesa zasnovano na činjenicama/podacima i praćenje rezultata i efekata, kako bi se izvukle pouke i revidirao miks politika.

U ovom radu se posebna pažnja posvećuje šestom koraku, procesu monitoringa i evaluacije (M&E). Problematiku M&E autori su već obrađivali u svojim ranijim radovima posvećenim raznim aspektima strategije pametne specijalizacije, kako u Srbiji, tako i šire u regionu Zapadnog Balkana [11, 12]. 4S (Strategija pametne specijalizacije Srbije; Smart Specialisation Strategy of Serbia – 4S) usvojena je 27.02.2020, a Akcioni plan za njenu realizaciju usvojen je godinu dana kasnije, 15.04.2021. U ovom radu se analiziraju metodološki aspekti praćenja realizacije 4S, poređenjem procedura predloženih u postavci „smart specialisation“ pristupa Evropske unije sa procedurama definisanim u usvojenoj 4S u Srbiji, uz primer pristupa primenjenog u

Hrvatskoj, koji je primeren uslovima realizacije strategije pametne specijalizacije u regionu Zapadnog Balkana.

## 2. M&E STRATEGIJE PAMETNE SPECIJALIZACIJE U EU

Pametna specijalizacija je definisana kao ex-ante uslov za korišćenje Evropskih razvojnih fondova (ERDF) u okviru Tematskog cilja 1 I&I – Istraživanje i inovacije (R&I – Research and Innovation) [European Commission, 2010], [3]. Preko 120 strategija pametne specijalizacije implementirano je tokom programskog perioda 2014-2020, koje su prethodile ulaganjima od preko 40 milijardi evra iz Evropskog fonda za regionalni razvoj (European Regional Development Fund – ERDF), a uključujući nacionalno sufinansiranje uloženo je preko 65 milijardi evra. M&E treba da pruži uvid u efikasnost i efektivnost implementacije politike pametne specijalizacije i trebalo je da izvuče pouke o toj politici Evropske unije za programiranje u narednom periodu, 2021-2027. Integrisani sistem praćenja i evaluacije ciklusa politike EU ima za cilj unapređenje kapaciteta učenja sistema i omogućava bolje prihvatanje i zadovoljavanje potreba širih grupa društva. Stoga, efikasni sistemi praćenja i evaluacije zahtevaju mehanizme koji mogu da mapiraju i angažuju relevantne zainteresovane strane [9].

U citiranoj studiji [9] analiziran je postupak M&E u regionima/državama EU korišćenjem ankete koju su popunjavale institucije nadležne za sprovođenje S3, uz analize mera za sprovođenje S3 i izveštaje o studijama slučaja. Od 120 postojećih strategija pametne specijalizacije, anketu je popunilo 79 nacionalnih ili regionalnih organa za sprovođenje iz devetnaest zemalja, dok studije slučaja pokrivaju trinaest regionalnih i 4 nacionalne strategije i prakse njihove implementacije. Pored ove studije, drugi autori su istraživali još četiri teme: uticaj pametne specijalizacije na upravljanje sistemima politike istraživanja i inovacija [8], proces preduzetničkog otkrivanja [14] i implementacija politike [7]. Glavni nalaz citirane studije M&E implementacije strategija pametne specijalizacije u EU [9] ukazuje na činjenicu da pametna specijalizacija predstavlja **kulturnu promenu** za većinu regiona, bilo da su razvijeni i već dobro upoznati sa praksama regionalne inovacione politike ili manje razvijeni sa nižim intenzitetom inovacionih aktivnosti. Te studije naglašavaju da je neophodno da se formira poseban **tim za M&E** u okviru javne uprave nadležne za implementaciju S3, koji je pri tome potpuno opremljen adekvatnim ljudskim i finansijskim resursima, kako bi se izvršila evaluacija rezultata S3 i efektivnosti logike intervencije u politici pametne specijalizacije. Za potrebe M&E neophodno je prikupljanje podataka o ponašanju inovacionih aktera u posmatranom regionu / državi za šta je, dalje, neophodno koristiti analitičke i informativne alate (big data, veb semantika, itd.) koji mogu da obezbede različite vrste i količine podataka [9].

### 2.1. Izazovi u praćenju i evaluaciji S3

Prvi veliki izazov u praćenju i evaluaciji S3 predstavlja složenost strategije pametne specijalizacije jer je to primer složene strategije inovacija u kojoj koegzistiraju instrumenti izvedeni iz različitih teoretskih logika uz uključivanje velikog broja regionalnih aktera iz četvorostrukog heliksa vlade, biznisa, istraživanja i civilnog društva. Ovo nameće pluralistički kontekst upravljanja u implementaciji strategije pametne specijalizacije. Samim time proces praćenja i evaluacije mora da obuhvati i zadovolji potrebe širih grupa u društvu, na koje politike mogu imati različit uticaj. Ovo zahteva mehanizme koji mogu da mapiraju relevantne zainteresovane strane i da se direktno angažuju sa njima i mere napredak ka novim industrijskim sistemima ka redefinisanim tržištima rada koja su otpornija na društvene, ekonomske i ekološke promene [9].

Sledeći veliki izazov je proces preduzetničkog otkrivanja (The entrepreneurial discovery process – EDP), koji se ne odnosi samo na identifikaciju prioriteta ulaganja u istraživanje i inovacije (prioritetne oblasti) i na istraživanje novih tehnoloških mogućnosti zahvaljujući angažovanju zainteresovanih strana, već je to društveni i politički proces, gde se moraju uzeti u obzir pitanja kao što su moć, interesi različitih grupa, itd. Pored toga, konceptualno, procesi preduzetničkog otkrivanja (EDP) evoluirali su od elementa faze dizajna pametne strategije u kontinuiranu aktivnost. Za sam proces M&E evaluacije ovo podrazumeva prelazak sa tradicionalnih pristupa ka participativnim pristupima u kojima je fokus na podršci kapacitetima za učenje sistema jačanjem povratnih informacija i poboljšanjem pristupa informacijama.

Drugi izazovi za evaluaciju odnose se na neizvesnost uticaja koji proističu iz intervencija i potrebu praćenja konteksta jer se ekonomski i inovacioni sistemi razlikuju u svojoj evoluciji, a zatim tu je i potreba da se radi u realnom vremenu kako bi procenili socio-ekonomske probleme i predložili efikasne strategije za njihovo rešavanje, odnosno, M&E bi trebalo da se odvija istovremeno sa razvojem i implementacijom programa S3 [9].

## 2.2. Kriterijumi za uspostavljanje efikasnog sistema praćenja i evaluacije S3

Zajednički istraživački centar EU u Sevilji ima posebnu ulogu u podršci državama članicama u kreiranju i implementaciji S3, u okviru kojeg je i proces M&E. U okviru te podrške, definisani su i sledeći kriterijumi efikasnog sistema praćenja i evaluacije implementacije S3 [9]:

Sistem M&E implementacije S3 treba da

- definiše ciljeve svake prioritetne oblasti S3, uključujući artikulaciju logike intervencije politike svakog prioriteta pametne specijalizacije;
- kvantifikuje rastojanje između očekivanja i stvarnosti intervencije, uključujući prikupljanje podataka o društveno-ekonomskom uticaju pametne specijalizacije;
- sistematski prikuplja, organizuje i prenosi informacije o razvoju političkih intervencija;
- proizvodi informacije koje podržavaju prilagođavanje i poboljšanje dizajna politike, čime se doprinosi cikličnom procesu učenja koji omogućava razumevanje odnosa između stvarnih i očekivanih rezultata.

Ovaj sistem mora da uspostavi mehanizam koji ima za cilj da potvrdi ispravnost logike političke intervencije i koji ima za cilj da identifikuje i podrži buduća poboljšanja u kreiranju politike i mehanizmima sprovođenja S3.

Kreiranjem mehanizama za funkcionisanje navedenih kriterijuma, uspostavlja se ciklus "Prioritetni ciljevi S3" – "Logika intervencije politike" – "Merenje rezultata" – "Okvir za upravljanje podacima" – "Proizvodnja informacija". Ovaj ciklus se ponavlja u zavisnosti od rezultata M&E na osnovu kojih se redefinišu ciljevi, a zatim i prilagođava S3 novouspostavljenim uslovima funkcionisanja tehno-ekonomskih faktora za implementaciju S3 u regionu/državi.

## 2.3. Nalazi i pouke primene sistema praćenja i evaluacije S3 u EU

Najvažniji nalazi i preporuke JRC u istraživanju postupaka i rezultata procesa praćenja i ocenjivanja primene strategija pametne specijalizacije u regionima / državama EU su sledeći [9]:

- Usvajanje koncepta pametne specijalizacije kao vodećih principa za implementaciju inovacionih strategija predstavljalo je promenu kulture za većinu regiona, bilo da su razvijeni i da su već dobro upoznati sa regionalnim praksama inovacione politike ili manje razvijene sa nižim inovacionim performansama. Međutim, praksa praćenja i evaluacije implementacije S3 i dalje zaostaje. Ovaj nalaz implicira potrebu nove promene kulture u kojoj kreatori politike zahtevaju studije evaluacije ne zato što moraju da se pridržavaju nekih administrativnih zahteva, već zato što su istinski zainteresovani za učenje i posvećeni učenju na rezultatima implementacije S3;
- Prilikom definisanja indikatora za sistem praćenja, neophodno je da oni budu adekvatno povezani sa prioritetima odabranim za strategiju pametne specijalizacije;
- Neophodno je identifikovati namenski tim odgovoran za praćenje i evaluaciju S3 u okviru javne uprave;
- Da bi se podržale aktivnosti evaluacije, važno je prikupiti tačne i blagovremene podatke koji se odnose na ponašanje inovacionih aktera;
- Potreba za blagovremeno dostupnim podacima priznata je u skoro svim studijama slučaja, pa je u narednom programskom periodu neophodna upotreba analitičkih i informativnih alata (big data, veb semantika, itd) koji bi mogli da obezbede različite vrste podataka i bržu reakciju na rezultate M&E;
- Tradicionalni pristupi evaluaciji zasnovani na odgovornosti treba da budu dopunjeni sofisticiranijim pristupima u kojima se praćenje i evaluacija ne posmatraju samo kao integrisani deo ciklusa kreiranja politike, već kao sam instrument koji ima za cilj poboljšanje kapaciteta za učenje sistema i, stoga, pomirenje implikacija sve veće društvene složenosti sa zahtevom za efikasnom intervencijom javne politike. Implikacija na vreme evaluacije je očigledna: evaluacija treba da se odvija istovremeno sa razvojem i implementacijom programa implementacije S3;
- Upotreba evaluacije u ciklusu politike mogla bi se poboljšati ako se uspostave dobri procesi za povratne informacije od evaluatora do kreatora politike, kako tokom programa implementacije S3, tako i na njegovom završetku.

Budući da je u početku (2014) koncept S3 ipak bio eksperiment, njegova primena je takođe podložna evaluaciji i oceni uspešnosti na kraju prvog perioda primene u EU (2014-2020), što je bio predmet mnogih zvaničnih i akademskih analiza [5, 13]. Već sredinom perioda implementacije rađene su analize i evaluacije samog koncepta. Reakcije na ono što se sprovelo ali i naučilo od 2014. verovatno će omogućiti predloge robustnijeg i jednostavnijeg koncepta, čije će posledice kada je u pitanju implementacija biti mnogo očiglednije i lakše što se tiče samih regiona.

Implementacija koncepta je bila predmet analize i jednog od njegovih tvoraca, Dominika Foreja [5]. Forej smatra da nakon masivne primene i neke vrste probe S3 koncepta počev od 2014. godine, osnovni smisao i

opšti principi ostaju nepromenjeni ali da se korisnicima – regionima može ponuditi jednostavniji i očigledniji proces koji će moći lakše da implementiraju. Ovaj pojednostavljeni proces primene S3 u prvom koraku neće zahtevati učešće svih regionalnih aktera, ali će i dalje zahtevati značajnu ulogu statistike i ekonomskog znanja budući da se mora voditi računa o delikatnom balansu između suviše široke i preterano uske definicije prioritarnih oblasti. U ovom koraku očekuje se maksimalni doprinos javnih i privatnih stejkholdera, onih koji će najbolje znati koje je rastuće tržište značajno za njihov region, i koja će istraživanja na najbolji način zadovoljiti potrebe biznisa i društva u datom region. To je sve u svrhu pojednostavljenja procesa a EDP će se veoma brzo materijalizovati u sledećoj fazi.

### 3. PRIMER: M&E U HRVATSKOJ

Svetska banka je za potrebe Vlade Republike Hrvatske organizovala analizu dizajna i realizacije upravljanja implementacijom strategije pametne specijalizacije Hrvatske. Od obimne studije u kojoj su analizirani svi značajni aspekti upravljanja implementacije S3, za potreba ovog rada izdvojeni su glavni nalazi i preporuke eksperata koji se odnose direktno na proces praćenja i ocenjivanja primene S3 [1, 18]:

- Uloga M&E u učenju i prilagođavanju politike tek treba da se uspostavi. Kritični izazov u ovom pogledu bio je tretiranje podataka praćenja kao aktivnog alata za donošenje odluka, a ne kao pasivnog toka informacija. Drugi izazov je vezan za integraciju različitih sistema praćenja što zahteva harmonizaciju okvira rezultata i protokola izveštavanja. Koordinacija je posebno teška za podatke o tematskim prioritetima jer ne postoji zajednički pristup prikupljanju ovih podataka;
- Institucionalni kapaciteti su neujednačeni u celom procesu politike S3. Kapacitet M&E je ograničen, delimično zbog fragmentacije odgovornosti za kreiranje, implementaciju i praćenje politike S3, što ne olakšava otkrivanje i rešavanje uskih grla. Koordinacija između kreiranja politike, implementacije, zajedničkog kreiranja i praćenja i evaluacije je nedovoljno razvijena, posebno zato što je funkcija Nacionalnog saveta za inovacije kao strateškog organa za koordinaciju i donošenje odluka nedovoljno iskorišćena;
- U pogledu upravljanja politikom, neophodno je integrisati aktere upravljanja politikom, implementacijom i procesom otkrivanja preduzetnika tako što će se:
  - uspostaviti S3 jedinica – to mora biti nezavisno telo sa jakim autoritetom nad pojedinačnim ministarstvima koje bi trebalo da olakša koordinaciju različitih faza ciklusa politike, poboljša praćenje i evaluaciju i obezbedi da se strateške odluke Nacionalnog saveta za inovacije sprovede;
  - Jačanje izveštavanja i korišćenja M&E. M&E bi trebalo da služi svrsishodnijoj, strateškoj ulozi u učenju, koja zahteva racionalizaciju i harmonizaciju sistema M&E i jačanje kapaciteta za sprovođenje efektivnog M&E. Jedinica za sprovođenje politike mogla bi da preuzme odgovornost i resurse za sprovođenje M&E. Tada bi mogla da koristi podatke koje prikuplja za zagovaranje prilagođavanja politike i implementacije;
  - Uspostavljanje mehanizma prilagođavanja u realnom vremenu. Tela uključena u upravljanje S3 treba da uspostave mrežu M&E kao strukturisanu platformu za diskusiju koja bi otkrila probleme u procesu otkrivanja preduzetništva i u implementaciji odmah kako se pojave.

### 4. M&E ZA 4S

Autorima ovog rada nije poznato da li je u Republici Srbiji pokrenut proces i uspostavljen sistem praćenja i ocenjivanja primene usvojene strategije pametne specijalizacije (4S). Ono što je javnosti raspoloživo su tekst 4S [16] i Akcioni plan [17]. U tekstu 4S u delu koji se odnosi na M&E (poglavlje VI: *Praćenje sprovođenja i vrednovanje efekata 4S*) navodi se: "*Monitoring i evaluacija su neodvojivi i integrisani deo svake RIS3 i neophodan element za svaku uspešnu strategiju. Shodno tome, Strategija pametne specijalizacije u Republici Srbiji biće praćena odgovornim i realnim akcionim planom za sprovođenje. Akcioni plan (AP) odnosiće se na period od 2020. do 2023. godine i biće usvojen do 90 dana nakon usvajanja Strategije, a podrazumevaće glavno sredstvo za praćenje sprovođenja 4S*". Akcioni plan je usvojen 15. aprila 2021.g., dakle godinu dana po usvajanju 4S.

Dalje, u tekstu 4S navodi se i sledeće: "*Ministarstvo prosvete, nauke i tehnološkog razvoja, kao nosilac Strategije 4S, biće odgovorna institucija i glavno telo zaduženo za praćenje i implementaciju 4S. Imaće podršku interne jedinice, koja će biti osnovana u okviru samog ministarstva.*" Na sajtu nadležnog ministarstva, u organizacionoj strukturi, nije navedeno postojanje takvog tela.

Sledeći deo teksta 4S odnosi se na specificiranje radnog tela nadležnog za proces M&E: "Radno telo MPTNR-a, Republičkog sekretarijata za javne politike (RSJP) i resornih ministarstva za RIS3 bilo bi zaduženo za koordinaciju i praćenje Strategije 4S. Mandat ove grupe trajaće tokom vremena trajanja Strategije i obuhvataće zadatak nadzora nad njenim sprovođenjem, zadržaće sličan sastav članova i time će osigurati nastavak već dobro vođenog procesa i uključivanja svih relevantnih zainteresovanih strana. Takođe, u okviru ove široke grupe, uspostaviće se i nova posebna radna grupa samo za prioritetnu oblast – HRANA ZA BUDUĆNOST. U ovu posebnu radnu grupu biće uključeni stručnjaci koji su učestvovali u procesu EDP-a i određeni predstavnici iz Ministarstva poljoprivrede, šumarstva i vodoprivrede". Na osnovu dostupnih javnih informacija zaključujemo da je ovo telo uspostavljeno, iako je krajem 2021. započeo rad na projektima koji su usklađeni sa prioriternim domenima S4.

Navođenje uspostavljanja sistema za M&E u tekstu 4S zaključuje se sledećim izvodom iz tog dokumenta: "Ovom strategijom predviđen je nastavak EDP procesa, koji će uključiti kontinuiranu komunikaciju i s predstavnicima EDP-a, kako bi se održala uspostavljena transparentnost procesa, ali i obezbedila kvalitetna implementacija Strategije 4S. Tokom sprovođenja Strategije 4S, RSJP (posebna organizacija Vlade Republike Srbije koja pruža stručnu podršku Vladi i organima državne uprave u procesu planiranja, razvoja, donošenja i koordinacije javnih politika i sprovođenja regulatorne reforme) obezbediće tehničku podršku MPNTR-u. Takođe, u toku prve godine primene Strategije 4S, ekspertsku podršku MPNTR-u pružaće i projekat Svetske banke „Konkurentnost i zapošljavanje”. Poslednji mehanizam kontrole biće evaluacije nezavisnog tela, koje će se vršiti prema potrebi. Nakon usvajanja Strategije 4S, već u prvoj polovini 2020. godine, biće lansiran prvi paket mera. Prvi izveštaj o praćenju planiran je početkom 2021. godine. Interim evaluacija sprovedeće eksterno telo 2023. godine, da bi se usaglasio ciklus evaluacije sa zemljama članicama EU". Na osnovu dostupnih javnih informacija zapažamo da nije došlo do realizacije navedenih planiranih aktivnosti.

Do sada je poznato da su programi Fonda za inovacionu delatnost Srbije (Katapult – sredstvima zajma od Svetske banke i bespovratnih sredstava EU, Pametni početak, Program saradnje nauke i privrede - iz Budžeta RS, Program za sufinansiranje inovacija, Program ranog razvoja – iz Budžeta RS) usklađeni u pogledu prioriternih domena sa S4. Projekti finansirani kroz ove programi su počeli da se realizuju u drugoj polovini 2021. i početkom 2022. godine, a finansijska sredstva su obezbeđena iz budžeta Republike Srbije i uz pomoć donatora (EU, Svetska banka, USAID, UNDP), s obzirom na to da su aktivnosti iz akcionog plana usaglašena sa donatorskim aktivnostima (<https://pametnaspecijalizacija.mpn.gov.rs/najava-dogadjaja/>).

## 5. ZAKLJUČAK

Analiza organizacije sistema i realizacije procesa praćenja i ocenjivanja primene strategije pametne specijalizacije u regionima / državama EU, uz istaknut primer takve analize na slučaju M&E strategije pametne specijalizacije u Republici Hrvatskoj, ukazali su na kompleksnost, neophodnost i značaj sistema i procesa M&E. Primena koncepta S3 još uvek je u razvojnoj fazi i u regionima / državama EU, iako ima ključnu ulogu u raspodeli sredstava razvojnih fondova EU (ERDF). Analiza implementacije S3 u Hrvatskoj ukazala je takođe na neophodnost formiranje jedinice za sprovođenje politike kao nezavisnog ekspertskog tima.

Situacija u Srbiji može da se oceni sa dva aspekta: Prvo, postavka sistema M&E u tekstu 4S odgovara teoretskim načelima koncepta strategije pametne specijalizacije. Naglašen je i značaj i definisani su mehanizmi i institucije nadležne za proces praćenja i ocenjivanja primene 4S; Drugo, proces i funkcionisanje praćenja i ocenjivanja primene 4S prema dostupnim informacijama nisu uspostavljeni, budući da javno nisu raspoloživa dokumenta koja se odnose na taj proces. Pošto je konceptom S3 definisana i ultimativno zahtevana javnost u svim koracima definisanja i primene S3, otuda se nameće i potreba za javnošću podataka i informacija o procesu praćenja i ocenjivanja primene 4S. Uspostavljanje sistema praćenja i evaluacije sprovođenja aktivnosti S4 naročito je važno kako za nosioce S4 i davaoce sredstava namenjenih S4 (te začuđuje njihova indiferentnost) tako i za istraživačke oblasti koje se bave podrškom kreiranja politika i strategija razvoja inovacionog ekosistema.

## LITERATURA

- [1] Aralica, Z., Božić, Lj. (2021). Monitoring of the innovation programmes in Croatia, *Tematski zbornik radova Tehnologija, kultura i razvoj* 28, pp. 4 - 21, ISBN: 978-86-915151-6-4 (UTD), 2021. <http://www.pupin.rs/cirnt/wp-content/uploads/2022/05/Zbornik-TKR28-f.pdf>

- [2] European Commission (2010). Regional Policy Contributing to Smart Growth in Europe 2020, COM (2010) 553 final
- [3] European Commission (2011). Europe 2020 Flagship Initiative Innovation Union SEC(2010) 1161, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, European Union, 2011
- [4] Foray, D. (2019). In response to ‘Six critical questions about smart specialisation’, European Planning Studies, DOI:10.1080/09654313.2019.1664037, <https://doi.org/10.1080/09654313.2019.1664037>
- [5] Foray D. & Ortega-Argilés R. (2011). Smart Specialisation. From Academic Idea to Political Instrument, the Surprising Career of a Concept and the Difficulties in its Implementation, MTEI WP 2011-001
- [6] Foray, D., David, P.A., and Hall, B. (2009). Smart Specialisation - The Concept, Knowledge Economists Policy Brief no.9, Knowledge for Growth Expert Group
- [7] Gianelle, C., Guzzo, F. and Fratesi, U. (2021). Lessons from the Smart Specialisation experience: Policy Implementation, Smart Specialisation – JRC Policy Insights, JRC124039.
- [8] Guzzo, F., and Gianelle, C. (2021). The impact of smart specialisation on the governance of research and innovation policy systems, JRC Policy Insights, JRC124072.
- [9] Hegyi, F., B., and Prota, F. (2021). Assessing Smart Specialisation: Monitoring and Evaluation Systems, EUR 30654 EN, Publications Office of the European Union, Luxembourg, 2021. ISBN 978-92-76-32592-5 , doi:10.2760/443642, JRC123734  
[http://ec.europa.eu/regional\\_policy/sources/docgener/presenta/smart\\_specialisation/smart\\_ris3\\_2012](http://ec.europa.eu/regional_policy/sources/docgener/presenta/smart_specialisation/smart_ris3_2012).  
<https://doi.org/10.1080/09654313.2016.1181717>
- [10] Kleibrink, A., Gianelle, C., & Mathieu Doussineau (2016). “Monitoring innovation and territorial development in Europe: emergent strategic management”, European Planning Studies, 24:8, 1438-1458, DOI: 10.1080/09654313.2016.1181717, Published online: 31 May 2016,
- [11] Kutlača, Đ., Semenčenko, D. (2021). "Razvoj zasnovan na znanju i strategije pametne specijalizacije zemalja Zapadnog Balkana", *Tematski zbornik radova Tehnologija, kultura i razvoj 28*, pp. 40 - 61, ISBN: 978-86-915151-6-4 (UTD), 2021. <http://www.pupin.rs/cirnt/wp-content/uploads/2022/05/Zbornik-TKR28-f.pdf>
- [12] Kutlača, Đ., Semenčenko, D., Živković, L. (2021). "Praćenje realizacije 4S – metodologija i prvi rezultati", *Tematski zbornik radova Tehnologija, kultura i razvoj 28*, , pp. 62 - 70, ISBN: 978-86-915151-6-4 (UTD), 2021.
- [13] McCann, Philpi, Ortega-Argilés, Raquel. (2016). The early experience of smart specialization implementation in EU cohesion policy, European Planning Studies, 24:8, 1407-1427, DOI: 10.1080/09654313.2016.1166177
- [14] Perianez-Forte, I. and Wilson, J. (2021). Assessing Smart Specialisation: EDP, JRC Policy Insights, JRC124101
- [15] S3 Platform (2012). Guide to Research and Innovation Strategies for Smart Specialisation (RIS3), May 2012,
- [16] Vlada Republike Srbije (2020). Strategija pametne specijalizacije u Republici Srbiji za period od 2020. do 2027. godine, "Službeni glasnik RS", br. 30/18, odluka Vlade 05 Broj: 153-1851/2020-1, 27. februara 2020. godine.
- [17] Vlada Republike Srbije (2021). Akcioni plan za sprovođenje Strategije pametne specijalizacije za period od 2021-2022 godine, "Službeni glasnik RS", broj 42/2021 od 27.04.2021, odluka Vlade od 15. aprila 2021. godine.
- [18] World Bank. (2021). “Analysis of Design and Implementation of Croatian S3 Governance”, World Bank Group, Croatia Country Office, April 2021

## ZAHVALNOST

Istraživanje opisano u ovom radu finansirano je od strane Ministarstva prosvete, nauke i tehnološkog razvoja Srbije.



## THE ROLE OF SMART SPECIALISATION STRATEGY IN ECONOMIC RECOVERY AND SUSTAINABLE DEVELOPMENT IN THE POST COVID-19 ERA

DIJANA ŠTRBAC<sup>1</sup>, NIKOLA VASILIC<sup>2</sup>

<sup>1</sup> University of Belgrade – Institute Mihajlo Pupin, dijana.strbac@pupin.rs

<sup>2</sup> University of Belgrade – Institute Mihajlo Pupin, nikola.vasilic@pupin.rs

**Abstract:** *Smart specialisation is an innovation policy concept which aims to develop competitive advantages of regions/countries based on their innovation potential and applying evidence-based analysis. The objective of this paper is to analyse the potential of smart specialisation strategies (S3) to support economic recovery after the crisis caused by the Covid-19 pandemic and support sustainability imposed as one of the most important pillars of the new EU growth strategy. The results of the conducted analysis indicate that: 1) smart specialisation strategies are economic transformation agendas based on identifying key national/regional priorities, 2) selection of S3 priorities is based on the analysis of innovative, scientific and economic potential as well as relevant societal challenges, 3) S3 is highly accepted approach not only in the EU and enlargement countries, but also worldwide, 4) S3 is an effective tool for the economic recovery after the Covid-19 crisis and achieving sustainable development goals.*

**Keywords:** *Smart Specialisation strategy (S3), economic recovery, sustainability, sustainable development, Covid-19.*

### 1. INTRODUCTION

Smart specialisation is a unique innovation policy experiment launched by the European Commission in 2010. It relies on the idea that each region or country has distinct characteristics, knowledge base and capacities which can be the source of their competitive advantage if the efforts are concentrated on a limited number of economic activities. Therefore, the goal is to link research & innovation potential and economic development by setting priorities with close involvement of various relevant stakeholders.

Smart specialisation strategies have a vital role in stimulating R&I and industrial transition of regions and countries. They are also an effective tool to help territories recover from the current pandemic-induced crisis. Moreover, S3 is recognised as a concept which can be successfully aligned with the European Green Deal and UN 2030 Agenda. The current development of S3 approach is based on introducing sustainability elements in its core phases.

This paper is organized as follows. Section 2 introduces the concept and origins of the smart specialisation strategy. The third section describes application of S3 approach in the EU and worldwide with emphasis on the Republic of Serbia as an example of a non-EU country involved in RIS3 design and implementation. The fourth section examines the potential of S3 in developing economic recovery paths after the economic crisis caused by the Covid-19 pandemic. Section 5 describes the need and possibilities for transformation from S3 to Smart Specialisation strategies for sustainability (S4). Finally, Section 6 presents main conclusions.

### 2. THE ORIGINS AND CONCEPT OF SMART SPECIALISATION

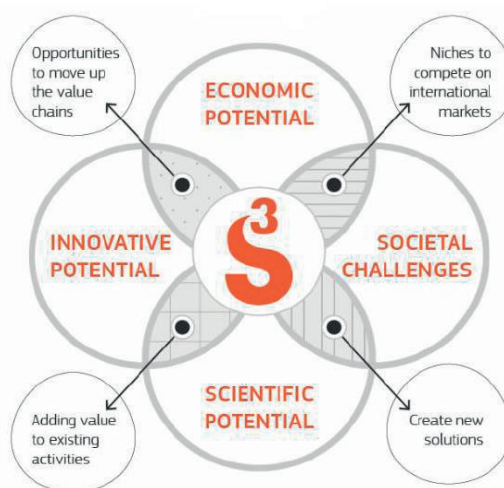
Over the last decade the Research and Innovation Strategy for Smart Specialisation (RIS3) has been considered as a policy framework which plays a central role in supporting innovative activities of regions and countries. The smart specialisation approach is incorporated in European Cohesion Policy and represents a precondition for receiving financial support from the European Structural and Investment Funds (ESIF).

The term “smart specialisation“ was formulated within the “Knowledge for Growth” expert group established by the former European Commissioner for Research, Janez Potočnik. In the context of the need to revitalise the Lisbon Strategy, in 2005 was established a group of prominent economists in order to provide him with high-level advice related to: contribution of knowledge to growth and prosperity; policy mix needed to promote creation, dissemination and use of knowledge and role of different actors in stimulating knowledge society [1]. Initially, the term was related to the need of concentration and specialisation of research and innovation with the aim to enhance entrepreneurial and innovation processes [2]. The concept of smart specialisation was first introduced in a policy brief prepared by Foray and Van Ark, as members of “Knowledge for Growth” expert group. The policy brief was published in 2007 and was devoted to developing the strategies

for attracting more R&D investments to Europe. The main conclusion was that there is a need to move towards the R&D system based on greater European-wide specialisation and that the European Research Area will only benefit those countries and regions that have clear visions and strategies as to how they can develop distinctive, original, modern areas of specialisation for the future [3].

The concept was further developed in 2009 within the expert group “Knowledge for growth“ with the aim to promote development strategies focused on investments in programs that complement existing production capacities and create future comparative advantages. It was clearly pointed out that the idea of smart specialisation approach does not include imposing top-down industrial policy or a foresight exercise, it is based on the entrepreneurial discovery process which reveals research and innovation areas in which certain country or region is likely to excel in the future [4].

The Guide for RIS3 was published in 2012 in order to offer the policy-makers and implementing bodies “methodological guidance on how to prepare for and how to design, draft and implement a national/regional research and innovation strategy for smart specialisation“ [5]. In this Guide, RIS3 is considered as a transformation agenda which brings to: 1) focusing policy support on key national/regional priorities and needs for knowledge-based development, 2) building competitive advantages of the country/region, 3) supporting technological and practice-based innovation with the aim to stimulate private sector investment, 4) full involvement of stakeholders, 5) evidence-based policy making with sound monitoring and evaluation system [5]. Therefore, smart specialisation strategies enable upgrading economic structure of countries and regions on the basis of their existing research and innovation potential. The focus is on development of competitive advantages on the basis of identified innovation niches. Therefore, smart specialisation approach includes unification of economic, innovative and scientific potential of a country/region with the aim to address societal challenges as well (Figure1).



**Figure 1:** Smart specialisation concept (Source: [6])

According to the official RIS3 guide, there are six steps in smart specialisation strategy development: regional analysis, governance structure, shared vision about the future, priority setting, policy mix and monitoring and evaluation mechanisms [5]. Although the S3 is often considered as a bottom-up approach, its authors describe it as a combination of two policy logics – a planning logic and a self-discovery logic [7]. In other words, the S3 approach is initiated and implemented on the basis of strategic decisions, but the process is marked with the interactions and initiatives of the actors participating in the strategy design, implementation, monitoring and evaluation.

### 3. INTERNATIONAL EXPERIENCES IN RIS3 IMPLEMENTATION

RIS3 as a place-based approach to innovation has initiated the development of more than 100 smart specialisation strategies in EU countries and regions. In total, 19 EU member states and 7 non-EU countries as well as 180 EU and 42 non-EU regions have registered to the S3 Platform of the Joint Research Centre [8]. The platform is specialised in offering advice to the countries and regions in RIS3 design and implementation. It also provides support in inter-territorial cooperation by supporting S3 thematic platforms and EU macro-regional strategies. The platform includes different mapping tools such as Eye@RIS3, Digital Innovation Hub

and Regional Benchmarking. These tools enable identification of specialisation domains for regions and countries and therefore foster collaboration among various actors in Europe.

In recent years, there has been a growing interest to implement smart specialisation approach in transforming innovation policies. The S3 has started to inspire different countries and regions belonging to the EU Enlargement policy, the EU Neighbourhood Policy and more widely, Australia, Brazil, Chile, Colombia, Mexico, Norway, Peru, China, Thailand, the United States of America, Canada and Africa, among others [6]. Therefore, the policy makers outside of the European continent have recognised the added value of the S3 approach in fostering economic transformation and building regional innovation ecosystems.

### 3.1. RIS3 in the Republic of Serbia

Building on the experiences of the EU countries, since 2016, the Joint Research Centre has been working with the countries under the EU Enlargement and Neighbourhood policy in order to provide support in development of smart specialisation strategies.

The Smart Specialisation process in Serbia started in 2017. It was initiated by the Ministry of Education, Science and Technological Development taking charge of the coordination of the process by establishing an Inter-ministerial working group for the development of the RIS3. The process was conducted with strong support of the JRC and in accordance with the „Smart specialisation framework for Enlargement and Neighbourhood countries“<sup>1</sup>.

The process of RIS3 design in Serbia started with a detailed quantitative analysis which included mapping of economic, innovation and research potential of the Republic of Serbia. On the basis of the identified areas within the framework of the quantitative analysis, in December 2017, an ICT Innovation Camp (software engineering) was held. In 2018, after a series of interviews with relevant stakeholders, a Qualitative Analysis was conducted, which identified potential priorities - areas with competitive advantage. In March 2019, the entrepreneurial discovery process (EDP) began and lasted until June 2019. During this period, 17 workshops were held, divided in three thematic areas: 1) SWOT (identification of strengths, weaknesses, opportunities and threats within the priority areas), 2) Vision (defining the priority area vision and defining the objectives to be achieved in the next 3-5 years), 3) Policy mix (selecting measures and activities for achieving of the identified objectives) [9].

Smart Specialisation Strategy of the Republic of Serbia (4S) was adopted in February 2020. On the basis of the quantitative and qualitative analyses, as well as EDP, the following priority areas in the Republic of Serbia are selected: Food for future, ICT, Future machines and manufacturing systems, Creative industries. The 4S is followed by the Action plan for the implementation of the Smart Specialisation Strategy in the Republic of Serbia for period 2021-2022.

Serbia is one of the first countries that has identified complementarities between RIS3 and sustainable development goals (SDGs). It is one of five countries in the Global Pilot Programme on STI for SDGs Roadmaps for SDGs from the United Nation's Inter-Agency Task Team on STI for the SDGs. The JRC has supported Serbia in this programme in developing a pilot methodology on adapting S3 to better contribute to SDGs.

## 4. RIS3 AND ECONOMIC RECOVERY AFTER COVID-19 PANDEMIC

The Covid-19 pandemic has created challenges for the health systems and citizen well-being, but also for economic activity on a global level. The spread of the coronavirus and the measures undertaken to protect population have caused different shocks on the supply and demand sides of the economic systems and therefore caused global economic crisis.

The crisis has also caused significant changes in research and innovation activity and its policies. Factors shaping the future of research and innovation are related to the uneven effect of the crisis across different R&D sectors, increased application of digital technologies and artificial intelligence, development of new forms of cooperation related to open science, transnational collaboration and public-private partnerships.

Although R&D expenditures are quite procyclical, the pandemic crisis did not affect decline in total R&D investment. On the contrary, research and innovation have become the central part of the policy response to the crisis. According to the OECD data, the OECD economies have increased investment in R&D in 2020 despite the strong decrease of the economic activities. The R&D expenditure in the OECD area grew by 1.8% in real terms in 2020. The growth of R&D investment in the OECD area in 2020 was primarily driven by growth in the United States at 5%, in contrast with R&D expenditures in Germany and Japan, which declined

<sup>1</sup> Elaborated in: Matusiak M. & Kleibrink A. (ed.) (2018). Supporting an Innovation Agenda for the Western Balkans: Tools and Methodologies, Publications Office of the European Union, Luxembourg (Annex 1).

at -5.3% and -2.7% respectively. In the EU27 area, business R&D performance was the main source of the aggregate fall in R&D. The reason for the low performance of the EU lies in the structure of its business R&D. During the crisis, the largest business R&D investments were in ICT and pharmaceutical industry, while automotive, aerospace and other industries have experienced large decline [10].

The academic literature recognises the following two stages in the response to the economic crisis caused by the pandemic: 1) resistance stage focused on macroeconomic and regulatory policies and 2) reconstruction and renewal stage which includes design of proactive policies to build resilience of the regions [11]. The research and innovation and its related policies, especially smart specialisation are beneficial for both phases since they offer support in the recovery of production process and value chains as well as in economic transformation in accordance with the regional advantages. Moreover, by focusing on a place-based approach (i.e. regional specificities and needs), smart specialisation strategies can provide more targeted support to the socioeconomic recovery. Although the pandemic has caused global economic decline, the scope of the crisis depends on the structure of the economy, policy measures implemented and general resilience of the innovation ecosystems. Therefore, RIS3 offers this urgently needed local recovery path.

The potential contribution of the RIS3 also lies in a unique character of the Covid-19 crisis. The crisis outbreak was in the health sector but it had spilled over on almost all aspects of business and lives. In the economic sector, the crisis has created new market trends and niches and innovation can play an important role to help companies in adapting to it and benefiting from it [12].

Since innovation is an interactive process, embedded in social, institutional and cultural framework, the full potential of smart specialisation can be achieved by (re)activating regional innovation ecosystems. The future smart specialisation strategies will have to focus on different perspectives of globalisation such as the new supply chains and developing sustainable and resilient economies. This includes the following:

- Creating a new sense of social solidarity and engaging different stakeholders in Entrepreneurial Discovery Process that can identify competitive advantages of the region and position the region in European value chains.
- Development of efficient innovation policy instruments to support the structural transformation of the economy at the regional/national level and on learning capacity.
- Transnational collaboration and search for “smart complementarities“ supported by the transnational cooperation programmes and strategic partnerships [13].

Smart specialisation strategies are recognised not only as support to the economic recovery, but also as a cornerstone of the needed green transition of the EU. The pandemic has started in the period when the EU has introduced a new growth strategy, the European Green Deal which is based on strong support to the research and innovation in order to tackle the environmental challenges.

## 5. THE EVOLUTION FROM S3 TO S4

The European Green Deal (EGD) was presented in 2019 as a new main growth strategy of the EU that will bring EU economy and society to a more sustainable path and enable that the EU becomes the first climate neutral continent by 2050. It also aims to ensure a strong contribution to the United Nation’s 2030 Agenda and the sustainable development goals (SDGs) [14]. On the other hand, smart specialisation is a bottom-up approach focused on place-based innovation and entrepreneurial discovery process. Therefore, the S3 can become a crucial path for achieving the SDGs and goals of European Green Deal by linking bottom-up processes with the top-down development agendas. In relation to this, the European Commission is working on transformation of S3 to Smart Specialisation Strategies for Sustainability (S4).

In order to shift from S3 to S4 and to meet the objectives of the UN 2030 Agenda and the EGD, the following aspects should be taken into consideration:

- focusing S3 on innovation and industry policy through the lens of a transitions and systems approach,
- formulating a holistic approach towards S4 development, combining a bottom-up approach (Entrepreneurial Discovery Process) with top-down guidance (missions and grand societal challenges),
- targeting policy coherence and coordination (within administrative decision authority and cross-borders),
- integrating EGD/SDGs goals and objectives into the S4 design and implementation,
- considering the three principles of sufficiency, resilience and efficiency,
- evaluating capacities, competences and available resources needed for effective implementation,
- considering environmental impacts within and across regional and national borders,
- going beyond ex-post reporting or checklist exercises on related SDGs through ex-ante planning,
- considering opportunities and the possible impacts of digitalisation [15].

New approach to smart specialisation strategies includes linking it with mission-oriented policies for sustainable development. Therefore, innovation should not be based only on aspirations for competitiveness but also on societal and environmental challenges in order to be “intermediate step towards the longer-term goals of fostering sustainability and inclusiveness” [16]. In this respect, it is recommended transformation of S3 to smart specialisation strategies for sustainable and inclusive growth (S4+). Table 1 presents intervention logic to support evolution from S3 to S4.

**Table 1:** Intervention logic for transformation from S3 to S4

Intervention logic	S3	S4
<i>SWOT analysis</i>	Appraise endowments in assets Assess innovation potential in a territory Appraise entrepreneurial base and dynamics Identify international networks and value chains	Strengthen S3, and ... S4: Position the SWOT analysis in the ecological and digital transitions of the energy, manufacturing, agri-food, housing, and mobility systems.
<i>Governance</i>	Management structure in place Participation of stakeholders in quadruple helix Institutional and human resources capacity	Strengthen S3, and ... S4: Role of the state goes beyond being facilitator and catalyst to co-create system transformation. This requires management reforms and capacity building to work cross-domains, cross-departments, cross-sectors and cross-disciplines.
<i>Vision</i>	Shared vision on present and future innovation challenges Strategy medium-term	Strengthen S3, and ... S4: Vision goes beyond the R&I system. Could be a societal vision or an industrial vision. The vision should be linked to overall target(s) set in time which is understandable, measurable, ambitious but realistic.
<i>Prioritisation</i>	Revision of previous priorities Identify areas of competitive advantage Verify critical mass of budget for achieving each priority	Strengthen S3, and ... S4: If the priorities are aligned or in the same overall direction as the overall EU-level investments, then the potential of reaching critical mass and of crowding-in of private investment and of EU funds increases.
<i>Implementation</i>	Broad definition of innovation Balance between focused and horizontal measures Upgrading existing industry using KETs and digital Experimentation in pilot actions Innovation ecosystems International collaboration search for value chains	Strengthen S3, and ... S4: The Implementation is driven by innovation but mobilises in synergy with other policy areas and investments, such as infrastructure, skills, etc. S4: The local framework conditions to innovation are now also European. When local entrepreneurs detect barriers to innovation for sustainability this can be channelled in “Green Deals” to the national and EU policy level. S4: Cooperation and mutual learning with other actors of change is facilitated by regional thematic network but the new network externalities are broader.

Source: [16].

## 6. CONCLUSION

Smart specialisation approach was developed a decade ago with the aim to identify priority areas for intervention based on a deep exploration of scientific, innovative and economic potential as well as entrepreneurial discovery process with wide involvement of relevant stakeholders. This innovation policy approach had a great success not only in the EU, but also in many other countries and regions around the world.

Smart specialisation can also be an effective tool to recover from the crisis caused by the Covid-19 pandemics in terms of formulating resistance and reconstruction measures. S3 offers support in the recovery of production process and value chains as well as in economic transformation in accordance with the regional

advantages. The S3 approach is place-based so it can offer targeted support on a regional level which is especially important since the Covid-19 crisis had uneven effect on countries and regions.

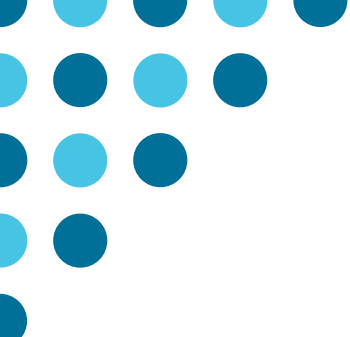
Smart specialisation strategies are place-based and innovation-led transformation agendas for achieving not only economic growth, but also sustainability and inclusiveness. Innovation are now considered as a useful tool for addressing environmental and societal challenges by creating new competitiveness foundations based on green technologies and digitalisation. Due to its bottom-up character, the S3 can become a crucial path for achieving the UN SDGs and goals of European Green Deal by linking bottom-up processes with the top-down development agendas.

### Acknowledgement

The research presented in this paper was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia

### REFERENCES

- [1] Expert Group "Knowledge for growth", Retrieved from: [https://ec.europa.eu/invest-in-research/monitoring/knowledge\\_en.htm](https://ec.europa.eu/invest-in-research/monitoring/knowledge_en.htm)
- [2] Gianelle, C., Kyriakou, D., McCann, P. & Morgan, K. (2020). Smart Specialisation on the move: reflections on six years of implementation and prospects for the future, *Regional Studies*, 54:10, 1323-1327, DOI: 10.1080/00343404.2020.1817364
- [3] Foray, D. & Van Ark, B. (2007). Smart specialisation in a truly integrated research area is the key to attracting more R&D to Europe, *Knowledge Economists Policy Brief*, No.1.
- [4] Foray, D., David, P. A., & Hall, B. (2009). Smart Specialisation – The concept (Knowledge for Growth Expert Group Policy Brief No. 9, June). European Commission.
- [5] Foray, D., Goddard, J., Beldarrain, X. G., Landabaso, M., McCann, P., Morgan, K., Nauwelaers, C., & Ortega-Argilés, R. (2012). Guide to Research and Innovation Strategies for Smart Specialization (RIS3). Publications Office of the European Union.
- [6] Gómez Prieto, J., Demblans, A. & Palazuelos Martínez, M. (2019). Smart Specialisation in the world, an EU policy approach helping to discover innovation globally, EUR 29773 EN, Publications Office of the European Union, Luxembourg, ISBN 978- 92-76-08393-1, doi:10.2760/962643, JRC117005.
- [7] Foray, D., Eichler, M. & Keller, M. (2021). Smart specialization strategies—insights gained from a unique European policy experiment on innovation and industrial policy design. *Review of Evolutionary Political Economy*, vol. 2, issue 1, 83-103.
- [8] Smart specialisation platform, <https://s3platform.jrc.ec.europa.eu/where-we-are>
- [9] Smart Specialisation Strategy of the Republic of Serbia for the period 2020-2027, Retrieved from: [https://pametnaspecijalizacija.mpn.gov.rs/wp-content/uploads/2021/06/Strategija-pametne-specijalizacije\\_EN\\_WEB.pdf](https://pametnaspecijalizacija.mpn.gov.rs/wp-content/uploads/2021/06/Strategija-pametne-specijalizacije_EN_WEB.pdf)
- [10] OECD (2022). OECD Main Science and Technology Indicators Highlights - March 2022. Retrieved from: <https://www.oecd.org/sti/msti-highlights-march-2022.pdf>
- [11] Wilson, J., Aranguren, M. J., Canto, P., Estensoro, M., Fernández, J., Franco, S., de San Vicente, I. G., Kamp, B., Larrea, M., Magro, E., Navarro, M., Rodríguez, A. (2020). Socioeconomic impacts of COVID-19: Reflections from the Basque Country, Basque Institute of Competitiveness-Deusto Foundation.
- [12] Marques Santos, A. (2021). Linking the ‘Recovery and Resilience Plan’ and Smart Specialisation. The Portuguese Case. JRC Working Papers on Territorial Modelling and Analysis No. 05/2021, European Commission, Seville, JRC126178.
- [13] Tuffs, R., Larosse, J., & Corpakis, D. (2020). Post-Covid-19 Recovery Policies: Place-based and Sustainable Strategies. *Symphonya. Emerging Issues in Management*, (2), 55-62.
- [14] European Commission, 2020, “The European Green Deal”, Communication from the Commission to the European Parliament, The European Council, The Council, The European Economic and Social Committee and the Committee of the Regions, COM(2019) 640 final, Brussels.
- [15] Nakicenovic, N., Zimm, C., Matusiak, M. & Ciampi Stancova, K. (2021). Smart Specialisation, Sustainable Development Goals and environmental commons. Conceptual framework in the context of EU policy, EUR 30882 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/766406, JRC126651
- [16] McCann, P. & Soete, L. (2020). Place-based innovation for sustainability, Publications Office of the European Union, Luxembourg, doi:10.2760/250023, JRC121271.



# KOMBINATORNA OPTIMIZACIJA

## COMBINATORIAL OPTIMIZATION







# INVERZNI NASUMIČNO-POHLEPNI ALGORITAM ZA REŠAVANJE PROBLEMA MINIMALNOG DOMINANTNOG SKUPA SA POZITIVNIM UTICAJEM U SOCIJALNOJ MREŽI A REVERSE RANDOMIZED GREEDY ALGORITHM FOR THE MINIMUM POSITIVE INFLUENCE DOMINATING SET PROBLEM IN SOCIAL NETWORKS

KRISTINA KOSTIĆ<sup>1</sup>, ZORICA STANIMIROVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Matematički fakultet, {kristina\_kostic, zoricast}@matf.bg.ac.rs

**Rezime:** U ovom radu je predložen heuristički pristup za rešavanje problema određivanja dominantnog skupa sa pozitivnim uticajem minimalne kardinalnosti (engl. Minimum Positive Influence Dominating Set - MPIDS) sa primenom u analizi socijalnih mreža. Za datu socijalnu mrežu predstavljenu grafom cilj je pronaći što manji skup uticajnih individua (čvorova) koji bi raširio pozitivni uticaj u čitavoj mreži. Obzirom na ubrzani porast socijalnih mreža i primat koji su preuzele u ljudskoj komunikaciji, javlja se potreba da se ovaj vid veze iskoristi na što bolji način. Kako je razmatrani problem NP-težak, u ovom radu je implementiran inverzni nasumično-pohlepni algoritam (engl. Reverse Randomized Greedy Algorithm - RRG), kao i višestartna metoda zasnovana na RRG algoritmu. Predloženi algoritmi su testirani na skupu realnih instanci iz literature i dobijeni rezultati su analizirani i upoređeni sa rezultatima postojećih pohlepnih algoritama za MPIDS problem.

**Ključne reči:** socijalne mreže, dominantni skup, pohlepni algoritam, heuristike

**Abstract:** In this paper we propose heuristic approach to finding a minimum positive influence dominating set (MPIDS) with application in social network analysis. For a given social network represented by a graph, the goal is to find a minimal set of influential individuals (nodes) that allows for spreading positive influence throughout the whole network. Given the fast growth of social networks and the importance they have in modern human communication, these connections should be used in the best possible way. As the considered problem is NP-hard problem, this paper proposes a reverse randomized greedy algorithm (RRG) and a multi-start method based on the RRG algorithm as solution approaches. The proposed algorithms are tested on a set of real-world test instances from literature and the obtained results are analysed and compared with the results of existing greedy algorithms for solving the MPIDS problem.

**Keywords:** social networks, dominating set, greedy algorithm, heuristics

## 1. UVOD

Sa porastom i ubrzanim razvojem socijalnih mreža problem nalaženja minimalnog dominantnog skupa sa pozitivnim uticajem (MPIDS) postaje jedan od popularnijih i značajnijih problema današnjice koji može pozitivno uticati na rešavanje raznih socioloških problema, kao i na globalni ekonomski razvoj. Problem MPIDS je uveo Wang početkom 21. veka sa ciljem rešavanja problema alkoholizma u jednoj relativno maloj adolescentskoj zajednici [5]. Socijalna mreža je opisana grafom u kom čvorovi grafa predstavljaju individue, dok se grane grafa odnose na veze i interakcije između njih. U osnovnoj varijanti problema, reč je o netežinskom grafu, iako je moguće dodati težine koje bi simulirale jačinu veze ili učestalost interagovanja između dve individue. Budući da se radi o NP-teškom problemu kombinatorne optimizacije, potrebno je naći brz i efikasan algoritam koji je u stanju da u kratkom vremenskom periodu pronađe rešenje zadovoljavajućeg kvaliteta na grafu koji ima veliki broj čvorova. U literaturi MPIDS problem je uglavnom rešavan korišćenjem pohlepnih algoritama sa akcentom na razvoju adekvatnog pohlepnog kriterijuma za odabir čvora (individue) koja postaje deo dominantnog skupa sa pozitivnim uticajem [1],[3]. Danas, kada je moguće lako dopreti do većeg dela populacije putem društvenih mreža, sve više raste potreba za razvojem efikasnih algoritama za rešavanje ovog problema.

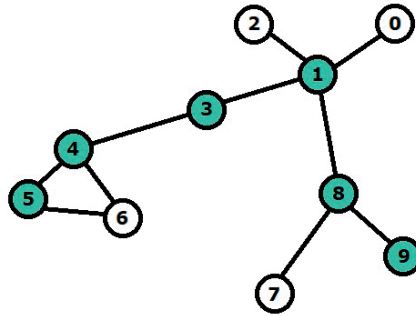
## 2. MATEMATIČKA FORMULACIJA MPIDS PROBLEMA

Neka je  $G = (V, E)$  prost neusmeren povezan graf sa skupom  $V$  od  $n$  čvorova i skupom grana  $E \subset V \times V$ . Za proizvoljna dva čvora  $u, v \in V$  kažemo da su susedi ukoliko postoji grana  $(u, v) \in E$  koja ih povezuje. Otvorena okolina čvora  $v \in V$ , u oznaci  $N(v)$ , definisana je sa  $N(v) = \{u \in V : (v, u) \in E\}$ . Broj čvorova susednih čvoru  $v$

predstavlja njegov stepen,  $deg(v) = |N(v)|$ . Podskup čvorova  $S \subseteq V$  je dominantan skup sa pozitivnim uticajem (engl. *Positive Influence Dominating Set - PIDS*) ukoliko svaki čvor  $v \in V$  ima bar polovinu svojih suseda u skupu  $S$ , preciznije:

$$S \subseteq V \text{ je PIDS} \Leftrightarrow |N(v) \cap S| \geq \left\lceil \frac{deg(v)}{2} \right\rceil, \forall v \in V. \quad (1)$$

Na slici 1 dat je primer dominantnog skupa sa pozitivnim uticajem na grafu od 10 čvorova. Čvorovi koji čine PIDS obojeni su zelenom bojom.



**Slika 1** Primer PIDS skupa na grafu od 10 čvorova

Na zadatom grafu koji modeluje određenu socijalnu zajednicu zadatak je naći dominantan skup sa pozitivnim uticajem minimalne kardinalnosti. Matematička formulacija MPIDS problema koristi sledeći skup binarnih promenljivih:

$$x_i = \begin{cases} 1, & \text{ukoliko čvor } i \text{ pripada PIDS skupu } S \\ 0, & \text{inače} \end{cases}$$

Koristeći prethodnu notaciju, problem nalaženja minimalnog dominantnog skupa sa pozitivnim uticajem može biti formulisan u vidu celobrojnog linearnog matematičkog programa:

$$\min \sum_{i=1}^n x_i, \quad (2)$$

$$\text{pri uslovima: } \sum_{v_j \in N(v_i)} x_j \geq \left\lceil \frac{deg(v_i)}{2} \right\rceil, \quad \forall v_i \in V, \quad (3)$$

$$x_i \in \{0, 1\}, \quad i = 1, \dots, n. \quad (4)$$

Funkcija cilja (2) minimizuje broj čvorova u PIDS skupu što omogućava nalaženje skupa minimalne kardinalnosti, dok je ograničenjima (3) obezbeđeno da svaki čvor iz  $V$  ima bar polovinu svojih suseda unutar PIDS skupa. Ograničenja (4) definišu tip korišćenih promenljivih.

### 3. PREDLOŽENI ALGORITMI ZA REŠAVANJE MPIDS PROBLEMA

Pri konstrukciji pohlepnih algoritama od presudnog značaja za efikasnost i kvalitet algoritma je odabir pohlepnog kriterijuma koji se koristi pri konstrukciji rešenja. Osnovna karakteristika uticajne individue je veliki broj veza i interakcija sa ostalim članovima zajednice. Iz tog razloga, jedan od glavnih kriterijuma je birati čvorove sa što većim stepenom pokrivanja, tj. što većim brojem nepokrivenih suseda. Takođe je bitno uzeti u obzir i potrebu nepokrivenih suseda onog čvora čije se dodavanje u PIDS razmatra. S tim u vezi, za svaki čvor  $v \in V$  potrebno je odrediti njegov stepen pokrivanja (engl. *cover-degree - cd*) i stepen potrebe njegovih nepokrivenih suseda (engl. *need-degree - nd*). Neka je  $S$  tekuće parcijalno rešenje MPIDS problema, tada je veličinom

$$h_S(v) = \left\lceil \frac{deg(v)}{2} \right\rceil - |N_S(v)|, \quad N_S(v) = N(v) \cap S. \quad (5)$$

određena pokrivenost čvora  $v$  parcijalnim rešenjem  $S$ . Preciznije, ukoliko je  $h_S(v) \leq 0$ , čvor  $v$  je pokriven tekućim rešenjem, inače je nepokriven. Koristeći uvedenu notaciju, za proizvoljan čvor  $v \in V$ , uvodimo dva kriterijuma: stepen pokrivenosti čvora  $v$  -  $cd(v)$  i stepen potrebe njegovih nepokrivenih suseda -  $nd(v)$  koji se definišu na sledeći način:

$$cd(v) = |\{u \in N(v) : h_S(u) > 0\}| \quad (6)$$

$$nd(v) = \sum_{u \in N(v)} \max\{h_S(u), 0\}. \quad (7)$$

Postojeći pohlepni algoritmi za rešavanje MPIDS problema koriste jedan ili kombinuju oba navedena kriterijuma pri čemu se u svakom koraku konstruktivnog algoritma pohlepno bira čvor sa najvećom vrednošću odabranog kriterijuma sve dok tekuće parcijalno rešenje ne postane PIDS skup.

Da bi željeni uticaj preovladao u čitavoj mreži potrebno je da svaka individua ima više pozitivnih nego negativnih suseda u svojoj okolini. Dosadašnje analize MPIDS problema pokazale su da je reč o skupu velike kardinalnosti, odnosno da PIDS skup neretko čini i više od 60% čvorova polaznog grafa [5]. Imajući to u vidu, u ovom radu je predložen novi pristup rešavanju MPIDS problema korišćenjem inverznog nasumično-pohlepnog algoritma (engl. *Reverse Randomized Greedy Algorithm - RRG*). Struktura RRG algoritma za rešavanje MPIDS problema prikazana je Algoritmom 1.

Algoritam uzima čitav skup  $V$  kao inicijalno rešenje i pohlepno izbacuje redundantne čvorove iz tekućeg rešenja čuvajući njegovu dopustivost. Treba primetiti da postoji određeni skup čvorova koji mora biti deo minimalnog dominantnog skupa sa pozitivnim uticajem. Iz tog razloga, pre procesa redukcije inicijalnog rešenja, potrebno je odrediti skup fiksiranih čvorova  $F$  i zabraniti njihovo izbacivanje. Status fiksiranog čvora imaju susedi svih čvorova sa stepenom 1. Čvor  $v \in V$  za koji je  $deg(v) = 1$  može biti pokriven od strane samo jednog čvora  $u$  koji zbog toga mora biti deo krajnjeg rešenja. Takođe, pre procesa redukcije, potrebno je formirati vektor pokrivenosti čvorova trenutnim rešenjem, koji se inicijalno poklapa sa vektorom njihovih stepena budući da su svi čvorovi u grafu maksimalno pokriveni.

Potencijalni kandidati za redukciju inicijalnog rešenja, nalaze se u skupu  $R = S \setminus F$ . Analogno  $cd$  kriterijumu, povoljan kandidat za izbacivanje iz PIDS skupa je onaj čvor čiji je stepen pokrivenosti najmanji, tj. čvor sa najmanjim stepenom. Iz tog razloga se kandidati za izbacivanje iz skupa najpre sortiraju rastuće po stepenu i tim redosledom uzimaju u razmatranje. Pri izbacivanju čvorova koristi se kriterijum inverzan  $nd$  kriterijumu - kriterijum redundantnosti. Izabrani čvor se izbacuje iz tekućeg rešenja sa verovatnoćom koja je proporcionalna stepenu njegove redundantnosti u odnosu na tekuće rešenje. Stepem redundantnosti čvora koji je deo rešenja određuje se na osnovu preopkrivenosti njegovih suseda i definiše se na sledeći način:

$$rd(v) = \begin{cases} 0, & \text{ako } \exists u \in N(v) : h_S(u) = 0 \\ \sum_{u \in N(v)} |h_S(u)|, & \text{inače} \end{cases} \quad (8)$$

Što je veći stepen redundantnosti čvora čije se izbacivanje razmatra, to je veća verovatnoća da taj čvor bude zaista i izbačen iz rešenja. Pod pretpostavkom da finalni MPIDS skup čine čvorovi visokog stepena, na verovatnoću izbacivanja redundantnog čvora, pored stepena njegove redundantnosti utiče i broj njegovih suseda. Neka je  $S$  tekuće rešenje. Verovatnoća izbacivanja čvora  $v \in S$  računa se po sledećoj formuli:

$$p(v) = \frac{d \cdot rd(v)}{\sum_{u \in N(v)} deg(u) - \left\lceil \frac{deg(u)}{2} \right\rceil} \quad (9)$$

gde je  $d = 0.5 \cdot \frac{deg(v) - d_{\min}}{d_{\max} - d_{\min}}$ , pri čemu je  $d_{\min}$  minimalni, a  $d_{\max}$  maksimalni stepen čvora u grafu  $G$ . Opisani algoritam ostavlja mogućnost da pronađeni redundantni čvor ne bude izbačen iz rešenja, pa je potrebno još jednom proći kroz sve čvorove koji su deo rešenja i izbaciti one koji su redundantni.

**Algorithm 1** Struktura RRG algoritma sa MPIDS problem

---

```

Ulaz: Prost, povezan, neusmeren graf  $G(V, E)$ 
Izlaz: Dominantan skup sa pozitivnim uticajem  $S$ 
function RRG( $G$ )
   $S \leftarrow V$ 
   $F \leftarrow \text{odrediti\_fiksirane\_cvorove}(G)$ 
   $R \leftarrow S \setminus F$ 
   $R\_sort \leftarrow \text{sortirati\_rastuce\_po\_stepenu}(R)$ 
   $H \leftarrow \text{stepeni\_cvorova}(G)$ 
  for svaki element  $v \in R$  do
    if  $v$  je redundantan then
       $rd \leftarrow \text{stepen\_redundantnosti}(v, S, G)$ 
       $p \leftarrow \text{izracuna\_j\_verovatnocu\_izbacivanja}(v, rd, G)$ 
       $\text{izbaci\_cvor\_v\_sa\_verovatnocom\_p}(S, p, v)$ 
      if  $v$  je izbacen then
         $Nv \leftarrow \text{susedi\_izbacenog\_cvora}(G, v)$ 
         $H(Nv) = H(Nv) - 1$ 
      end if
    end if
  end for
   $\text{izbaci\_preostale\_redundantne\_cvorove}(S, G, H)$ 
  return  $S$ 
end function

```

---

Predloženi RRG algoritam elemente skupa  $R$  sortira rastuće po njihovom stepenu. Takvo sortiranje nije jedinstveno budući da u opštem slučaju imamo više čvorova sa istim stepenom čiji se raspored može menjati. U cilju poboljšanja kvaliteta rešenja, implementirana je višestartna RRG procedura (engl. *Reverse Randomized Greedy Multi-Start Algorithm - RRGMS*) čija je struktura prikazana Algoritmom 2. U svakoj iteraciji pravi se novi redosled čvorova koji učestvuju u redukciji rešenja tako što se za skupove čvorova sa istim stepenom uzima proizvoljna permutacija njihovog redosleda. Ovim postupkom dodatno razmrđavamo algoritam i omogućavamo pronalaženje novih potencijalno boljih rešenja.

**Algorithm 2** Struktura RRGMS algoritma za MPIDS problem

---

```

Ulaz: Prost, povezan, neusmeren graf  $G(V, E)$  i maksimalan broj iteracija  $max\_iter$ 
Izlaz: Dominantan skup sa pozitivnim uticajem  $S$ 
function RRGMS( $G$ )
   $S\_best \leftarrow V$ 
   $iter = 0$ 
  while  $iter < max\_iter$  do
     $R\_prm \leftarrow \text{napravi\_novi\_redosled}(R)$ 
     $new\_S \leftarrow \text{RRG}(G, R\_prm)$ 
     $iter = iter + 1$ 
    if  $|S\_new| < |S\_best|$  then
       $S\_best \leftarrow S\_new$ 
    end if
  end while
  return  $S\_best$ 
end function

```

---

Prednost predloženih algoritama je da u svakom trenutku raspoložemo nekim dopustivim rešenjem, što nije slučaj u postojećim algoritmima iz literature. Klasičan pohlepni algoritam konstruiše rešenje polazeći od praznog skupa i pohlepno birajući naredni čvor koji se dodaje tekućem parcijalnom rešenju. U svakom koraku klasičnog pohlepnog algoritma je neophodno proveriti da li je tekuće parcijalno rešenje ujedno i PIDS skup, i ako jeste, proces formiranja rešenja se završava.

#### 4. EKSPERIMENTALNI REZULTATI

Testiranja su izvršena na računaru sa procesorom AMD Ryzen 5 3500U 2.10 GHz i 8GB RAM. Predloženi algoritmi implementirani su u programskom jeziku C++. Korišćen je skup instanci iz literature koje su javno dostupne na Stanford Large Network Dataset Collection (<https://snap.stanford.edu/data/>) i Network Data Repository (<https://networkrepository.com>). U tabeli 1 prikazani su parametri test instanci koje su korišćene u eksperimentalnoj analizi. U prvoj koloni dat je naziv instance, druga kolona sadrži broj čvorova -  $n$ , dok se u trećoj koloni nalazi broj grana u zadatom grafu -  $m$ . Kako je MPIDS problem definisan na prostom, povezanom i neusmerenom grafu, na svim instancama je najpre izvršeno pretprocesiranje radi uklanjanja izolovanih čvorova, petlji i paralelnih grana.

**Tabela 1:** Test instance korišćene u eksperimentalnoj analizi

Instanca	$n$	$m$
Karate	34	78
Dolphins	62	159
Football	115	613
Jazz	198	2742
socfb-nips-ego	2888	2981
socfb-Mich67	3748	81903
socfb-Brandeis99	3898	137567
CA-GrQc	5241	14484
CA-HepTh	9875	25973
CA-HepPh	12006	118589
CA-AstroPh	18771	198050
CA-CondMat	23133	93439

U tabeli 2 date su vrednosti funkcije cilja najboljih rešenja koja su dobijena postojećim pohlepnim algoritmi-ma: Wang's Greedy [5], Raeli's Greedy [4], Fei's Greedy [2], Pan's Greedy [3] i IGA-PIDS [1], redom. U tabeli 3 su prikazani rezultati predloženih algoritama RRG i RRGMS na razmatranim test instancama. Pri testiranju RRGMS metode korišćene su sledeće vrednosti za kriterijum zaustavljanja:  $max\_iter = 500$  za instance kod kojih je  $n < 1000$ , dok je na ostalim instancama  $max\_iter = 1000$ .

Druga i treća kolona tabele 3 sadrže vrednost funkcije cilja RRG rešenja -  $sol_{RRG}$  i odgovarajuće vreme izvršavanja -  $t(s)$  u sekundama. Poslednjih pet kolona tabele 3 odnose se na rezultate RRGMS metode i prikazuju: vrednost funkcije cilja najboljeg RRGMS rešenja dobijenog kroz 10 pokretanja -  $best\_sol_{RRGMS}$ , prosečnu vrednost funkcije cilja rešenja dobijenih RRGMS metodom kroz 10 pokretanja -  $avg\_sol_{RRGMS}$ , prosečno vreme izvršavanja RRGMS metode -  $t(s)$  u sekundama, prosečno procentualno odstupanje RRGMS rešenja od najboljeg rešenja dobijenog pohlepnim algoritmi-ma -  $agap(\%)$  i standardna devijacija -  $std(\%)$ . U tabelama 2 i 3, za svaku navedenu instancu, vrednosti funkcije cilja najboljih dostignutih rešenja su podebljana.

**Tabela 2:** Rezultati postojećih pohlepni algoritama za rešavanje MPIDS problema

Instanca	Wang's Greedy	Raeli's Greedy	Fei's Greedy	Pan's Greedy	IGA-PIDS
Karate	<b>15</b>	<b>15</b>	<b>15</b>	<b>15</b>	<b>15</b>
Dolphins	31	31	<b>30</b>	32	31
Football	68	68	69	69	68
Jazz	<b>81</b>	82	<b>81</b>	83	<b>81</b>
socfb-nips-ego	<b>1398</b>	<b>1398</b>	<b>1398</b>	<b>1398</b>	<b>1398</b>
socfb-Mich67	1481	1478	1473	1458	<b>1427</b>
socfb-Brandeis99	1535	1539	1529	1522	<b>1502</b>
CA-GrQc	2626	2623	2622	2612	2607
CA-HepTh	4598	4602	4582	4565	<b>4544</b>
CA-HepPh	4887	4886	4876	4857	4817
CA-AstroPh	7081	7085	7062	7030	<b>6953</b>
CA-CondMat	9869	9853	9837	9816	<b>9748</b>

Posmatrajući rezultate jednog pokretanja RRG algoritma, može se zaključiti da nov pristup redukcije rešenja nalazi rešenja dobrog kvaliteta u veoma kratkom vremenu izvršavanja, čak i na instancama sa velikim brojem

**Tabela 3:** Eksperimentalni rezultati dobijeni predloženim algoritmima RRG i RRGMS

Instanca	$sol_{RRG}$	$t(s)$	$best\_sol_{RRGMS}$	$avg\_sol_{RRGMS}$	$t(s)$	$agap(\%)$	$std(\%)$
Karate	<b>15</b>	0.0004	<b>15</b>	15.00	0.3889	0.0000	0.0000
Dolphins	<b>30</b>	0.0005	<b>30</b>	30.00	0.1328	0.0000	0.0000
Football	70	0.0005	<b>67</b>	67.40	0.2474	0.5882	0.7204
Jazz	83	0.0007	<b>81</b>	81.10	0.4714	0.1235	0.3704
socfb-nips-ego	<b>1398</b>	0.0444	<b>1398</b>	1398.00	8.4415	0.0000	0.0000
socfb-Mich67	1441	0.0128	1434	1434.20	13.0124	0.5046	0.0280
socfb-Brandeis99	1519	0.0166	1513	1513.80	14.5844	0.7856	0.0266
CA-GrQc	2610	0.0159	<b>2605</b>	2605.20	16.6129	0.0077	0.0154
CA-HepTh	4562	0.0527	4549	4549.30	76.9856	0.1166	0.0141
CA-HepPh	4829	0.0789	<b>4813</b>	4814.40	77.5240	0.0291	0.0212
CA-AstroPh	6983	0.1890	6969	6971.00	199.4324	0.2589	0.0157
CA-CondMat	9768	0.5957	9764	9766.50	484.9744	0.1897	0.0147

čvorova. Na skoro svim instancama rezultati RRG algoritma bolji su od svih pohlepnih algoritama, izuzev IGA-PIDS koji je superiorniji od RRG u nalaženju PIDS skupova najmanje kardinalnosti. Prednost RRG algoritma u odnosu na ostale pohlepne algoritme iz literature je mogućnost generisanja većeg broja različitih rešenja dobrog kvaliteta što je od velikog značaja za formiranje metaheurističkih metoda. Predložena RRGMS metoda (višestartni pristup RRG algoritma) koristi tu osobinu RRG pristupa u cilju dobijanja rešenja boljeg kvaliteta. Iz tabele 3 se može uočiti da je RRGMS metoda poboljšala rezultate IGA-PIDS algoritma na sledećim instancama: *Dolphins* za 3.23%, *Football* za 1.47%, *CA-GrQc* za 0.076% i *CA-HepPh* za 0.083%. Na instancama *Karate*, *Jazz* i *socfb-nips-ego* RRGMS metoda je pronašla PIDS skupove iste kardinalnosti kao i IGA-PIDS algoritam. Na svim instancama prosečne vrednosti parametra *agap* ne prelaze 0.79%, dok prosečna vrednost standardne devijacije nije veća od 0.73%. Na osnovu toga se zaključuje da RRGMS metoda pronalazi rešenja koja su po kvalitetu uporediva sa rešenjima IGA-PIDS algoritma. Na 3 od 12 instanci su vrednosti parametara *agap* i *std* jednake nuli, tj. RRGMS metoda je u svakom pokretanju dolazila do PIDS skupova iste kardinalnosti.

## 5. ZAKLJUČAK

U ovom radu je razmatran problem nalaženja minimalnog dominantnog skupa sa pozitivnim uticajem na primerima realnih i veštačkih socijalnih mreža. Predstavljen je nov način rešavanja ovog problema korišćenjem inverznog nasumično-pohlepnog algoritma (RRG) i njegove višestartne varijante RRGMS. Izvršeno je testiranje na poznatim grafovima iz literature sa preko 20 000 čvorova. Dobijeni rezultati su upoređeni sa rezultatima postojećih algoritama za rešavanje MPIDS problema i pokazano je da RRG pristup pohlepnog izbacivanja čvorova (pronalaženja najneuticajnijih individua) rezultuje rešenjima dobrog kvaliteta. Višestartni pristup RRG algoritma je pokazao stabilnost u pronalaženju rešenja visokog kvaliteta u kratkom vremenu izvršavanja čak i na instancama velikih dimenzija. Jedan od pravaca daljeg istraživanja je hibridizacija predloženih RRG i RRGMS algoritama sa drugim heurističkim pristupima. Drugi pravac je nadogradnja MPIDS modela uvođenjem težina koje bi predstavljale jačinu veze između individua i definisanje novog kriterijuma uticajnosti u mreži i razvoj adekvatnih heurističkih metoda za njegovo rešavanje.

## LITERATURA

- [1] Bouamama, S., & Blum, C. (2021). An Improved Greedy Heuristic for the Minimum Positive Influence Dominating Set Problem in Social Networks. *Algorithms*, 14(3),79
- [2] Fei, M., & Weidong, C. (2016). An improved algorithm for finding minimum positive influence dominating sets in social networks. *J. South China Norm. Univ*, 48, 59–63.
- [3] Pan, J., & Bu, T. (2019). A Fast Greedy Algorithm for Finding Minimum Positive Influence Dominating Sets in Social Networks. *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 360-364.
- [4] Raei, H., Yazdani, N., & Asadpour, M. (2012). A New Algorithm for Positive Influence Dominating Set in Social Networks. *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 253-257.
- [5] Wang, F., Camacho, E., & Xu, K. (2009). Positive influence dominating set in online social networks. In: *International Conference on Combinatorial Optimization and Applications*, 313-321.

## HYBRID OPTIMIZATION ALGORITHMS BASED ON GREEDY HEURISTICS AND THE BRANCH-AND-BOUND METHOD FOR LOGICAL ANALYSIS OF DATA

IGOR MASICH<sup>1</sup>, LEV KAZAKOVTSSEV<sup>2</sup>, ALENA STUPINA<sup>3</sup>, EKATERINA KRAEVA<sup>4</sup>

<sup>1</sup> Siberian Federal University – Laboratory "Hybrid Methods of Modelling and Optimization in Complex Systems", is.masich@gmail.com

<sup>2</sup> Siberian Federal University – Laboratory "Hybrid Methods of Modelling and Optimization in Complex Systems", levklevk@gmail.com

<sup>3</sup> Siberian Federal University – Laboratory "Hybrid Methods of Modelling and Optimization in Complex Systems", h677hm@gmail.com

<sup>4</sup> Reshetnev University – Institute of Space Technology, ekaterina.masich@gmail.com

---

**Abstract:** Logical analysis of data (LAD) consists in identifying patterns in a dataset that cover a portion of similar observations. These patterns are combinations of some attribute values and can be used to solve machine learning problems. The most useful patterns are those that are optimal according to some criterion, e.d., patterns with maximum coverage. To search for them, both mixed linear programming and approximate algorithms, mostly greedy heuristics, are used. We propose a hybrid algorithm for pattern search that combines greedy heuristics and the branch-and-bound method. The use of the hybrid algorithm improves the quality of generated patterns compared to greedy heuristics. At the same time, the complexity remains significantly lower than the complexity of finding the exact solution using integer or mixed linear programming. The proposed hybridization scheme allows us to control the trade-off between the quality of generated patterns and the complexity of searching for patterns in LAD.

**Keywords:** Logical analysis of data, Branch-and-bound method, Greedy heuristics





## NEW IDEAS TO SPEED-UP FLOYD-WARSHALL SHORTEST PATHS ALGORITHM

GIUSEPPE LANCIA<sup>1</sup>, FRANCA RINALDI<sup>2</sup>

<sup>1</sup> University of Udine – D.M.I.F, giuseppe.lancia@uniud.it

<sup>2</sup> University of Udine – D.M.I.F, franca.rinaldi@uniud.it

**Abstract:** Floyd and Warshall's algorithm for the all-pairs shortest path problem is a  $\Theta(n^3)$  procedure which revisits  $n$  times all the cells of an  $n \times n$  distance matrix. At each pass, all the cells are checked but only some of them get updated. In this paper, we report some preliminary results on a new version of the algorithm, designed to avoid checking cells which will not be updated, in order to reduce the overall time. Our procedure uses heaps to quickly identify which cells can be good candidates for an update. The new version improves over Floyd-Warshall's original for those input graphs in which the number of cells updated over all passes is substantially smaller than the number of checks. However, our procedure is worse than the original if the ratio between cell checks and updates is not large enough. To obtain an improvement independently of the particular instance type, we propose a hybrid combination of the two approaches, which starts with the original Floyd and Warshall version and then switches to the new one after some iterations. Preliminary experiments show the effectiveness of this strategy.

**Keywords:** Shortest path algorithm, Floyd Warshall, All-pairs shortest paths, Smart force

### 1. INTRODUCTION

We are given in input a directed graph  $G = (V, A)$  with arc costs  $c(i, j)$ . Wlog,  $V = \{1, 2, \dots, n\}$ . The costs do not need to be nonnegative, but for the results to be correct there must be no negative-length cycle in  $G$ .

The goal is to compute the shortest path length between each pair of vertices  $i, j$ . We will denote the length of a shortest  $i \rightarrow j$  path by  $L[i, j]$ .

This is a classical problem in combinatorial optimization, which is solved via a classical algorithm, named Floyd-Warshall (FW) after the authors who inspired it [3, 7]. The algorithm updates a matrix  $L[i, j]$  initialized as  $L[i, j] := c(i, j)$  for  $(i, j) \in A$ , while  $L[i, j] := +\infty$  for  $(i, j) \notin A$  and  $L[i, i] := 0$  for all  $i$ . It then iterates three nested for-cycles:

```
1. for  $k := 1, \dots, n$ 
2.   for  $i := 1, \dots, n$ 
3.     for  $j := 1, \dots, n$ 
4.       if  $(L[i, j] > L[i, k] + L[k, j])$  then
5.          $L[i, j] := L[i, k] + L[k, j]$ ;
```

At the end  $L[i, j]$  is the shortest length of an  $i \rightarrow j$  path, for each  $i, j \in V$ . The complexity of the algorithm is  $\Theta(n^3)$ . This is a worst-case complexity, but also the best-case and average-case are the same.

Major modifications of the basic Floyd-Warshall scheme, to reduce either its worst- or average-case computational complexity, have been proposed, for instance, by Brodnik et al. in [1] for complete graphs and by Sao et al. in [6] for sparse graphs. In this extended abstract, we introduce a relatively simple idea with a straightforward implementation, to speed-up the basic FW procedure. As discussed in Section 5, the idea yields quite promising results.

### 2. MAIN IDEA

Let us call line 4 of Floyd-Warshall algorithm an execution of  $\text{test}(i, j|k)$  and line 5 an execution of  $\text{update}(i, j|k)$  (the update of cell  $[i, j]$  due to  $k$ ).

The Floyd-Warshall algorithm can be seen as an enumeration of all triples  $(i, j, k)$  to execute the tests  $\text{test}(i, j|k)$ , some of which will cause an update and some will not. We could say that a test (and the triple) *succeeds* if it causes an update, while it *fails* otherwise. It is clear that when a triple fails the time spent to

consider it was wasted, but we don't know in advance which triples will succeed, or we would enumerate only them. If  $\mathcal{S}$  is the set of succeeding triples and  $\mathcal{F}$  is the set of failing triples, it is expected (and observed in our experiments) that  $|\mathcal{S}|$  is much smaller than  $\binom{n}{3}$ . In particular, after some iterations of the algorithm (as  $k$  approaches  $n$ ), the number of new succeeding triples drops significantly, as the matrix  $L[\cdot, \cdot]$  is almost entirely filled with the final correct values.

FWalgorithm is very lean and it appears very difficult to speed it up while maintaining its general scheme. Indeed, there are  $\binom{n}{3}$  triples overall, and the time it takes Floyd-Warshall to create and test one triple is  $O(1)$ . This constant, call it  $a$ , is very small since there is a minimal cost for bookkeeping. Let us call  $b$  the constant time to execute  $\text{update}(i, j|k)$ . Then, the total running time is  $a(|\mathcal{S}| + |\mathcal{F}|) + b|\mathcal{S}|$ .

Ideally, in the best case, to beat it we would only enumerate the triples in  $\mathcal{S}$ . Suppose finding them costs us, on average,  $f(n)$  each. The total cost would be  $f(n)|\mathcal{S}|$  and so the new algorithm would be competitive as long as  $f(n)|\mathcal{S}| < a(|\mathcal{S}| + |\mathcal{F}|) + b|\mathcal{S}|$ . However, finding precisely only the successful triples seems quite hard to do. In this paper we propose a heuristic strategy, that exploits a necessary condition that all successful triples must satisfy, in order to lower the probability that a triple elected for the test will in fact fail. In particular, while not all triples will be enumerated, still our scheme will not miss any succeeding one.

### 3. THE SMART-FORCE ALGORITHM

The general question of Floyd-Warshall is:

QFW: *Having fixed  $k$ , which are the pairs  $(i, j)$  that complete a succeeding triple  $(i, j, k)$ ?*

Notice that there are  $\Theta(n)$  choices for  $k$ , while the answer takes  $\Theta(n^2)$  time.

We will rephrase QFW into two questions:

Q1: *Having fixed  $(k, i)$ , which are the indices  $j$  that complete a succeeding triple  $(i, j, k)$ ?*

Q2: *Having fixed  $(k, j)$ , which are the indices  $i$  that complete a succeeding triple  $(i, j, k)$ ?*

There are  $\Theta(n^2)$  choices for  $(k, i)$  (or  $(k, j)$ ), but we aim at finding the answer in less than  $\Theta(n)$  time, thus realizing a speed-up with respect to Floyd-Warshall.

An entry  $(i, j)$  gets updated if  $L[i, k] + L[k, j] < L[i, j]$ . Therefore, a necessary condition for a triple  $(i, j, k)$  to be succeeding is that

$$\left( L[i, k] < \frac{L[i, j]}{2} \right) \vee \left( L[k, j] < \frac{L[i, j]}{2} \right)$$

which we can rewrite as  $\left( L[i, j] > 2L[i, k] \right) \vee \left( L[i, j] > 2L[k, j] \right)$ .

When  $(k, i)$  are fixed we will check the first part of the condition, over all  $j$ , and when  $(k, j)$  are fixed, we will check the second part over all  $i$ . This way, all succeeding triples would be found. The above checks are not done by complete enumeration (brute-force) but rather by direct access to only the indices that can fulfill the condition (a technique we called "smart-force" (SF) in our previous papers [4, 5]).

Let us consider the case of  $(k, i)$  being fixed, so that  $2L[i, k] =: \alpha$  is a constant. The condition for some  $j$  to be part of a succeeding triple with  $i$  and  $k$  is that  $L[i, j] > \alpha$ . This implies, for instance, that if we knew that  $\max_j L[i, j] \leq \alpha$ , then we could avoid enumerating the  $n$  possible values for  $j$ . We therefore want to sample only values  $j_1, j_2, \dots, j_t$  such that  $L[i, j_h] > \alpha$  for  $h = 1, \dots, t$ . We achieve this by keeping a (row-)heap  $H_i^R$  for each  $i$ . The heap contains all elements  $j$  of row  $i$ , sorted by their  $L[i, j]$  value. This way popping the maximum index takes time  $O(1)$ , while readjusting the heap after each pop takes  $O(\log n)$ . The total number of triples thus sampled is  $t$  rather than  $n$  at a cost of  $O(t \log n)$  which is no worse than  $\Theta(n)$  as long as  $t$  is  $O(n/\log n)$ .

In a perfectly analogous way, we work for the case in which  $(k, j)$  are fixed, by keeping  $n$  (column-)heaps  $H_j^C$ , one for each  $j$ , and looking for indices  $i$  such that  $L[i, j] > \alpha := 2L[k, j]$ .

Building the heaps at the very beginning has cost  $O(n)$  per heap, for a total cost  $O(n^2)$ . Each row-heap (and, respectively, column-heap) is implemented in the customary way as an array  $H_i^R[\cdot]$  of  $n$  records  $\langle \text{ndx}, \text{value} \rangle$ , where  $\text{ndx}$  represents a column- (respectively, row-) index, while  $\text{value}$  is the value of an entry of  $L[i, \cdot]$  (respectively,  $L[\cdot, j]$ ). E.g., if  $H_3^R[1] = \langle 5, 100 \rangle$  it means that the highest value in row 3 is at column 5, and it is  $L[3, 5] = 100$ . The second- and third- highest values of row 3 are in the next two entries of  $H_3^R[\cdot]$  (not necessarily in this order). In general, each index  $h$  identifies a node of the heap whose left subtree is rooted at  $2h$  while the right subtree is rooted at  $2h + 1$ .

In the course of the algorithm, we need to keep the heaps consistent with  $L$ , i.e., each time an entry  $L[i, j]$  is updated, we need to also update  $H_i^R$  and  $H_j^C$ . In order to make these updates effectively, we use a back-pointer data

**Algorithm 1** SF-FW

---

```

    /* create row- and column- heaps */
1. for  $i \leftarrow 1$  to  $n$  do
2.   for  $j \leftarrow 1$  to  $n$  do
2.      $H_i^R[j] \leftarrow \langle j, L[i, j] \rangle$ ;  $\text{Pos}_{H_i^R}[j] \leftarrow j$ 
3.   for  $v \leftarrow \lfloor n/2 \rfloor$  downto 1 do heapify( $H_i^R, v$ )
4. for  $j \leftarrow 1$  to  $n$  do
5.   for  $i \leftarrow 1$  to  $n$  do
5.      $H_j^C[i] \leftarrow \langle i, L[i, j] \rangle$ ;  $\text{Pos}_{H_j^C}[i] \leftarrow i$ 
6.   for  $v := \lfloor n/2 \rfloor$  downto 1 do heapify( $H_j^C, v$ );
7. for  $k \leftarrow 1$  to  $n$  do
8.   for  $i \leftarrow 1$  to  $n$  do /* run updates due to column  $k$  */
9.      $\alpha \leftarrow 2L[i, k]$ 
10.     $t \leftarrow 0$ 
11.    while (  $H_i^R[1].\text{value} > \alpha$  )
12.      popped[ $t$ ]  $\leftarrow$  popHeapTop(  $H_i^R$  )
13.       $j :=$  popped[ $t$ ].ndx
14.      popped[ $t$ ].value  $\leftarrow$   $L[i, j] \leftarrow$  eval( $L, i, j, k$ )
15.       $t++$ 
    /* restores heap. Popped entries get new value */
16.   for  $e := 1$  to  $t$  do insertHeap(  $H_i^R$ , popped[ $e$ ] )
17. for  $j \leftarrow 1$  to  $n$  do /* run updates due to row  $k$  */
18.    $\alpha \leftarrow 2L[k, j]$ 
19.    $t \leftarrow 0$ 
20.   while (  $H_j^C[1].\text{value} > \alpha$  )
21.     popped[ $t$ ]  $\leftarrow$  popHeapTop(  $H_j^C$  )
22.      $i :=$  popped[ $t$ ].ndx
23.     popped[ $t$ ].value  $\leftarrow$   $L[i, j] \leftarrow$  eval( $L, i, j, k$ )
24.      $t++$ 
    /* restores heap. Popped entries get new value */
25.   for  $e := 1$  to  $t$  do insertHeap(  $H_j^C$ , popped[ $e$ ] )

```

---

structure into the heap, i.e., we keep "parallel" arrays  $\text{Pos}_{H_i^R}[\cdot]$  (respectively,  $\text{Pos}_{H_j^C}[\cdot]$ ) that tell the position of each column index  $j$  (respectively, row index  $i$ ) in the heap  $H_i^R$  (respectively,  $H_j^C$ ). In our example,  $\text{Pos}_{H_3^R}[5] = 1$  since  $j = 5$  is the root of heap  $H_3^R$ . This way, whenever  $L[i, j]$  is changed, we go to the node corresponding to  $j$  in  $H_i^R$ , and change its value (which necessarily decreases). In doing so, we let the node slide down in the heap, by exchanging it with the largest of its two sons, if necessary, until it finds its correct position. We then record the new position of  $j$  in  $\text{Pos}_{H_i^R}[j]$ . The overall cost of this update is  $O(\log n)$ .

## 4. PSEUDOCODE

### 4.1. Procedure SF-FW.

The main procedure SF-FW is outlined in Procedure 1. We assume the input to be an  $n \times n$  cost matrix  $L[\cdot, \cdot]$ , and the final result, i.e., all pairs shortest paths lengths, will be overwritten on  $L[\cdot, \cdot]$ .

The procedure starts by creating the row-heaps  $H_i^R$  for each row  $i$ , and column-heaps  $H_j^C$  for each column  $j$  (lines 1–6). At this time, also the parallel arrays  $\text{Pos}_{H_i^R}$  and  $\text{Pos}_{H_j^C}$  are initialized. This is done within the procedure HEAPIFY (Procedure 2), which, while moving the elements in the heaps, keeps track of their positions.

The main loop starts at line 7 and ends at line 25. It is composed of two sub-loops. The first one (lines 8–16) takes care of all the updates to  $L[\cdot, \cdot]$  due to the elements on column  $k$ , i.e., the pairs  $(i, k)$  for varying  $i$ . A perfectly analogous loop (lines 17–25) concerns the updates due to row  $k$  of  $L[\cdot, \cdot]$ , i.e., pairs  $(k, j)$  for varying  $j$ . We will comment only on the first loop since similar considerations hold also for the second.

The loop 8–16 starts by determining the threshold  $\alpha = 2L[i, k]$  which must be beat by an index  $j$  to be considered in a possibly succeeding triple. In lines 11–15, all elements  $X := \langle j (= \text{ndx}), \text{value} \rangle$  in the heap such that  $L[i, j] > \alpha$  are popped from the heap and put in a local buffer  $\text{popped}[\cdot]$ . The triple  $(i, j, k)$  is then

evaluated and the new value of the entry  $L[i, j]$  is stored as the new value of  $X$ . At the end  $t$  is the total number of elements that were popped (i.e., of triples evaluated). At this point, the loop in line 16 puts back in the heap  $H_i^R$  the elements that were popped, with their new values. The back-pointers  $\text{Pos}_H^R$  are updated within the procedure INSERTHEAP (Procedure 3).

#### 4.1.1 Procedures for heaps management.

A heap  $H$  is a binary trees stored in consecutive cells of an array (see [2]). In general, the left son of node  $H[i]$  is node  $H[2i]$ , while the right son is node  $H[2i + 1]$ . The heap property requires that the value of each node is not smaller than the value of any of its sons. Thus, the largest value of any subtree is at the root node of the subtree.

HEAPIFY, depicted in Procedure 2, is a standard procedure that, given an index  $v$  and an array of  $n$  records such that the subtrees rooted at  $2v$  and  $2v + 1$  are heaps, turns the subtree rooted at  $v$  into a heap. This is obtained by letting  $H[v]$  "slide" down in the heap until it finds its final position. This version of HEAPIFY is such that the back-pointers to the heap are updated while the elements are moved around. Note that by calling HEAPIFY for  $v = \lfloor n/2 \rfloor, \dots, 1$ , an entire array can be turned into a heap.

---

#### Procedure 2 HEAPIFY ( $H, v$ )

---

```

1.  $ls \leftarrow 2v$  /* left son */
2.  $rs \leftarrow 2v + 1$  /* right son */
3. if ( $H[ls].value > H[v].value$ ) then  $large \leftarrow ls$  else  $large \leftarrow v$ 
4. if ( $H[rs].value > H[large].value$ ) then  $large \leftarrow rs$ 
5. if ( $large \neq v$ ) then
6.    $\text{Pos}_H[H[v].\text{ndx}] \leftarrow large$ 
7.    $\text{Pos}_H[H[large].\text{ndx}] \leftarrow v$ 
8.    $H[v] \leftrightarrow H[large]$  /* swaps  $H[v]$  with the largest of its sons */
9.    $\text{heapify}(H, large)$ 

```

---

The procedure INSERTHEAP (procedure 3) is used to insert a new element  $e_l$  into a heap  $H$  of less than  $n$  elements. The element is initially put at the rightmost bottom leaf (position  $H[H.SIZE]$ ) and then is lifted up to its final position, i.e., when its value is not larger than its parent's. This is done in the loop 3–6. The position of each element which is moved down in the heap while  $e_l$  moves up is updated within the loop. Eventually, the element finds its final position in lines 7–8.

---

#### Procedure 3 INSERTHEAP ( $H, e_l$ )

---

```

1.    $H.SIZE++$ 
2.    $curr \leftarrow H.SIZE$ 
3.   while ( $curr > 1$ ) AND ( $H[\lfloor \frac{curr}{2} \rfloor].value < e_l.value$ )
4.      $\text{Pos}_H[H[\lfloor \frac{curr}{2} \rfloor].\text{ndx}] \leftarrow curr$ 
5.      $H[curr] \leftarrow H[\lfloor \frac{curr}{2} \rfloor]$ 
6.      $curr \leftarrow \lfloor \frac{curr}{2} \rfloor$ 
7.    $\text{Pos}_H[e_l.\text{ndx}] \leftarrow curr$ 
8.    $H[curr] \leftarrow e_l$ 

```

---

The procedure POPHEAPTOP (not reported, for space limitations) extracts from the heap, and returns, the element of maximum value (which is the root node, i.e.,  $H[1]$ ). It then rebuilds the heap by replacing the root with the leaf  $H[H.SIZE]$  which, by calling  $\text{heapify}(H, 1)$ , is moved down along the tree until the heap property is again fulfilled. The positions in the heap of the elements moved are updated accordingly, while the position of the element extracted from the heap is set to -1. The cost of this procedure is  $O(\log(H.SIZE))$ .

## 5. DISCUSSION AND COMPUTATIONAL EXPERIMENTS

The cost of producing a triple to test is  $O(\log n)$  for the SF\_FW procedure, while for FW it is  $O(1)$ . Furthermore: (i) the hidden multiplicative constant for the complexity of SF\_FW is certainly higher than for FW; (ii) for SF\_FW there is the possibility that some triples  $(i, j, k)$  are evaluated twice (this happens when both  $L[i, k] < L[i, j]/2$  and also  $L[k, j] < L[i, j]/2$ ). For the above reasons, the only hope for SF\_FW to outperform

FW is that the input graph is such that only a *very small number of triples* (compared to  $n^3$ ) are updated in the course of FW procedure, and not many more than these are evaluated by SF\_FW.

**Graphs with few updates.** For an extreme example, consider a complete graph in which the edge costs are random numbers in a range  $\{m, \dots, M\}$  with  $m \geq M/2$ . For such a graph, each edge is in fact a shortest path and hence no triple would lead to an update. As far as SF\_FW is concerned, for any target  $\alpha = 2L[k, j]$  (or  $\alpha = 2L[i, k]$ ) no entry of  $L[, ]$  would ever be popped from the heaps, with the exception of the phases in which  $k = j$  or  $k = i$  (i.e., the phases in which the target is given by one of the  $n$  diagonal elements). In these phases,  $\alpha = 0$  and  $n - 1$  heap elements get popped). Therefore, overall, only  $2n(n - 1)$  triples get checked for SF\_FW while  $n^3$  triples get checked for FW. Notice that the time to recognize that each edge is indeed a shortest path cannot be smaller than  $\Theta(n^2)$  and so SF\_FW is optimal for these inputs.

If we perturb this type of input by altering a few entries, still SF\_FW outperforms Floyd-Warshall. For an example, consider graphs of the above type in which we pick at random  $n$  arcs and give them a new random value in  $[1, m]$ , so that some triples will probably lead to an update (i.e., they are successful). Then we obtain the following average results (for  $m = 200$ ,  $M = 399$ , over 5 instances for each  $n$ , rounded to integer)

$n$	successful triples	FW attempted triples	FW time	SF_FW attempted triples	SF_FW time
1000	322,487	1,000,000,000	5.6s	3,708,892	0.8s
2000	1,313,753	8,000,000,000	42.4s	15,009,216	3.0s
3000	3,011,423	27,000,000,000	135.4s	34,221,686	7.2s

To characterize the best inputs for SF\_FW, it would be important to identify classes of graphs on which FW would run into a large ratio of attempted vs successful triples. One such type of graph could be a graph with some “fast” nodes (also called *hubs*) intermixed with general nodes. In such a graph each edge incident on a hub would be “short”, while all the other edges can be of any length (like, in real life, large cities connected by highways are quicker to reach than small remote villages) and a shortest path most likely uses only hubs as intermediate nodes. Assuming the nodes are sorted according to their average distance from the rest, the hubs are processed first, and so for the remaining iterations probably very few entries of  $L[, ]$  will be updated.

We simulate a graph of this type by the following parameters:  $n_h$  is the number of hubs;  $m$  is the minimum cost of any arc;  $M_h$  is the maximum cost for an arc incident with a hub;  $M$  is the maximum cost for any other arc. We have performed tests on this type of graphs, with  $n = 2000$  nodes and either 10, 20 or 40 hubs. We have set  $m = 100$  and  $M = 10000$  and have varied the length of the arcs incident on the hubs in  $\{120, 150, 200\}$ . The results (times in seconds) show a speed improvement from 3 times to 5 times of SF\_FW versus the original Floyd-Warshall procedure:

$M_h$	40 hubs			20 hubs			10 hubs		
	FW	SF_FW	Speed-up	FW	SF_FW	Speed-up	FW	SF_FW	Speed-up
200	41.9s	16.1s	2.6×	42.0s	14.1s	3.0×	41.9s	13.4s	3.1×
150	41.8s	14.0s	3.0×	42.0s	12.0s	3.5×	41.5s	10.1s	4.1×
120	41.3s	12.6s	3.3×	41.6s	10.0s	4.2×	41.2s	8.1s	5.1×

**More general graphs and the hybrid approach.** The overhead due to the heap management compared to the minimal bookkeeping of FW is such that when the arc costs vary in a large interval (and therefore also the lengths of the paths can be very different among them), the original FW easily outperforms SF\_FW. However, as soon as this variability in the arc costs starts to decrease, SF\_FW becomes competitive and eventually can be better than Floyd-Warshall. Consider the following experiment run on complete graphs of size  $n = 1000$  whose arc costs are drawn randomly in  $\{1, \dots, \text{MAX}\}$  for various decreasing values of MAX:

MAX	successful triples	FW attempted triples	FW time	SF_FW attempted triples	SF_FW time
100	7,843,881	1,000,000,000	5.6s	61,760,461	10.4s
50	5,889,226	1,000,000,000	5.6s	55,583,872	9.3s
10	2,248,129	1,000,000,000	5.6s	20,532,006	3.4s
5	993,010	1,000,000,000	5.5s	8,289,043	1.6s

The table shows how SF\_FW benefits from the arc costs varying in a small range (“small” should be intended in a relative way, as shown by our next experiment. That, is, if we keep MAX fixed, SF\_FW benefits from the increase of  $n$ ). For  $n = 1000$ , we see that SF\_FW beats Floyd-Warshall for  $\text{MAX} \leq 10$ , but this range is indeed too small to be really interesting.

In order to obtain a procedure which is effective also when the arc costs vary in a reasonably large range, we made the following observation. During the execution of FW algorithm, the entries get updated easily in the early phases, but then, once the values get closer to the true shortest path lengths, it becomes more and more difficult to beat them and so the number of updates drops dramatically. Therefore, it would make sense to use the original FW for some iterations, in order to improve the current matrix, and then continue on this matrix with the SF\_FW version for the remaining iterations.

We then performed an experiment, for various values of  $n$ , on complete graphs with independent random arc costs drawn in the range  $\{1, \dots, 100\}$ . For each instance, we have tested what would happen if we start with  $S$  iterations of FW followed by  $n - S$  iterations of SF\_FW, with  $S = \frac{h}{10}n$  and  $h = 0, 1, \dots, 10$ . Notice that when  $h = 0$  the hybrid approach degenerates into SF\_FW, while for  $h = 10$  it is FW.

switch point	$n = 1000$	$n = 2000$	$n = 3000$	$n = 4000$	$n = 5000$
$n$ (FW)	5.6s	42.0s	130.1s	304.3s	607.4s
$0.9 n$	5.1s	39.6s	131.6s	290.3s	562.0s
$0.8 n$	5.4s	37.8s	128.9s	290.7s	534.0s
$0.7 n$	5.1s	35.9s	123.1s	276.3s	512.8s
$0.6 n$	4.8s	32.9s	118.2s	257.7s	489.1s
$0.5 n$	<b>4.5s</b>	31.0s	111.3s	245.5s	465.1s
$0.4 n$	4.6s	30.8s	104.9s	232.8s	438.3s
$0.3 n$	4.7s	<b>30.5s</b>	99.2s	220.1s	418.4s
$0.2 n$	5.5s	32.1s	<b>95.9s</b>	<b>204.2s</b>	392.7s
$0.1 n$	6.8s	36.9s	105.5s	208.4s	<b>395.5s</b>
0 (SF_FW)	11.0s	61.2s	167.5s	322.0s	585.6s

From the table, we can see that there is always some combination of FW+SF\_FW which beats FW (indeed, with the exception of just one entry for  $n = 1000$ , all combinations we tested are better than FW alone). We have highlighted the best combination for each  $n$ , and it looks that, when  $n$  increases, the number of iterations of FW needed to get the best speed-up of the search decreases. A heuristic rule that yields good results for  $1000 \leq n \leq 5000$  appears to be "run 500 iterations of FW and then switch to SF\_FW". Another interesting observation from the table is that, for increasing  $n$ , pure SF\_FW becomes more and more competitive with pure FW, and eventually it beats it for  $n = 5000$ .

## 6. CONCLUSIONS

We have suggested a new "smart-force" strategy to sample triples in Floyd-Warshall algorithm, in order to increase their probability of success. Preliminary results show the strategy to be very effective on some particular graphs, and a research question would be to characterize other classes of graphs on which the approach should perform well. Furthermore, we proposed a hybrid approach which is effective on more general graphs. An important question would be to characterize the best switching point from the old to the new strategy. This could be done by extensive testing and tuning, with possibly the use of machine-learning techniques. Other lines of future research concern finding permutations of the nodes that might help the running time, and incorporating simple changes to the strategy that can avoid the possibility of enumerating the same triple twice. Finally, we want to investigate if this type of idea can be applied to other algorithms, for this or other optimization problems.

## REFERENCES

- [1] Brodnik A., Grgurovič M. and Požar R. (2022). Modifications of the Floyd-Warshall algorithm with nearly quadratic expected-time. *Ars Mathematica Contemporanea* 22, 1–22.
- [2] Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C. *Introduction to Algorithms, Third Edition*, 3rd ed.; The MIT Press: Cambridge, MA, USA, 2009.
- [3] Floyd, R. W. (1962). Algorithm 97: Shortest Path. *Communications of the ACM*, 5(6), 345.
- [4] Lancia, G., Dalpasso, M., (2020). Finding the Best 3-OPT Move in Subcubic Time. *Algorithms*, 13(11), 306–321.
- [5] Lancia, G., Vidoni, P., (2020). Finding the largest triangle in a graph in expected quadratic time. *European Journal of Operational Research*, 286, 458–467.
- [6] Sao P., Kannan R., Gera P., Vuduc R. (2020) A Supernodal All-Pairs Shortest Path Algorithm in Proceedings of the 25th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming.
- [7] Warshall, S. (1962). A theorem on Boolean matrices. *Journal of the ACM*, 9(1), 11–12.

## THE CHOICE OF HEURISTICS FOR FORMATION OF PARTIAL LOGICAL RULES IN THE METHOD OF LOGICAL ANALYSIS OF DATA

ROMAN KUZMICH<sup>1</sup>, ALENA STUPINA<sup>2,3</sup>, KATERINA PONOMAREVA<sup>4</sup>, VLADISLAV STASIUK<sup>5</sup>

<sup>1</sup> Siberian Federal University – Department of Business Informatics and Business Process Modeling, romazmich@gmail.com

<sup>2</sup> Siberian Federal University – Department of Digital Management Technologies, astupina@sfu-kras.ru

<sup>3</sup> The Siberian Fire and Rescue Academy of State Firefighting Service of the Ministry of Russian Federation for Civil Defence, Emergencies and Elimination of Consequences of Natural Disasters – Department of Civil Defense and Emergency Management, h677hm@gmail.com

<sup>4</sup> Siberian Federal University – Department of Digital Management Technologies, ponomareva\_katerina@mail.ru

<sup>5</sup> Siberian Federal University – Department of Digital Management Technologies, vlad1394@bk.ru

---

**Abstract:** *The paper proposes a number of heuristics for the formation of logical rules in the method of logical analysis of data. Each heuristic is presented as a combinatorial optimization model, on the basis of which logical rules should be obtained. The composition of the obtained rules forms a classifier, on the basis of which the decision-making process on the classification of a new observation takes place. Thus, the quality of classification of new observations directly depends on the informativity of the resulting logical rules. Therefore, in order to improve the quality of classification, it is necessary to make a choice of the best heuristic for the formation of logical rules. The choice of the best heuristics is carried out by checking their suitability for the formation of a set of rules that allows you to build a classifier with a high generalizing ability.*

**Keywords:** *Logical Rule, Classifier, Heuristics, Optimization Model*

### 1. INTRODUCTION

Currently, in practice, it is increasingly necessary to solve classification problems in which the cost of a mistake in making a decision is critical, for example, medical and economic problems [1, 2, 3]. In medical problems, human life and health depend on the quality of decision-making. An incorrect solution to an economic problem can lead to significant financial damage or bankruptcy of the enterprise. Thus, methods are increasingly used for such problems that allow not only to obtain a solution (to determine a class for a new observation), but also to substantiate it. This family of methods includes logical classification algorithms [4].

Logical classification algorithms rely on initial data in the decision making process. Raw data is a sample of observations obtained during the data collection stage. For each observation in the sample, the values of the input attributes and the resulting attributes (class) are known.

The classification problem is a supervised learning problem. Originally, the classification algorithm is trained on the initial data, and then new observations are presented to it as input, for which the value of the resulting attribute is unknown. The task of the algorithm is to determine the value of the resulting attribute (class) for a new observation using the knowledge gained in the learning process.

A feature of logical classification algorithms is the representation of the acquired knowledge in the form of a set of logical rules. Logical rules are most often searched for in the form of conjunctions of a small number of terms that represent elementary logical statements (conditions). This is due to the need for easy interpretation by a person of the received logical rules.

The main tasks of the paper are: firstly, the proposal of heuristics for constructing logical rules in the method of logical analysis of data [2], related to logical classification algorithms; secondly, comparison of the proposed heuristics by testing them on a real classification problem; thirdly, the choice of the best heuristic for building logical rules from the proposed heuristics in the paper.

### 2. MATERIALS AND METHODS

Let us introduce  $\varphi: X \rightarrow \{0, 1\}$ , a predicate defined on the set of observations  $X$ . If the condition  $\varphi(x) = 1$  is satisfied, then the predicate  $\varphi$  covers the observation  $x$ . A predicate  $\varphi$  will be called a logical rule  $\varphi$  if it has the informativity property: it covers a sufficiently large number of observations of its “own” class and

practically does not cover observations of other classes. It follows that any resulting logical rule covers only a part of the observations of the set  $X$ . Thus, to classify any observation from the set  $X$ , it is necessary, first, to construct various logical rules based on the observations of this set. Secondly, to combine the obtained rules into a composition – a classifier, on the basis of which a decision is made about whether a new observation belongs to a specific class.

A class for a logical rule is called its “own” class if the rule is built on the basis of observations of this class. Observations of “one’s own” class are also called positive, while others are called negative.

Each logical rule has common properties: coverage (the number of observations of its class that satisfy its conditions) and degree (the number of terms in the conjunction).

## 2.1. Setting the problem of finding informative logical rules

To formulate the problem of finding informative logical rules, we introduce the following notation [4]:

1.  $P_k$  is the number of class  $k$  observations in the sample  $X^\ell = (x_i, w_i)$ , where  $W = \{1, \dots, M\}$  – is a finite set of class labels,  $P_k > 1, i = 1, \dots, \ell$ ;
2.  $p_k(\varphi)$  is the number of observations of class  $k$  for which the condition  $\varphi(x) = 1$  is satisfied;
3.  $N_k$  is the number of observations of all other classes  $W \setminus \{k\}$  in the sample  $X^\ell$ , where  $N_k > 1, P_k + N_k = \ell$ ;
4.  $n_k(\varphi)$  is the number of observations of all other classes  $W \setminus \{k\}$  for which the condition  $\varphi(x) = 1$  is satisfied.

According to the indicated notation, the problem of constructing an informative logical rule  $\varphi$  consists in optimizing according to two criteria:  $p_k(\varphi) \rightarrow \max$  and  $n_k(\varphi) \rightarrow \min$ .

Let us introduce the notation  $E_k(\varphi, X^\ell)$  and  $D_k(\varphi, X^\ell)$ . These are the proportions of negative and positive observations among all observations covered by the logical rule  $\varphi$  in the sample  $X^\ell$ :

$$E_k(\varphi, X^\ell) = \frac{n_k(\varphi)}{p_k(\varphi) + n_k(\varphi)};$$

$$D_k(\varphi, X^\ell) = \frac{p_k(\varphi)}{p_k(\varphi) + n_k(\varphi)}.$$

Based on the introduced notation, we define the concept of a logical  $\varepsilon$ -,  $\delta$ -rule for the class  $k$ . If  $E_k(\varphi, X^\ell) \leq \varepsilon$  and  $D_k(\varphi, X^\ell) \geq \delta$  for given sufficiently small  $\varepsilon$  and sufficiently large  $\delta$  from the interval  $[0, 1]$ , then the logical rule  $\varphi$  is called the logical  $\varepsilon$ -,  $\delta$ -rule.

A logical rule is called pure if  $n_k(\varphi) = 0$ . If  $n_k(\varphi) > 0$ , then the logical rule  $\varphi$  is called partial. If the initial data of real practical classification problems contain missing values of features and mistakes in the numerical values of features, then it is better to build partial logical rules to increase stability when working with such initial data [1].

## 2.2. Heuristics for building partial logic rules

In the previous paragraph, it was indicated that the problem of constructing an informative logical rule is reduced to optimization according to two criteria:  $p_k(\varphi) \rightarrow \max$  and  $n_k(\varphi) \rightarrow \min$ .

The first way to solve this problem is to convolve the criteria  $p_k(\varphi) \rightarrow \max$  and  $n_k(\varphi) \rightarrow \min$  into a single scalar criterion, which we will further use as the objective function of the optimization model for constructing logical rules. It turned out to be difficult to offer an adequate convolution. For example, in [5] the following versions of the simplest heuristics are proposed:  $p_k(\varphi) - n_k(\varphi) \rightarrow \max$ ;  $p_k(\varphi) - 5n_k(\varphi) \rightarrow \max$ ;  $\frac{p_k(\varphi)}{P_k} - \frac{n_k(\varphi)}{N_k} \rightarrow \max$ . These heuristics have a significant drawback, since two varieties of logical rules unsuitable for classification can be constructed and turn out to be preferable. The first ones cover too few observations, the second cover positive and negative observations approximately in the proportion in which they are represented in the training sample  $X^\ell$ .

However, in [6], it was possible to propose an adequate convolution in the form of a boosting criterion:

$$\sqrt{p_k(\varphi)} - \sqrt{n_k(\varphi)} \rightarrow \max. \quad (1)$$

The basis for the adequacy of the boosting criterion for formation of logical rules is the fact that it has been successfully used for a long time in the algorithm of the same name, which belongs to the weighted rules voting algorithm [7].

The second way to solve the problem of constructing informative logical rules is to consider one of the criteria as an objective function, and use the second one as a constraint function.

We will demonstrate the implementation of the second method on a model for constructing partial logical rules with the maximum coverage of observations of its “own” class. As the name implies, the criterion



$p_k(\varphi) \rightarrow \max$  is considered as the objective function, and the criterion  $n_k(\varphi) \rightarrow \min$  will be considered as the optimization model constraint function.

An optimization model has been developed for the method of logical analysis of data [2]. Before presenting this optimization model, let us consider a number of features of the classification method itself related to its preparation for the stage of constructing logical rules.

For clarity, we will consider a binary classification problem, in which the initial data sample consists of disjoint sets  $\Omega^+$  (a set of observations of a positive class) and  $\Omega^-$  (a set of observations of a negative class) of  $n$ -dimensional vectors. Vector elements can take quantitative, nominal, binary values. Due to the different types of feature values, a procedure for their binarization is necessary for the correct operation of the method [8].

The next step is the task of constructing a reference set of features, the meaning of which is to select a subset of features  $S$  that can separate the observation sets of positive and negative classes with high accuracy. The projections  $\Omega_s^+$  and  $\Omega_s^-$  of the sets  $\Omega^+$  and  $\Omega^-$  onto  $S$  are used in the further work of the method [2].

At the next stage of the method, we present a model for constructing partial logical rules with the maximum coverage of observations of its “own” class. Let’s present it on the example of constructing partial positive logical rules. The problem of constructing partial negative logical rules is formulated in a similar way.

A logical rule built on the basis of a positive observation  $\alpha$  is called a positive logical  $\alpha$ -rule for  $\alpha \in \{0,1\}^t$ . The task is to form for each observation  $\alpha \in \Omega_s^+$  the maximum logical  $\alpha$ -rule, that is, capturing the maximum number of observations of the positive class. Let’s define a logical  $\alpha$ -rule using binary variables  $Y = (y_1, y_2, \dots, y_t)$ :

$$y_i = \begin{cases} 1, & \text{if the } i\text{-th attribute is in the rule.} \\ 0, & \text{otherwise.} \end{cases}$$

For pure rules, coverage of observations of another class is not allowed. Therefore, for each observation of the negative class  $\beta \in \Omega_s^-$ , the variable  $y_i$  must be equal to 1 for at least one index  $i$  for which  $\beta_i \neq \alpha_i$ :

$$\sum_{\substack{i=1 \\ \beta_i \neq \alpha_i}}^t y_i \geq 1 \text{ for any } \beta \in \Omega_s^-.$$

For a stronger restriction, you can change the number 1 to another positive integer  $d$ .

To establish the fact that the logical  $\alpha$ -rule covers the observation of the positive class  $\gamma \in \Omega_s^+$  it is necessary that the value of the variable  $y_i$  be equal to 0 only for indices  $i$  that are not fixed in the rule. Thus, the objective function of the optimization model can be represented as the sum of observations of a positive class captured by the logical  $\alpha$ -rule:

$$\sum_{\substack{\gamma \in \Omega_s^+ \\ \gamma_i \neq \alpha_i}} \prod_{i=1}^t (1 - y_i).$$

Combining the objective function and the constraint function into the optimization model, we obtain [9]:

$$\sum_{\substack{\gamma \in \Omega_s^+ \\ \gamma_i \neq \alpha_i}} \prod_{i=1}^t (1 - y_i) \rightarrow \max, \quad (2)$$

$$\sum_{\substack{i=1 \\ \beta_i \neq \alpha_i}}^t y_i \geq d \text{ for any } \beta \in \Omega_s^-, y \in \{0,1\}^t. \quad (3)$$

Based on such an optimization model (2-3), we will obtain pure patterns. Since our aim is to form partial logical rules, we modify the constraint function as follows:

$$\sum_{\beta \in \Omega_s^-} z_\beta \leq D, \quad (4)$$

$$\text{where } z_\beta = \begin{cases} 0, & \text{if } \sum_{\substack{i=1 \\ \beta_i \neq \alpha_i}}^t y_i \geq d, \\ 1, & \text{otherwise.} \end{cases}$$

where  $D$  is a parameter of the optimization model, which is equal to the number of observations of the opposite class admissible for coverage.

Thus, using the obtained constraint function (4) together with the objective function (1), we obtain an optimization model for constructing partial logical rules based on the boosting criterion. And in the case of using the obtained restriction function (4) together with the objective function (2), we obtain an optimization model for constructing partial logical rules with the maximum coverage of observations of its “own” class.

### 3. RESULTS

Experimental studies are carried out on the problem of hepatitis domain, taken from the open Machine Learning Repository UCI [10]. Hepatitis domain is caused by the problem of an increase in the blood levels

of an intermediate product of the breakdown of hemoglobin – bilirubin, as a result of which liver function may be impaired, which, in extreme cases, can lead to death.

To compare and select the best one, two models for constructing partial logical rules are used: a model with the maximum coverage of observations of its “own” class and a model based on the boosting criterion.

The initial sample consists of 155 observations: 32 positive (fatal case) and 123 negative (survivors). There are missing feature values in the sample. 15% of the original sample was used as an examining sample.

**Table 1:** Accuracy of diagnosis of the disease “hepatitis domain”

Optimization model	Set of rules	Parameter D	Average degree of rule	Cover of negative patterns	Cover of positive patterns	Classification accuracy, %
Objective function (1), constraint (4)	Neg.	5	2	76	3	57
	Pos.		2	3	7	78
Objective function (2), constraint (4)	Neg.	5	2	80	5	79
	Pos.		4	5	8	89
Objective function (1), constraint (4)	Neg.	10	1	86	7	79
	Pos.		2	4	8	67
Objective function (2), constraint (4)	Neg.	10	1	88	8	79
	Pos.		3	9	10	78
Objective function (1), constraint (4)	Neg.	15	1	88	8	79
	Pos.		1	5	8	56
Objective function (2), constraint (4)	Neg.	15	1	101	14	71
	Pos.		3	14	12	56

According to Table 1, the model for constructing partial logical rules with the maximum coverage of observations of its “own” class works better when the value of the parameter D in the constraint function (4) is small, and the model for constructing partial logical rules based on the boosting criterion performs better for large values of the parameter D.

#### 4. CONCLUSION

The paper presents a number of heuristics for constructing partial logical rules. Some of the heuristics presented in the paper turned out to be poor due to a significant drawback: the possibility of constructing logical rules unsuitable for classification, covering few observations or covering positive and negative observations approximately in the proportion in which they are represented in the training set.

As a result, at the stage of real testing, only two models for constructing partial logical rules are compared: a model with the maximum coverage of observations of its “own” class and a model based on the boosting criterion. The best in terms of classification accuracy is the model for constructing partial logical rules with the maximum coverage of observations of its “own” class. However, for large values of the parameter D in the constraint function (4), the model for constructing partial logical rules based on the boosting criterion works better. Thus we conclude that it is expedient to use each of the compared models for constructing partial logical rules in solving practical problems of classification.

#### Acknowledgement

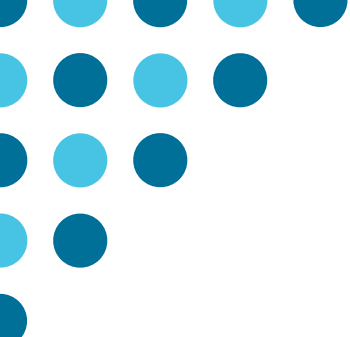
This work was supported by the Ministry of Science and Higher Education of the Russian Federation (Grant No. 075-15-2022-1121).

#### REFERENCES

- [1] Golovenkin, S.E., Gulakova, T.K., Kuzmich, R.I., Masich, I.S., & Shulman, V.A. (2010). Model of logical analysis for solving the problem of predicting myocardial infarction. *Vestnik Sib-SAU*, 4(30), 68-73.
- [2] Hammer, P.L., & Bonates, T. (2005). Logical Analysis of Data: From combinatorial optimization to medical applications. *RUTCOR Research Report*, 10, 1-27.

- 
- [3] Kochedykov, D.A., Ivakhnenko, A.A., & Vorontsov, K.V. (2005). Credit scoring system based on logical classification algorithms. *Mathematical methods of pattern recognition-12. MAKSS Press*, 349-353.
- [4] Vorontsov, K. (2010). Lectures on logical algorithms of classification. Available at: <http://www.machinelearning.ru/wiki/images/3/3e/Voron-ML-Logic.pdf>.
- [5] Zagoruiko, N.G. (2013). Cognitive data analysis. *The Russ. Academy of Sciences, Sib. Department, Institute of Mathematics named after S.L. Sobolev. Novosibirsk: Geo*, 186 p.
- [6] Furnkranz, J., & Flach, P.A. (2005). Roc „n“ rule learning-towards a better understanding of covering algorithms. *Machine Learning*, 58(1), 39-77.
- [7] Sun, B., Chen, S., Wang, J., & Chen, H. (2016). A robust multi-class AdaBoost algorithm for mislabeled noisy data. *Knowledge-Based Systems*, 102, 87-102.
- [8] Rastrigin, L., & Freymanis, E. (1988). Solving problems of multiple-scale optimization using random-search methods. *Problems of Random Search*, 11, 9-25.
- [9] Kuzmich, R.I., Masich, I.S., & Stupina, A.A. (2017). Patterns of formation of patterns in the method of logical analysis of data. *Control systems and information technology*, 1(67), 33-37.
- [10] UCI Machine Learning Repository. Available at: <http://archive.ics.uci.edu/ml/index.html>.





# LOGISTIKA I LANCI SNABDEVANJA

---

## LOGISTICS AND SUPPLY CHAIN MANAGEMENT





## AN APPROACH TO OPTIMAL SCHEDULING OF COLLECTION VEHICLES FOR CONSTRUCTION AND DEMOLITION WASTE

MILORAD VIDOVIĆ<sup>1</sup>, NENAD BJELIĆ<sup>2</sup>, BRANISLAVA RATKOVIĆ<sup>3</sup>

<sup>1</sup> University of Belgrade – Faculty of Transport and traffic engineering mvidovic@sf.bg.ac.rs

<sup>2</sup> University of Belgrade – Faculty of Transport and traffic engineering n.bjelic@sf.bg.ac.rs

<sup>3</sup> University of Belgrade – Faculty of Transport and traffic engineering mvidovic@sf.bg.ac.rs

**Abstract:** We consider pick-up and delivery routes performed by fleet of vehicles used to serve construction and demolition sites. The quantity of the construction and demolition waste (CDW) is usually larger than the capacity of a single waste collection container. Therefore, CDW collection process assumes multiple visits to the same network node, delivering an empty and picking up a loaded container. Containers are emptied at disposal site, and then transferred to the waste generation site or to the depot when there is no site demanding empty containers. The collection process is based on the full truckload (FTL) concept. The CDW collection problem analyzed here is considered as set partitioning problem with dynamic multiple time windows, where the next time window depends on the previous. We formulate the problem, and solve some numerical examples.

**Keywords:** CDW collection problem, Dynamic TW, MINLP problem formulation

### 1. INTRODUCTION

European construction sector produces 820 million tones of construction and demolition waste (CDW) every year, which is around 46% of the total amount of total waste generated according to Euro stat, (2017) – source [1]. In spite of low negative environmental effects, huge quantities of CDW are still a cause for concern, because of logistic activities and necessary land occupation. Those facts make numerous challenges for governments and attracted considerable interest in the professional and academic community. One of the most cited review papers [2] examined the status of CDW research and found many different aspects and disciplines were investigated. Although CDW collection and transportation is a source of significant costs, and the first and crucial activity triggering the others, those problems have not been considered as intensively as in case of other waste streams, particularly municipal solid waste (MSW) collection. The CDW collection process is quite different from MSW which is collected on a route with multiple stops, ending when vehicle reaches load capacity limit. In contrast, CDW is collected in direct routes from waste generation sites based on full truckload (FTL) concept. Furthermore, in the most of collection routes a vehicle picks up loaded container, after previously had delivered an empty container to the waste generation site. Containers are emptied at disposal or treatment site, and then transferred to the waste generation site or to the depot when there is no site demanding empty containers. CDW collection can be considered as a dynamic transport system where pick-up of loaded and delivery of emptied containers times are driven by intensity of construction-demolition activities and generated quantities of waste. Since the quantity of waste is usually larger than the capacity of a single container, the waste collection process may assume multiple visits to the same network node. The CDW collection problem and the industrial waste collection problem had been referred to as the skip collection problem, the rollon–rolloff collection, and the rollon–rolloff vehicle routing problem [6]. The problem was first described by Cristallo [4], according to De Meulemeester et al. [3], who formulated it as an integer linear program and provided a constraint relaxation algorithm. Formulation proposed by Bodin et al. [5] combines a set partitioning problem with a bin packing problem. To solve the problem they propose four heuristic.

Characteristics of the mentioned problem class opens, also, a space for considering CDW collection problem as a vehicle scheduling problem, while the time moments, when loaded waste containers should be picked-up and in the most cases replaced with an empty skip, give opportunity to be considered as multiple time windows (TW). Hence, in this paper we consider pick-up and delivery routes performed by fleet of vehicles used to serve construction and demolition sites with the idea to contribute in formulating the CDW collection problem as a very specific vehicle scheduling problem with the objective of earliest possible collecting containers. Our research differs from the previous approaches to solving similar problems in several ways. Although the model is based on the set partitioning introduced by Balinski and Quandt [10], it

is non linear in its nature because of the objective function type and introducing dynamic time windows in the formulation. In this way, the proposed model introduces certain contributions not only related to the CDW collection, where set partitioning formulation was only basis.

The remainder of this paper is organized as follows. The problem description is given in Section 2. The model formulation is given in Section 3. Experimental results for smaller problem instances are presented in Sections 4. Finally, conclusions are provided in Section 5.

## 2. PROBLEM DESCRIPTION

Waste collection area could be viewed as a transportation network where each waste generation site is connected only with the depot and the waste disposal site. Since the quantity of waste is (usually) larger than the capacity of a single waste collection container, the CDW collection process assumes multiple visits to the same network node, delivering an empty and picking up a loaded container. Containers are emptied at disposal site, and then transferred to the waste generation site or to the depot when there is no site demanding empty containers. The collection process is based on the full truckload (FTL) concept.

Let  $J = \{0, 1, 2, \dots, j, \dots, |J|, |J|+1\}$  be the set of network nodes  $j \in J$ , where the node  $j=0$  is a depot and the node  $j=|J|+1$  is a waste disposal site, while the nodes  $j \in J \setminus (\{0\} \cup \{|J|+1\})$  represent waste generation sites. Since the quantity of waste is larger than the capacity of a single waste collection container, waste collection process assumes multiple visits to the same network node, delivering an empty and picking up a loaded waste container. This gives opportunity for transforming graph into another, in which each node  $j \in J \setminus (\{0\} \cup \{|J|+1\})$  is replaced with the set  $n_j$  of task nodes  $i \in n_j$  whose number  $|n_j|$  corresponds to the number of tours visiting the node. After this transformation, waste collection process assumes single visits to the task nodes  $i \in n_j$ , performed by the set of collection vehicles  $k \in K$  which make one or more tours in which each vehicle visits waste generation sites, i.e. one of the task nodes  $i \in n_j$ , delivering there an empty container and picking up loaded. We assume that the vehicle fleet is homogenous, transporting containers of the same type and the capacity of  $q^c$ . Loaded container being picked up at waste generation site is then transported to the waste disposal site to be unloaded. Then, depending on the defined process dynamic, container is returned to the depot, or delivered to another waste generation site.

In order to provide CDW containers collection we consider three types of routes: Type I, Type II and Type III that could be performed by a vehicle. Routes of Type I only deliver empty containers to the waste disposal sites, providing the beginning of the collection process. Type II routes deliver an empty container to the site, pick up loaded, transfer it to disposal location and return emptied container to the next location where it is needed. Routes of Type III are the last route in the sequence performed by a vehicle. Vehicle collect container, transfer it to disposal site, where is emptied and then transfer to the depot. One vehicle in one route – sequence of task nodes can visit up to five task nodes. Note that according to the nature of the process, routes creation should strictly respect precedence constraints. For example, task node 3, cannot be visited before the task node 2, since task node 3 appears after container from the task node 2 is collected.

Except delivering an empty container at the beginning of working day, any particular tour request, i.e. task node visit, appears as the consequence of the filling up a container that should be replaced with an empty one in order to prevent slowdown or stoppage of the working process. Time needed to fill up an empty container  $t_i^F$ , delivered to a network node  $i \in N$ , is a consequence of waste loading intensity  $q_i^L$  at this location, and the container capacity  $q^c$ , i.e.  $t_i^F = q^c / q_i^L$ . To prevent slowdown or stoppage of the working process an empty container should be delivered just after the previous is fully loaded, which obviously introduces time window, in which the service should be realized, as an inherent collection process characteristic. Since a waste generation sites, depending on the CDW quantity, may require to be served several times during the planning period  $[0, T]$ , to each waste generation site, node  $j \in J \setminus (\{0\} \cup \{|J|+1\})$  are assigned multiple time windows corresponding to task nodes  $i \in n_j$ . We consider the earliest moment when a loaded container need to be replaced. Note that this problem type belongs to the class of VRP with multiple time windows [7], where service is realized in each time window. However, in the proposed transformed network, multiple time windows are also transformed, and each task node has a single time window.

For any two task nodes  $p, r \in n_j$  belonging to the same waste generation site  $j \in J$  which are served in this order, time  $w_r$ , when an empty container is delivered to serve tasks node  $r \in n_j$ , should not be before the



container in the previous task node  $p \in n_j$  is not filled up. That is,  $w_r \geq w_p + t_p^F$ , where  $t_p^F$  is container filling up time. Also, it should be noted that the moment when a loaded container is picked up from an arbitrary task node follows the moment when container is delivered to the same task node, increased by containers' replacement time.

The right bound of the time window, the latest moment when a loaded container need to be replaced or picked up (in the case of the last task at an location) depends mostly on the agreement with the customer which can accept certain period of time, usually constant  $\tau^c$ , in which loaded container should be picked up,  $u_i = l_i + \tau^c$ . In the case when  $w_i > u_i$ , the interval  $w_i - u_i$  may be also penalized, but this possibility here is not considered.

The tour duration of collection vehicles includes container loading times  $t^L$  at the depot, travel time  $t_i^{DP}$  from a depot to a task node  $i \in N$  carrying an empty container, one or more containers replacement times  $t^R$ , same for all task nodes, which includes delivering an empty and collecting an loaded container, one or more travel times  $t_i^{DS}$  between the task nodes  $i \in N$  and the waste disposal site, and one or more container unloading times  $t^U$  at the disposal site. Vehicle tour ends, because of the end of working shift, with the travel time  $t^{DSDP}$  from the disposal site to the depot. In addition to those travel and handling times, vehicles tours will include different idle times caused by time windows constraints and proposed vehicles scheduling.

Hence, in this research, we consider the problem of optimal scheduling of CDW collection vehicles under objective of minimizing slowdown and stoppage of the waste collection process. To formulate the problem we propose MINLP model to optimize collection vehicles schedule.

### 3. MINLP MODEL FORMULATION

The MINLP model formulation of optimal scheduling of CDW collection vehicles, is considered here as the FTL multiple trip vehicle scheduling with dynamic time windows under objective of minimizing lateness in containers collection, expressed through minimizing total waiting time of loaded containers in task nodes. To formulate the problem we use following notation.

#### Sets

$N^D, N^{PD}$	sets of only delivery $i \in N^D$ and pickup delivery task nodes $i \in N^{PD}$
$N$	set of all task nodes $i \in N = N^D \cup N^P \cup N^{PD}$
$J$	set of task nodes locations $j \in J$
$n_j$	set of task nodes $i \in n_j$ at location $j \in J$
$A^D, A^{PD}$	sets of routes visiting only delivery $r \in A^D$ , and pickup delivery task nodes $r \in A^{PD}$
$K$	set of collection vehicles $k \in K$

#### Parameters

$T_k^A$	available working time of the vehicle $k \in K$
$t_i^F$	time needed to fill up an empty container delivered to a task node $i \in N$ , $t_i^F = q^c / q_i^L$ , where $q^c$ is the container capacity, and $q_i^L$ loading intensity at the task node $i \in N$
$t_i^{DP}$	travel time of a vehicle from the depot to a task node $i \in N$
$t^R$	containers replacement time in a task node, and unloading time of an empty container at the depot
$t_i^{DS}$	travel time between the task node at location $j \in J$ and the waste disposal site
$t^L$	container loading time at the depot
$t^U$	container unloading time at the disposal site
$t^{DSDP}$	travel time from the disposal site to the depot
$t_i^{arr}$	travel time needed to arrive in a task node $i \in N$

#### Variables

$y_{rk}^D, y_{rk}^{PD}$	binary decision variables which take the value of 1 when a vehicle $k \in K$ visits the sequence of task nodes in route $r \in A^D$ , $r \in A^{PD}$ , respectively, otherwise 0.
$w_{ir}^D, w_{ir}^{PD}$	time when an vehicle visits task node $i \in r$ in route $r \in A^D$ , $r \in A^{PD}$ respectively

$w_k^{D \max}, w_k^{PD \max}$	largest time when an vehicle $k \in K$ visits the last task node in the route in the sets of routes $r \in A^D, r \in A^{PD}$ respectively, $w_k^{D \max} = \max_{r \in A^D \wedge i \in r} (w_{ir})$ , $w_k^{PD \max} = \max_{r \in A^{PD} \wedge i \in r} (w_{ir})$
$x_k$	binary decision variable which take the value of 1 when a vehicle $k \in K$ serves task nodes, otherwise 0
$z_{irk}^D, z_{irk}^{PD}, z_{irk}^{D \max}, z_{irk}^{PD \max}$	continuous variables which represent products of binary and continuous variable $z_{irk}^D = w_{ir} y_{rk}^D$ , $z_{irk}^{PD} = w_{ir} y_{rk}^{PD}$ , and $z_{irk}^{D \max} = y_{rk}^D w_k^{D \max}$ , $z_{irk}^{PD \max} = y_{rk}^{PD} w_k^{PD \max}$

Based on the notation given, the MINLP model can be formulated as follows.

$$\text{Min} \left( \sum_{r \in A^{PD}} \sum_{i \in r} \sum_{k \in K} w_{ir} y_{rk}^{PD} + \sum_{r \in A^D} \sum_{i \in r} \sum_{k \in K} w_{ir} y_{rk}^D + \gamma \sum_k x_k \right) \quad (1)$$

s.t.

$$\sum_{r \in A^{PD} \wedge i \in r} \sum_{k \in K} y_{rk}^{PD} = 1 \quad \forall i \in N \quad (2)$$

$$\sum_{r \in A^D \wedge i \in r} \sum_{k \in K} y_{rk}^D = 1 \quad \forall i \in N^D \quad (3)$$

$$w_{ir}^D \geq t_{ir}^{arr} - M_1(1 - y_{rk}^D) \quad \forall r \in A^D \quad \forall i \in r \quad \forall k \in K \quad (4)$$

$$w_{ir}^{PD} \geq w_{is}^D + t_i^F - M_2(1 - y_{rk}^{PD}) - M_2(1 - y_{sv}^D) \quad \forall r \in A^{PD} \quad \forall s \in A^D \quad \forall i \in r, s \quad \forall v, k \in K \quad (5)$$

$$w_{ir}^{PD} \geq t_i^{arr} - M_3(1 - y_{rk}^{PD}) \quad \forall r \in A^{PD} \quad i \in r \quad \forall k \in K \quad (6)$$

$$w_{ir}^{PD} \geq w_{i-1r}^{PD} + t_{i-1}^F - M_4(1 - y_{rk}^{PD}) \quad \forall r \in A^{PD} \quad \forall i, i-1 \in r \quad \forall k \in K \quad (7)$$

$$w_{ir}^{PD} \geq w_{i-1r}^{PD} + t_i^{arr} - M_5(1 - y_{rk}^{PD}) \quad \forall r \in A^{PD} \quad \forall i, i-1 \in r \quad \forall k \in K \quad (8)$$

$$w_{ir}^{PD} \geq y_{sk}^{PD} t_{i-1}^{arr} - M_6(1 - y_{rk}^{PD}) \quad \forall r \in A^{PD} \quad \forall s \in A^{PD} \quad r \neq s \quad \forall i \in r \quad \forall i-1 \in s \quad \forall k \in K \quad (9)$$

$$w_{ir}^{PD} \geq y_{sk}^{PD} t_{i-1}^F - M_6(1 - y_{rk}^{PD}) \quad \forall r \in A^{PD} \quad \forall s \in A^{PD} \quad r \neq s \quad \forall i \in r \quad \forall i-1 \in s \quad \forall k \in K \quad (10)$$

$$y_{rk}^D \left( \max_{r \in A^D \wedge i \in r} (w_{ir}) + t_i^{DP} \right) + y_{rk}^{PD} \left( \max_{r \in A^{PD} \wedge i \in r \wedge |i|=|j|} (w_{ir}) + t_i^{DS} + t^{DSDP} + 2t^U \right) \leq x_k T_k^A \quad \forall k \in K \quad (11)$$

$$w_{ir}, w_k^{D \max}, w_k^{PD \max}, z_{irk}^D, z_{irk}^{PD}, z_{irk}^{D \max}, z_{irk}^{PD \max} \geq 0, \quad y_{rk}^D, y_{rk}^{PD}, y_{rk}^{D \max}, y_{rk}^{PD \max}, x_k = \{0, 1\} \quad (12)$$

$$r \in A^D \vee r \in A^{PD} \quad \forall i \in r \quad \forall k \in K$$

The quadratic terms which are products of binary and continuous variable  $y_{rk} w_{ir}$  can be linearized replacing the products with the new variables  $z_{irk} = y_{rk} w_{ir}$  and by adding additional linearization constraints (12a), (12b) and (12c), as in [8]. The same concept is applied for products  $z_{rk}^{D \max} = y_{rk}^D w_k^{D \max}$  and  $z_{rk}^{PD \max} = y_{rk}^{PD} w_k^{PD \max}$

$$z_{irk} \leq M_9 w_{ir} \quad \forall r \in A \quad i \in r \quad \forall k \in K \quad (12a)$$

$$z_{irk} \leq w_{ir} \quad \forall r \in A \quad i \in r \quad \forall k \in K \quad (12b)$$

$$z_{irk} \geq w_{ir} - M_{10}(1 - y_{rk}) \quad \forall r \in A \quad i \in r \quad \forall k \in K \quad (12c)$$

Nonlinear maximization functions of type  $w_k^{\max} = \max_{r \in A \wedge i \in r} (w_{ir}) \quad \forall k \in K$  are linearized by following transformation, adding linearization constraints as in [9].

$$w_k^{D\max} \geq w_{ir} \quad \forall r \in A^D \quad \forall i \in r \quad \forall k \in K \tag{12d}$$

$$w_k^{PD\max} \geq w_{ir} \quad \forall r \in A^{PD} \quad \forall i \in r \quad \forall k \in K \tag{12e}$$

The objective function (1) tries to minimize total waiting time of loaded containers in task nodes, which corresponds to minimizing total lateness in task nodes servicing. The second term of objective (1) is bin packing part which tries to minimize number of vehicles used. Constraints (2) and (3) provide that each task node, both only delivery and pickup delivery are visited exactly once. Constraints (4) to (10) ensure the feasibility of the schedules and define arrival times at task nodes. Constraints (4) are related to arrival times at only delivery task nodes. Constraints (5) and (6) prevent vehicle to visit first task nodes in a route before it is possible because of the travel time from the depot, and before the container is filled up. Similarly, constraints (7), (8), and (9), (10) also ensure feasibility of the schedules and define arrival times at task nodes. Constraints (7) and (8) describe the case when a task node and its predecessor are in the same route, while constraints (9), (10) cover the case when a task node and its predecessor are in different routes. Constraints (11) state that the vehicle must visit all nodes without breaking available working time, while the same constraints are also bin packing constraints. Maximization functions  $\max_{r \in A^D \wedge i \in r} (w_{ir})$  and  $\max_{r \in A^{PD} \wedge i \in r \wedge i \neq |r|} (w_{ir})$  define the times of visit the last task node in a route. Constraints (12) define the domains of variables.

#### 4. NUMERICAL RESULTS

In order to analyze performances of the optimization approach proposed, we used 7 smaller randomly generated test instances (Inst\_1 to Inst\_7) which have 3 to 9 task nodes located at two and three CDW locations (Table 1). For example, the list [3,2,3] which represents instance means that there are three waste generation locations with 3, 2 and 3 task nodes in each, respectively.

Distances between the depot and CDW locations are between 6 and 19km, between the waste disposal site and waste generation locations are 12-22 km. Container capacity is assumed to be 7t, distance between the depot and disposal site is 20km. Times of loading – unloading containers at waste generation site was 7min, loading container at the depot 5min, container unloading at disposal site 4min, while the average vehicle speed was 45 km/h. Container loading intensity at waste generation site was between 0.04-0.07 t/min.

All problem instances are solved using the CPLEX 12.10, while the parameters calculation and preparation of the linearized MINLP model, through Python API, was made in Python 3.7., on Fujitsu lifebook e554, x64, Intel core i7 2.3GHz, with 12Gb RAM.

**Table 1:** Test instances used to analyze the model performances

Instance	Instance structure	Total number of routes (max. number of task nodes in a route is 5)	Objective function value (min)	Solving time (sec)
<i>Inst_1</i>	[2,1]	11	341.828	0.456
<i>Inst_2</i>	[2,2]	32	371.161	0.828
<i>Inst_3</i>	[2,3]	85	391.161	2.688
<i>Inst_4</i>	[2,1,2]	149	474.731	7.375
<i>Inst_5</i>	[2,3,2]	1079	533.398	149.541
<i>Inst_6</i>	[3,2,3]	2345	553.398	768.435
<i>Inst_7</i>	[3,3,3]	4755	574.731	3189.279

Obviously, proposed optimization model behave very well, delivering results for set of smaller test instances. However, it is obvious that with increasing number of task nodes solution times dramatically increase which clearly indicate the problem complexity and the need of heuristic approaches development.

#### 5. CONCLUSION

In this paper we consider the CDW collection problem which is modeled here as a modified set partitioning problem with different additional constraints that ensure feasibility of the schedules and define arrival times at task nodes.

The model includes dynamic time windows where the next depends on the previous. It has been formulated as MINLP and then linearized by applying different transformations. On the set of smaller test instances it is shown that the model behave very well, but with dramatic solving time increase when the number of task nodes increase.

The model complexity and the need of giving opportunity to solve the more realistic problems with larger number of task nodes and waste generation sites obviously refer to the necessity of heuristic and metaheuristic approaches. This could be understood, also as a one of main directions for future research. Besides, analyze and potential improvements of the current model, and solving it without linearization, as a quadratic MINLP are also important future research steps.

### Acknowledgement

This research was partially supported by the Ministry of Education, Science and Technological Development of the Government of Serbia through the grant No.36006 – institutional financing.

### REFERENCES

- [1] José-Luis Gálvez-Martosa, David Stylesb, Harald Schoenbergerd, Barbara Zeschmar-Lahle (2018). Construction and demolition waste best management practice in Europe, *Resources, Conservation & Recycling* 136, 166–178.
- [2] Wu, H., Zuo, J., Zillante, G., Wang, J., Yuan, H., (2019). Status quo and future directions of construction and demolition waste research: a critical review. *J. Clean. Prod.* 240, 118-163.
- [3] De Meulemeester L, Laporte G, Louveaux FV and Semet F (1997). Optimal sequencing of skip collections and deliveries. *J Opl Res Soc* 48: 57–64.
- [4] Cristallo G (1994). Optimisation de tournées de véhicules de transport container. Mémoire de licence en Sciences Economiques et Sociales, Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium.
- [5] Bodin, L., Mingozzi, A., Baldacci, R., Ball, M., 2000. The rollon–rolloff vehicle routing problem. *Transportation Science* 34 (3), 271–288.
- [6] Wy, J., Kim, B.I., Kim, S, (2013). The rollon–rolloff waste collection vehicle routing problem with time windows. *European Journal of Operational Research* 224, 466–476
- [7] H.S.Ferreira, E.T.Bogue, T.F.Noronha, S. Belhaiza, C. Prins, (2018) Variable Neighborhood Search for Vehicle Routing Problem with Multiple Time Windows, *Electronic Notes in Discrete Mathematics* Vol.66, 207–214
- [8] L.C. Coelho, (2013) Linearization of the product of two variables, <https://www.leandro-coelho.com/linearization-product-variables/>
- [9] P.A. Rubin, (2011) Max and Min Functions in a MIP, <https://orinanobworld.blogspot.com/2011/01/max-and-min-functions-in-mip.html>
- [10] M. L. Balinski, R. E. Quandt, (1964) On an Integer Program for a Delivery Problem. *Operations Research* 12(2):300-304.

## APPLICABILITY OF INDUSTRY 4.0 TECHNOLOGIES IN LOGISTICS CENTERS

SNEŽANA TADIĆ<sup>1</sup>, SMILJKA MIŠKIĆ<sup>2</sup>, ŽELJKO STEVIĆ<sup>3</sup>, MLADEN KRSTIĆ<sup>4</sup>

<sup>1</sup> University of Belgrade - Faculty of Transport and Traffic Engineering s.tadic@sf.bg.ac.rs; m.krstic@sf.bg.ac.rs

<sup>2</sup> University of East Sarajevo - Faculty of Transport and Traffic Engineering miskicmiljka@gmail.com

<sup>3</sup> University of East Sarajevo - Faculty of Transport and Traffic Engineering zeljkostevic88@yahoo.com

<sup>4</sup> University of Belgrade - Faculty of Transport and Traffic Engineering m.krstic@sf.bg.ac.rs

---

**Abstract:** Logistics activities were included in the wave of changes brought about by Industry 4.0. Through the fourth industrial revolution, technological innovations and methods have become available, which enabled the development of complex logistics systems. The following Industry 4.0 technologies have been applied in logistics centers: Internet of Things (IoT), Automated Guided Vehicles (AGV), Autonomous Vehicles (AV), Artificial Intelligence (AI), Big Data (BD), Blockchain (BC), Cloud Computing (CC), Electronic Marketplace/Mobile Marketplace (E/M Marketplace) and Advanced Robotics. In this paper, the applicability of the previously mentioned nine Industry 4.0 technologies in logistics centers was evaluated using the combined MEREC-fuzzy MARCOS model. The technologies were evaluated on the basis of a technological, socio-political and economic-operational group of criteria, which include 15 sub-criteria. Based on the obtained results, it can be concluded that Cloud Computing is the most applicable technology in logistics centers, while advanced robotics is the least applicable.

**Keywords:** Logistics center 4.0, MCMD, MEREC, fuzzy MARCOS

### 1. INTRODUCTION

The term Industry 4.0 (I4.0) collectively refers to a wide range of current concepts and is considered a new industrial stage in which vertical and horizontal production integration of processes and product connectivity is carried out, which can help companies achieve higher industrial performance. This new industrial paradigm merges the digital and physical worlds through cyber-physical systems enhanced by the Internet of Things and is expected to have a major impact on industry, the market and the economy, improving production processes and increasing productivity, affecting the entire product life cycle, creating new business models, changing the working environment and restructuring the labor market [10]. The most applicable I4.0 technologies in logistics are [8]: Internet of Things (IoT), Autonomous (AV) and Automated Guided Vehicles (AGV), Artificial Intelligence (AI), Virtual (VR) and Augmented Reality (AR), Big Data (BD), Data mining, Blockchain (BC), Cloud Computing (CC), 3D printing, Advanced Robotics, etc.

Industry 4.0 technologies are widely accepted in many areas such as: logistics, agriculture, medicine, tourism, trade, energy sector, public sector, etc. with the aim of increasing the quality, productivity and efficiency of business. By applying I4.0 technologies in the field of logistics, a new concept called Logistics 4.0 is being developed. Logistics 4.0 implies the active use of I4.0 technologies with the aim of reducing delays in the management and execution of logistics business operations, thus strengthening the organizational responsiveness, logistics capabilities and competitiveness [2]. In general, the concept of Industry 4.0 is still poorly researched in the field of logistics and there are many aspects of logistics where it can find application, which have not been explored so far, such as logistics centers (LC).

Logistics centers are places of the highest concentration of logistics activities and logistics flows [12, 17] and accordingly it is very important that they apply modern technology in their operations in order to achieve the highest possible level of business efficiency. The key technologies of I4.0 have been insufficiently analyzed in terms of the possibility of application in logistics centers. This study tries to fill those gaps in research on the application of I4.0 technologies in LC. Therefore, the aim of this research is to assess the applicability of key technologies of Industry 4.0 in LC. The assessment of the applicability of these technologies was performed using the integrated MEREC-fuzzy MARCOS model. By applying the MEREC method, the weights of the criteria and subcriteria were calculated. Based on the obtained results, the impact of technology on the labor market was singled out as the most important criterion. After that, the key technologies of I4.0 were ranked using the fuzzy MARCOS method. Based on the obtained results, Cloud Computing is the most applicable technology in logistics centers, followed by Internet of Things and Big Data. The remainder of this study is structured through four more sections. The concept of logistics centers 4.0 is defined in Section 2, while

the applied methodology is presented in Section 3. Section 4 presents the results of the application of the integrated MEREC-fuzzy MARCOS model. Section 5 concludes by giving suggestions for future research.

## 2. LOGISTICS CENTER 4.0

A logistics center is a place that has been adapted and modernized to take on multiple logistical tasks, such as receiving, transporting, storing, preparing and distributing goods. They can appear in different numbers, forms, sizes, with different functions, etc. [13]. The formation of LCs enables the rational division of work on the transport market, the concentration of commodity work, the choice of the most favorable transport carrier in macrodistribution and the concentration of activities with a unique service to urban areas and industrial zones, then in macrodistribution with a unique and convenient information system in all links of the logistics chain. Due to the increasing application of modern I4.0 technologies in this area, logistics 4.0 has already become a reality, so logistics centers 4.0 are being talked about more and more often. The term **logistics center 4.0** refers to a logistics center that applies various modern I4.0 technologies in the main segments of its business with the aim of modernizing and digitizing processes and activities, reducing all business costs, and improving productivity and business efficiency. Logistics centers 4.0 not only digitize their processes, but also use continuous information flows, data analytics and intelligent machine control, among other tactics. They have IT systems that can track millions of items in real time, as well as storage systems and automated handling equipment that perform multiple tasks automatically [5]. The implementation of I4.0 technologies is spreading more and more in logistics centers by areas of activity. The huge amount of material flows and the high concentration of logistics activities in the logistics hub make automation and digitization vital for achieving competitiveness in the market.

**IoT** technologies in the LC shorten the time needed to perform tasks, enable optimization and simplification of work processes, provide accurate data on cargo flows, inventory status, equipment and means of transport in real time, which further enables the rational use of the capacity of means of transport and warehouses [15]. IoT enables vehicle-to-vehicle communication, which facilitates the process of organization of transportation, and in addition, it also enables increased cooperation and information exchange between suppliers, carriers and customers, as all information about inventory is available in real time [1].

**Big Data** technology enables fast and efficient processing and analysis of the huge amount of data it collects through the application of various IoT technologies. BD can be used in logistics centers for: predictive analysis of transport, evaluation and definition of the cost structure for future business, analysis of information coming from customers in the entire market, which helps to improve the quality of services and collection of various types of data necessary for better decision-making. The integration of BD technology and **Cloud Computing** technology can improve the transparency and quality of logistics center service management [14]. The main advantages of CC are that anyone with the appropriate authorization can access them regardless of location and time, and cloud resources can be scaled very easily and without delay. This technology enables the connection of all decentralized LC objects into one unit. The high level of data and process security guaranteed by established "cloud" providers is another advantage of CC technology [3]. Some of the most famous CC software solutions used in the field of logistics are: Enterprise Resource Planning (ERP), Warehouse Management System (WMS), Inventory Management System (IMS), Electronic Data Interchange (EDI), Transport Management System (TMS), Intelligent Transport System (ITS), telematics systems, Package Management System (PMS), etc.

**Blockchain** is a technology that is based on the security of confidential information, information that is usually of a financial nature. BC in the logistics center enables faster and more reliable data exchange, improved tracking and visibility of goods in supply chains, automation of commercial contracts between business partners through smart contracts based on BC, data transparency to all partners and users, development of new business models, etc.

By applying **artificial intelligence** in logistics centers, it is possible to achieve an enviable degree of automation of all logistics activities, thus achieving better performance and greater efficiency [9]. Automating logistics processes using AI-based control frees up employees working in warehousing, transportation, planning or procurement of these activities and instead gives them more time to concentrate on more complex, value-adding tasks. The emergence and development of AI has influenced the development of Advanced robotics and AV. By using intelligent robots in internal transport and storage, significant improvements in quality, speed of handling goods, better use of storage capacity, greater safety and efficiency of all realized activities are achieved.

Using **Advanced robotics** in internal transport and storage, significant improvements in quality, speed of handling goods, better use of storage capacity, greater safety and efficiency of all realized activities are achieved. The use of robots is of particular importance in performing more physically demanding work tasks,

where it greatly reduces the possibility of error, and on the other hand, increases the safety of the process itself [4].

**Automated guided vehicles** have long been used in internal transport and storage, reducing labor costs, speeding up and facilitating processes in given subsystems, increasing safety, accuracy and productivity. **Autonomous vehicles** rely on AI software based on deep learning techniques. The use of such vehicles brings many advantages, which are reflected in the following: reduction of transport costs, faster delivery time, better use of vehicle capacity, greater safety and security of goods during transport, higher quality of delivery, reduction of vehicle maintenance costs, reduction of the negative impact of vehicles on the environment, greater delivery flexibility [16].

The development of the **E/M-marketplace** platform led to the intensive development of the logistics market. Logistics service markets are actually becoming part of electronic markets because the customer now automatically buys the logistics service with the purchase of the product. LCs play a key role in e-commerce because e-commerce companies carry out activities related to storage, transportation and distribution of goods using LCs, and thus it depends on the LC where the company will store its inventory, how quickly it can deliver customer orders and the different types of shipping options you can offer your customers.

### 3. METHODOLOGY

The methodology of the paper consists of the three main steps:

**Step 1:** Define the structure of the decision-making problem.

**Step 2:** Calculate the criteria weights by applying the MEREC method.

**Step 3:** Rank the technologies by applying fuzzy MARCOS method.

Defining the structure of the problem includes: defining alternatives and criteria, defining scales for evaluating alternatives, and choosing methods for evaluating market values of criteria and ranking alternatives. The applicability of the following nine I4.0 technologies in LCs was analyzed: Internet of Things ( $A_1$ ), Automated Guided Vehicles ( $A_2$ ), Autonomous Vehicles ( $A_3$ ), Artificial Intelligence ( $A_4$ ), Big Data ( $A_5$ ), Blockchain ( $A_6$ ), Cloud Computing ( $A_7$ ), E/M Marketplace ( $A_8$ ) and Advanced Robotics ( $A_9$ ). The applicability of technologies was analyzed on the basis of three groups of criteria that include five sub-criteria each. The technological group of criteria includes the following sub-criteria [7]: degree of development ( $C_1$ ), possibility of integration (modularity) ( $C_2$ ), complexity of implementation ( $C_3$ ), possibility of standardization ( $C_4$ ) and adaptability ( $C_5$ ). The socio-political group of criteria includes the following sub-criteria: safety ( $C_6$ ), labor market impact ( $C_7$ ), environmental impact ( $C_8$ ), cultural framework ( $C_9$ ) and political framework ( $C_{10}$ ). The economic-operational group of criteria includes: implementation costs ( $C_{11}$ ), energy consumption efficiency ( $C_{12}$ ), security ( $C_{13}$ ), organizational readiness ( $C_{14}$ ) and logistics service quality ( $C_{15}$ ). Sub-criteria  $C_3$ ,  $C_8$ ,  $C_{11}$ ,  $C_{14}$  are cost-related, while all other sub-criteria are beneficial.

The weight values of the defined criteria were calculated using the MEREC method (step 2), and the ranking of the selected technologies of I4.0. was performed using the fuzzy MARCOS method (step 3). The scales that were used to evaluate the alternatives in relation to the observed criteria when forming the initial matrix of the MEREC method and the initial matrix of the fuzzy MARCOS method are shown in table 1. A total of nine linguistic terms are defined and for each term a numerical value, i.e. a triangular fuzzy number.

**Table 1:** Evaluation scale

Evaluation scale for MEREC				Evaluation scale for fuzzy MARCOS			
Linguistic Evaluation	Abbreviation	Numerical Value		Linguistic Evaluation	Abbreviation	Triangular fuzzy number	
		Benefit	Cost			Benefit	Cost
None	N	1	9	Extremely poor	EP	(1,1,1)	(7,9,9)
Very Low	VL	2	8	Very poor	VP	(1,1,3)	(7,7,9)
Low	L	3	7	Poor	P	(1,3,3)	(5,7,7)
Fairly Low	FL	4	6	Medium poor	MP	(3,3,5)	(5,5,7)
Medium	M	5	5	Medium	M	(3,5,5)	(3,5,5)
Fairly High	FH	6	4	Medium good	MG	(5,5,7)	(3,3,5)
High	H	7	3	Good	G	(5,7,7)	(1,3,3)
Very High	VH	8	2	Very good	VG	(7,7,9)	(1,1,3)
Extremely High	EH	9	1	Extremely good	EG	(7,9,9)	(1,1,1)

MEREC is an objective method for obtaining criteria weights that includes the following six steps [6]: forming the initial decision matrix, normalizing the decision matrix ( $N$ ), calculating the total performance of alternatives ( $S_i$ ), calculating the performance of alternatives by removing each criterion, calculating absolute deviations and determining the final weight of criteria. Fuzzy MARCOS method is implemented by applying the following ten steps [11]: formation of the initial fuzzy matrix for decision-making, formation of the extended initial fuzzy matrix, normalization of the extended initial fuzzy matrix  $\tilde{N}$ , calculation of the difficult fuzzy matrix  $\tilde{V}$ , calculation of the  $\tilde{S}_i$  fuzzy matrix, calculation of the degree of utility of alternatives  $\tilde{K}_i$ , calculation of the fuzzy matrix  $\tilde{T}_i$ , determination of utility functions in relation to the ideal  $f(\tilde{K}_i^+)$  and anti-ideal  $f(\tilde{K}_i^-)$  solution, determination of the utility function of alternatives  $f(\tilde{K}_i)$  and ranking of alternatives based on the final values of the utility functions.

**4. THE RESULTS**

In this part of the paper, expert evaluations of alternatives based on defined criteria and the results of applicability assessment of nine previously mentioned I4.0 technologies using the integrated MEREC-fuzzy MARCOS model are presented. At the very beginning, the initial decision-making matrix of the MEREC method was formed, for three groups of criteria, i.e. technological criteria -  $C_T$ , socio-political criteria -  $C_{SP}$  and economic-operational criteria -  $C_{EO}$  (table 2), and then an initial matrix was formed for all 15 sub-criteria (table 3).

**Table 2:** Initial matrix of main criteria

Linguistic Ratings			
	$C_T$	$C_{SP}$	$C_{EO}$
$A_1$	VH	M	EH
$A_2$	H	FH	VH
$A_3$	VH	H	VH
$A_4$	EH	H	H
$A_5$	EH	VH	EH
$A_6$	H	FH	VH
$A_7$	EH	H	VH
$A_8$	FH	VH	H
$A_9$	VH	VH	H

**Table 3:** Initial matrix of subcriteria

	Technological					Socio-political					Economic-operational				
	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	$C_7$	$C_8$	$C_9$	$C_{10}$	$C_{11}$	$C_{12}$	$C_{13}$	$C_{14}$	$C_{15}$
$A_1$	VH	EH	H	M	EH	VH	FH	VH	L	H	FH	H	M	H	EH
$A_2$	EH	M	FL	FH	L	H	L	FL	FH	VH	FL	VH	M	M	H
$A_3$	L	L	L	M	L	M	L	M	FH	FL	L	M	FL	FL	FH
$A_4$	L	H	FL	L	FH	FH	FL	H	H	M	M	FH	FH	H	FH
$A_5$	FH	FH	VH	FH	VH	H	H	FL	FL	FH	FH	H	M	EH	VH
$A_6$	H	VH	FH	FH	VH	FH	VH	L	FL	M	FH	VH	H	VH	H
$A_7$	H	VH	VH	H	EH	FH	VH	L	M	M	H	H	FH	H	H
$A_8$	VH	FL	VH	M	FL	M	VL	H	H	H	M	M	FL	M	L
$A_9$	M	L	L	M	L	H	N	FL	H	L	FH	FL	M	L	FL

By applying the MEREC method, the following criteria/subcriteria weight values were obtained:

$$\begin{aligned}
 & C_T=0.382; & C_{SP}=0.453; & C_{EO}=0.165; \\
 & w_1=0.223; & w_2=0.212; & w_3=0.192; & w_4=0.192; & w_5=0.182; \\
 & w_6=0.070; & w_7=0.444; & w_8=0.126; & w_9=0.184; & w_{10}=0.176; \\
 & w_{11}=0.175; & w_{12}=0.181; & w_{13}=0.096; & w_{14}=0.282; & w_{15}=0.266
 \end{aligned}$$

After combining the previous criteria and sub-criteria values, we get the following final criteria weight values:

$$\begin{aligned}
 & w_1=0.085; & w_2=0.081; & w_3=0.073; & w_4=0.073; & w_5=0.069; \\
 & w_6=0.032; & w_7=0.201; & w_8=0.057; & w_9=0.083; & w_{10}=0.080; \\
 & w_{11}=0.029; & w_{12}=0.030; & w_{13}=0.016; & w_{14}=0.046; & w_{15}=0.044
 \end{aligned}$$



Based on the results obtained, it can be concluded that the most important is the seventh criterion related to the impact on the labor market, followed by other criteria in the following order: degree of development, cultural framework, possibility of integration (modularity), political framework, complexity of implementation, possibility of standardization, adaptability, environmental impact, organizational readiness, quality of logistics services, safety, efficiency of energy consumption, implementation costs and security.

After the weight values of the criteria were calculated, the initial matrix for the fuzzy MARCOS method was formed (table 4). The results obtained using the fuzzy MARCOS method are shown in table 5.

**Table 4:** The initial decision matrix of the fuzzy MARCOS method

	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>	A <sub>9</sub>
C <sub>1</sub>	VG	EG	P	P	MG	G	G	VG	M
C <sub>2</sub>	EG	M	P	G	MG	VG	VG	MG	P
C <sub>3</sub>	G	MP	P	MP	VG	MG	VG	VG	P
C <sub>4</sub>	M	MG	M	P	MG	MG	G	M	M
C <sub>5</sub>	EG	P	P	MG	VG	VG	EG	MP	P
C <sub>6</sub>	VG	G	M	MG	G	MG	MG	M	G
C <sub>7</sub>	MG	P	P	MP	G	VG	VG	VP	EP
C <sub>8</sub>	VG	MP	M	G	MP	P	P	G	MP
C <sub>9</sub>	P	MG	MG	G	MP	MP	M	G	G
C <sub>10</sub>	G	VG	MP	MP	MG	M	M	G	P
C <sub>11</sub>	MG	MP	P	M	MG	MG	G	M	MG
C <sub>12</sub>	G	VG	M	MG	G	VG	G	M	MP
C <sub>13</sub>	M	M	MP	MG	M	G	MG	MP	M
C <sub>14</sub>	G	M	MP	G	EG	VG	G	M	P
C <sub>15</sub>	EG	G	MG	MG	VG	G	G	P	MP

**Table 5:** The final results of applying the fuzzy MARCOS method

	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>	A <sub>9</sub>
<b>Ki</b>	0.714	0.387	0.195	0.335	0.671	0.648	0.785	0.396	0.179
<b>Rank</b>	2	6	8	7	3	4	1	5	9

Based on the table, we can conclude that the most applicable technology in logistics centers is CC, followed by the IoT and BD. Advanced robotics is in the last place in the ranking, which means that it is the least applicable technology in logistics centers.

## 5. CONCLUSION

The paper assessed the applicability of nine I4.0 technologies in LCs based on three groups of criteria that include five sub-criteria each. The assessment of acceptability of technologies was performed using the integrated MEREC-fuzzy MARCOS model. Based on the obtained weight values of criteria and/or sub-criteria using the MEREC method, the impact of technology on the labor market was singled out as the most important criterion. Observing the results obtained by applying the fuzzy MARCOS method, we can conclude that the most applicable technology in LCs is Cloud Computing. The least applicable technology is advanced robotics. Therefore, the goal of this paper to analyze and evaluate the applicability of key technologies was successfully realized by applying an integrated MCMD (multi-criteria decision-making) model. One of the future directions of research may include the analysis and assessment of the impact of I4.0 technologies on the efficiency of operations of logistics centers. In addition, future research can be manifested through the analysis of the practical application of key I4.0 technologies with the aim of solving numerous problems in business that may appear not only in LC but also in the entire supply chain.

## REFERENCES

- [1] Agarwal, V., Sharma, S., & Agarwal, P. (2021). IoT based smart transport management and vehicle-to-vehicle communication system. In *Computer Networks, Big Data and IoT* (pp. 709-716). Springer, Singapore.

- 
- [2] Chen, M., Sinha, A., Hu, K., & Shah, M. I. (2021). Impact of technological innovation on energy efficiency in industry 4.0 era: Moderation of shadow economy in sustainable development. *Technological Forecasting and Social Change*, 164, 120521.
- [3] Chen, Y. H. (2020). Intelligent algorithms for cold chain logistics distribution optimization based on big data cloud computing analysis. *Journal of Cloud Computing*, 9(1), 1-12.
- [4] Echelmeyer, W., Kirchheim, A., Lilienthal, A. J., Akbiyik, H., & Bonini, M. (2011). Performance indicators for robotics systems in logistics applications. In *IROS workshop on metrics and methodologies for autonomous robot teams in Logistics (MMARTLOG)* (p. 55).
- [5] Feng, B., & Ye, Q. (2021). Operations management of smart logistics: A literature review and future research. *Frontiers of Engineering Management*, 8(3), 344-355.
- [6] Keshavarz-Ghorabae, M., Amiri, M., Zavadskas, E. K., Turskis, Z., & Antucheviciene, J. (2021). Determination of objective weights using a new method based on the removal effects of criteria (MERECE). *Symmetry*, 13(4), 525.
- [7] Krstić, M., Agnusdei, G. P., Miglietta, P. P., Tadić, S., & Roso, V. (2022). Applicability of Industry 4.0 Technologies in the Reverse Logistics: A Circular Economy Approach Based on COmprehensive Distance Based RANking (COBRA) Method. *Sustainability*, 14(9), 5632.
- [8] Krstić, M., Tadić, S., & Zečević, S. (2021). Technological solutions in logistics 4.0. *Ekonomika preduzeća*, 69(6-7), 385-401.
- [9] Pandian, A. P. (2019). Artificial intelligence application in smart warehousing environment for automated logistics. *Journal of Artificial Intelligence*, 1(02), 63-72.
- [10] Pereira, A., & Romero, F. (2017). A review of the meanings and the implications of the Industry 4.0 concept. *Procedia Manufacturing*, 13, 1206-1214.
- [11] Stanković, M., Stević, Ž., Das, D. K., Subotić, M., & Pamučar, D. (2020). A new fuzzy MARCOS method for road traffic risk analysis. *Mathematics*, 8(3), 457.
- [12] Tadić, S., Krstić, M., & Brnjac, N. (2019). Selection of efficient types of inland intermodal terminals. *Journal of Transport Geography*, 78, 170-180.
- [13] Tadić, S., Krstić, M., Roso, V., & Brnjac, N. (2019). Planning an intermodal terminal for the sustainable transport networks. *Sustainability*, 11(15), 4102.
- [14] Wang, Y., Feng, L., Chang, H., & Wu, M. (2017). Research on the impact of big data on logistics. In *MATEC Web of Conferences* (Vol. 100, p. 02015). EDP Sciences.
- [15] Weihua, G., Tingting, Z., & Yuwei, Z. (2013). On RFID application in the information system of rail logistics center. *Int. J. Educ. Manag. Eng*, 3(2), 52-58.
- [16] Willems, L. (2021). Understanding the Impacts of Autonomous Vehicles in Logistics. *The Digital Transformation of Logistics: Demystifying Impacts of the Fourth Industrial Revolution*, 113-127.
- [17] Zečević, S., 2006. Freight Terminals and Freight Villages (in Serbian). Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade.

## DEFINISANJE SKUPA FAZI PRAVILA U FAZI SISTEMU ZA IZBOR SNABDEVAČA DEFINING A SET OF FUZZY RULES IN THE SUPPLIER SELECTION FUZZY SYSTEM

GORDANA RADIVOJEVIĆ<sup>1,2</sup>, MILICA MITROVIĆ<sup>1</sup>, DRAŽEN POPOVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Saobraćajni fakultet, Beograd, {milica.mitrovic, g.radivojevic, d.popovic}@sf.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Institut Mihajlo Pupin, Beograd, gordana.radivojevic@pupin.rs

**Rezime:** *Preduslov za uspešno i efikasno upravljanje lancem snabdevanja je adekvatan izbor snabdevača. Izbor snabdevača je problem višekriterijumskog odlučivanja i predstavlja stratešku odluku kompanija. U ovom radu je primenom fazi logike određen indeks rangiranja 50 snabdevača. Cilj ovog rada je razvoj dva fazi sistema za izbor snabdevača koji se razlikuju u određivanju vrednosti izlazne promenljive u fazi pravilima. Snabdevači su ocenjeni prema sledećim kriterijumima: kvalitet, cena, geografska udaljenost, reputacija i profesionalizam. Rezultati fazi sistema su pokazali da se model može koristiti za rešavanje realnih problema izbora snabdevača.*

**Ključne reči:** *Fazi logika, Fazi sistem, Izbor snabdevača, Lanac snabdevanja*

**Abstract:** *A prerequisite for successful and efficient supply chain management is adequate supplier selection. Suppliers selection is a problem of multi-criteria decision-making and represents a strategic decision of companies. In this paper, the ranking index of 50 suppliers was determined by applying the fuzzy logic. The aim of this paper is to develop two-fuzzy systems for supplier selection that differ in determining the value of the output variable in the fuzzy rules. Suppliers are evaluated according to the following criteria: quality, price, geographical distance, reputation and professionalism. The results of the fuzzy system showed that the model can be used to solve real problems of supplier selection*

**Keywords:** *Fuzzy logic, Fuzzy system, Supplier selection, Supply chain*

### 1. UVOD

Upravljanje lancem snabdevanjem na pravi način predstavlja ključ uspeha svake kompanije. Izbor adekvatnog snabdevača je glavni uslov za postizanje efikasnog lanca snabdevanja. U današnje vreme, lanci snabdevanja su više dinamični i nepredvidivi nego što su nekada bili. Prema tome, izbor snabdevača je jedna od ključnih aktivnosti sektora nabavke u bilo kojoj organizaciji jer zaposleni u ovom sektoru mogu uticati na ostvarenje ušteda i smanjenje rizika u poslovanju kompanije. Dva glavna zadatka u procesu izbora snabdevača su izbor kriterijuma i metoda za evaluaciju snabdevača. Najčešće metode koje autori koriste za ocenjivanje snabdevača su metode višeatributivnog i višekriterijumskog odlučivanja, fazi logike i kombinovanje više metoda.

Cilj ovoga rada je opis postupka za definisanje fazi pravila u uslovima kada postoji veći broj kriterijuma odlučivanja i kada je teško obezbediti konzistentnost ekspertske ocene za svako fazi pravilo. Rad obuhvata četiri dela. U prvom delu je opisan problem koji se rešava. Drugi deo obuhvata postupak definisanja fazi pravila a treći deo je primer primene fazi sistema. U četvrtom delu su data zaključna razmatranja.

### 2. OPIS PROBLEMA

U radu se posmatra jedna logistička kompanija i problem izbora snabdevača za određenu grupu artikala koji se kasnije distribuiraju maloprodajnim objektima. U praksi se ovaj problem rešava odlučivanjem grupe menadžera (ekspertata). Izbor snabdevača skoro nikada ne zavisi od jednog ili dva kriterijuma već od skupa kriterijuma koji opisuju razmišljanje i zaključivanje ekspertata. Odluka se donosi na osnovu različitih kvantitativnih i kvalitativnih kriterijuma: cene, troškova, kvaliteta, prethodnog iskustva, partnerskih odnosa, reputacije, profesionalnosti u poslovanju, pozicije na tržištu, mišljenja korisnika i kupaca, itd.

Proces odlučivanja ekspertata je veoma teško modelirati iz više razloga:

- Veliki broj kriterijuma utiče na odlučivanje ali je njihov značaj različit,
- Kriterijume je teško kvantifikovati za svakog potencijalnog snabdevača,
- Razmišljanje i zaključivanje ekspertata nije uvek moguće matematički formulisati,

- Eksperti često ne mogu da identifikuju sve svoje kriterijume jer je njihovo odlučivanje na osnovu znanja, iskustva, prethodnih stanja, intuicije itd.

U literaturi postoji veliki broj istraživanja na ovu temu. Autori su za rešavanje ovakvih problema razvijali heurističke algoritme i/ili primenjivali različite metode višekriterijumskog odlučivanja. Haddad i ostali [3] su primenili fazi TOPSIS (engl. Technique for Order Performance by Similarity to Ideal Solution) metodu za ocenu tri snabdevača nafte i gasa posmatranjem četiri kriterijuma iz oblasti zdravlja, bezbednosti i životne sredine. Rezultati su pokazali da je fazi TOPSIS odgovarajuća metoda za problem izbora snabdevača, jer omogućava veću efikasnost i agilnost u procesu izbora snabdevača u posmatranim oblastima. Fagundes i ostali [2] su predložili model zasnovan na FEAHP metodi (engl. Fuzzy Extended Analytic Hierarchy Process) za izbor snabdevača posmatranjem relevantnih rizika. Model automatizuje izbor snabdevača utvrđivanjem hijerarhije kriterijuma, podkriterijuma i alternativa na racionalan, fleksibilan i agiln način.

Puška i ostali [6] su predstavili model za izbor snabdevača zasnovan na primeni intervalne fazi logike i modela grupnog odlučivanja. Intervalna fazi logika je primenjena na AHP i TOPSIS metode. Metodima AHP su određene težine kriterijuma, dok je TOPSIS metodom urađeno rangiranje snabdevača. Jain i ostali [4] su razvili model zasnovan na fazi sistemu zaključivanja (engl. Fuzzy Interfes System – FIS) i fazi MCDM (engl. Multi-criteria decision-making) pristupu za ocenu performansi održivosti snabdevača u industriji gvožđa i čelika u Indiji. Autori su koristili tri dimenzije održivosti: ekonomsku, ekološku i socijalnu. Çalik [1] je razvio model za izbor zelenog snabdevača zasnovan na komponentama Industrije 4.0 kombinujući AHP i TOPSIS metodu sa Pitagorinim fazi brojevima. Autor je koristio Pitagorinu fazi AHP metodu za definisanje težina kriterijuma, a Pitagorinu fazi TOPSIS metodu za određivanje najboljeg snabdevača. U razvijenom modelu je posmatrano i ocenjeno pet snabdevača.

### 3. FAZI SISTEM

U ovom radu se predlaže fazi sistem za izbor snabdevača u logističkoj kompaniji. Fazi sistemi se zasnivaju na teoriji fazi skupova i fazi logike. Fazi skupovi su pogodan alat za modeliranje veličina koje je veoma teško kvantifikovati tj. pouzdano opisati odgovarajućim realnim vrednostima. Primena fazi logike omogućava modeliranje odlučivanja eksperata u realnim uslovima [5]. U osnovi fazi sistema je skup pravila zaključivanja kojima se opisuju ishodi odlučivanja u zavisnosti od vrednosti ulaznih veličina.

Postupak razvoja fazi sistema (slika 1) obuhvata:

- Određivanje ulaznih i izlazne veličine i
- Definisane fazi pravila.

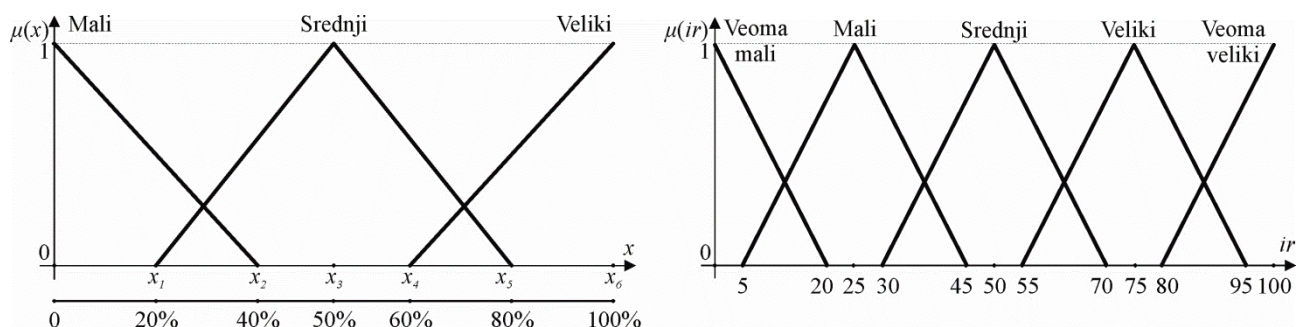


Slika 1: Koraci u razvoju fazi sistema

#### 3.1. Ulazne i izlazne veličine

Razvijeni fazi sistem ima pet ulaznih veličina ( $N = 5$ ) i jednu izlaznu veličinu. Ulazne veličine su kriterijumi odlučivanja: kvalitet ( $K_1$ ), cena ( $K_2$ ), geografska udaljenost ( $K_3$ ), reputacija ( $K_4$ ) i profesionalizam ( $K_5$ ). Relativni značaj kriterijuma  $i$  ( $i = 1, \dots, 5$ ) je  $z_i$  i on pokazuje uticaj u procesu odlučivanja. Sve ulazne veličine se opisuju fazi skupovima: mali ( $X_M$ ), srednji ( $X_S$ ) i veliki ( $X_V$ ). Na osnovu relativnog značaja  $z_i$  ( $0 < z_i \leq 1$ ) određuje se interval vrednosti kriterijuma  $i$ , leva i desna granica i vrednost sa najvećim stepenom pripadnosti fazi skupova  $X_M$ ,  $X_S$  i  $X_V$  prema relacijama:  $x_{i1} = 0.2 \cdot z_i$ ,  $x_{i2} = 0.4 \cdot z_i$ ,  $x_{i3} = 0.5 \cdot z_i$ ,  $x_{i4} = 0.6 \cdot z_i$ ,  $x_{i5} = 0.8 \cdot z_i$  i  $x_{i6} = z_i$ .

Izlazna veličina je indeks rangiranja snabdevača ( $IR$ ) i opisuje se fazi skupovima veoma mali ( $IR_{VM}$ ), mali ( $IR_M$ ), srednji ( $IR_S$ ), veliki ( $IR_V$ ) i veoma veliki ( $IR_{VV}$ ). Vrednost indeksa rangiranja snabdevača može biti u intervalu  $(0, 100)$ . Na slici 2 su prikazane funkcije pripadnosti fazi skupova koji opisuju ulazne i izlaznu veličinu.



Slika 2: Funkcije pripadnosti fazi skupova  $X_M$ ,  $X_S$ ,  $X_V$ ,  $IR_{VM}$ ,  $IR_M$ ,  $IR_S$ ,  $IR_V$  i  $IR_{VV}$

Za jedan skup vrednosti ulaznih veličina izlaz je rezultujući fazi skup na koji se primenjuje postupak defazifikacije. U ovom radu se koristi centroid metoda defazifikacije [7].

### 3.2. Definisane fazi pravila

Skup fazi pravila opisuje sistem zaključivanja eksperata pri donošenju odluke o izboru snabdevača. Jedno fazi pravilo predstavlja fazi frazu definisanu na skupu  $K_1 \times K_2 \times K_3 \times K_4 \times K_5$ . Potpun skup fazi pravila obuhvata  $NP$  pravila:

$$NP = \prod_{i=1}^N NK_i \quad (1)$$

gde je  $N$  broj ulaznih veličina (kriterijuma odlučivanja) a  $NK_i$  broj fazi skupova (nivoa kriterijuma) koji opisuju veličinu  $i$ . U ovom fazi sistemu je  $NP = 243$ .

Fazi pravilo u pretpostavkama sadrži ulazne veličine a u konsekvenci izlaznu veličinu fazi sistema. Vrednost izlazne veličine se određuje na osnovu vrednosti i značaja ulaznih veličina i predstavlja modeliranje ponašanja eksperata u konkretnim uslovima. U fazi pravilima cilj je da se neke ulazne veličine maksimiziraju (kvalitet, reputacija i profesionalizam), a neke minimiziraju (cena i geografska udaljenost) kako bi se postigla što bolja ocena indeksa rangiranja snabdevača. Podaci o ulaznim veličinama su prikazani u tabeli 1. U ovom radu je značaj kriterijuma (ulaznih veličina) dobijen anketiranjem eksperata u posmatranoj kompaniji.

Tabela 1: Karakteristike ulaznih veličina

Kriterijum	$K_1$	$K_2$	$K_3$	$K_4$	$K_5$
Značaj ( $z_i$ )	0.3449	0.2427	0.0724	0.1282	0.2118
Min/Max	Max	Min	Min	Max	Max

Kada je skup fazi pravila relativno mali tada je moguće obezbediti konzistentnost ekspertskog mišljenja pri definisanju vrednosti izlazne veličine. U uslovima kada postoji veći broj fazi pravila, određivanje vrednosti izlazne veličine je dugotrajan posao i teško se može obezbediti da svako fazi pravilo u potpunosti odgovara ekspertskom znanju u realnim uslovima odlučivanja.

U ovom radu je definisan indeks težine ( $IT_j$ ) za svako fazi pravilo  $j$  ( $j = 1, \dots, NP$ ) koji se određuje prema relaciji:

$$IT_j = \sum_{i=1}^N nk_{ij} \cdot z_i \cdot g_i \quad (2)$$

gde je

$$nk_{ij} = \begin{cases} 1, & \text{mali} \\ 2, & \text{srednji, fazi skup ulazne veličine } i \text{ u pravilu } j \\ 3, & \text{veliki} \end{cases}$$

$$g_i = \begin{cases} -1, & \text{ulazna veličina se minimizira} \\ +1, & \text{ulazna veličina se maksimizira} \end{cases}$$

$z_i$  – značaj ulazne veličine  $i$ ,

$N$  – broj ulaznih veličina.

U zavisnosti od indeksa težine  $IT_j$  za svako pravilo se određuje nivo izlazne promenljive tj. indeks rangiranja  $IR_j$  koji može uzeti vrednosti: 1 ( $IR_{VM}$  – veoma mali), 2 ( $IR_M$  – mali), 3 ( $IR_S$  – srednji), 4 ( $IR_V$  – veliki) i 5 ( $IR_{VV}$  – veoma veliki). Indeks težine  $IT_j$  je u intervalu ( $IT^{min}$ ,  $IT^{max}$ ). Postoji više načina za određivanje indeksa rangiranja u svakom pravilu a ovde se posmatraju dva:

- Interval indeksa težine se podeli na pet jednakih segmenata koji odgovaraju vrednostima indeksa rangiranja (1, 2, 3, 4 i 5).
- Interval indeksa težine se podeli na pet segmenata različite dužine tako da nivoi indeksa rangiranja (1, 2, 3, 4 i 5) budu ravnomerno zastupljeni u skupu pravila.

Ovako se formiraju dva skupa fazi pravila tj. dva fazi sistema koji se razlikuju u određivanju vrednosti izlazne promenljive u fazi pravilima. Izlazna promenljiva  $IR1$  se dobija na osnovu istih intervala vrednosti indeksa težine  $IT_j$  pravila  $j$ . U tabeli 2 su prikazani podaci o intervalima indeksa težine i strukturi  $IR1$  u fazi pravilima. Najveći broj pravila (87) ima srednji indeks rangiranja, nešto manje pravila ima mali (61) i veliki (61) indeks rangiranja, a najmanje pravila ima veoma mali (17) i veoma veliki (17) indeks rangiranja.

**Tabela 2:** Struktura  $IR1$  u fazi pravilima

$IR1$		$IT$		Broj pravila	%
1	Veoma mali	-0.2603	0.1397	17	7.00
2	Mali	0.1397	0.5397	61	25.10
3	Srednji	0.5397	0.9397	87	35.80
4	Veliki	0.9397	1.3397	61	25.10
5	Veoma veliki	1.3397	1.7397	17	7.00
Ukupno:				243	100

Drugi skup fazi pravila ima izlaznu promenljivu  $IR2$  koja se dobija na osnovu različitih intervala vrednosti indeksa težine. Širina intervala se određuje tako da nivoi indeksa rangiranja (veoma mali, mali, srednji, veliki i veoma veliki) budu ravnomerno raspoređeni u skupu pravila. U tabeli 3 su dati podaci o intervalima indeksa težine i strukturi  $IR2$  u skupu fazi pravila.

**Tabela 3:** Struktura  $IR2$  u fazi pravilima

$IR2$		$IT$		Broj pravila	%
1	Veoma mali	-0.2603	0.3800	49	20.16
2	Mali	0.3800	0.6350	48	19.75
3	Srednji	0.6350	0.8450	49	20.16
4	Veliki	0.8450	1.1000	48	19.75
5	Veoma veliki	1.1000	1.7397	49	20.16
Ukupno:				243	100

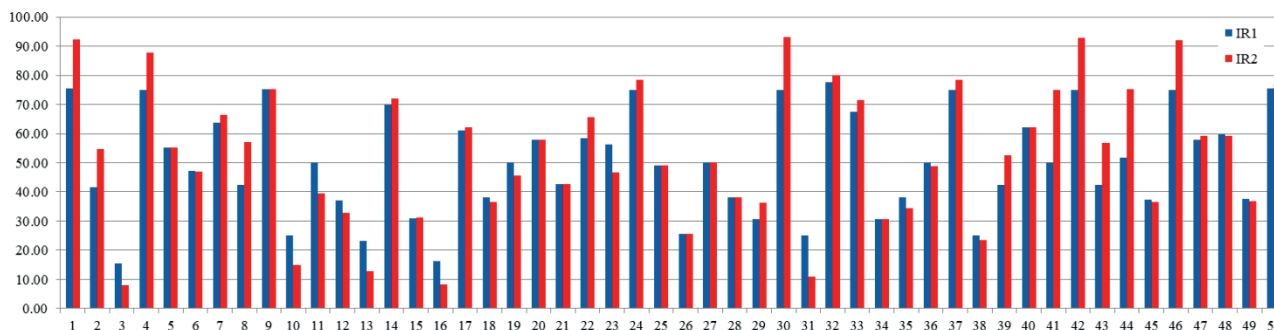
Poređenje opisanih skupova fazi pravila pokazuje da je najveći broj fazi pravila isti u oba skupa (141). Kod 51 pravila je  $IR1 > IR2$  a kod 51 pravila je  $IR1 < IR2$ . Razlika između  $IR1$  i  $IR2$  nikada nije veća od 1 što znači da indeks rangiranja varira između susednih fazi skupova. U tabeli 4 je prikazano prvih 10 fazi pravila i njihovi indeksi rangiranja.

**Tabela 4:** Ilustracija definisanja fazi pravila

Pravilo	$K_1$	$K_2$	$K_3$	$K_4$	$K_5$	$IT$	$IR1$	$IR2$
1	1	1	1	1	1	0.3699	2	1
2	1	1	1	1	2	0.5817	3	2
3	1	1	1	1	3	0.7935	3	3
4	1	1	1	2	1	0.4981	2	2
5	1	1	1	2	2	0.7099	3	3
6	1	1	1	2	3	0.9217	3	4
7	1	1	1	3	1	0.6262	3	2
8	1	1	1	3	2	0.8381	3	3
9	1	1	1	3	3	1.0499	4	4
10	1	1	2	1	1	0.2974	2	1

#### 4. NUMERIČKI PRIMER

Opisani skupovi fazi pravila – dva fazi sistema su testirani na većem broju ulaznih podataka. Snabdevači su ocenjivani po svakom kriterijumu i sa tim ulaznim podacima su testirani fazi sistemi. Ocenjivanje snabdevača je urađeno na način kako to rade eksperti u posmatranoj kompaniji. Poređenje vrednosti indeksa rangiranja  $IR1$  i  $IR2$  za 50 snabdevača je prikazano na slici 3.



Slika 3: Poređenje  $IR1$  i  $IR2$  za sve snabdevače

Za prvi fazi sistem  $IR1 \in [15.40; 77.64]$  a za drugi fazi sistem  $IR2 \in [8.05; 92.99]$ . Vrednosti indeksa rangiranja  $IR1$  i  $IR2$  se kod većine snabdevača razlikuju (41) dok su kod nekih iste (9). Kod bolje rangiranih snabdevača je  $IR2 > IR1$ , dok je kod lošije rangiranih  $IR1 > IR2$ .

U tabeli 5 je prikazano 10 najbolje rangiranih snabdevača po oba indeksa rangiranja. Redosled snabdevača je različit ali se 9 istih snabdevača nalazi u ovom skupu. Detaljnija analiza pokazuje da drugi fazi sistem daje indekse rangiranja koji više odgovaraju odlučivanju eksperata u realnim uslovima.

Tabela 5: Rangiranje 10 najboljih snabdevača

Prvi fazi sistem			Drugi fazi sistem		
Rang	Snabdevač	$IR1$	Rang	Snabdevač	$IR2$
1	S32	77.64	1	S30	92.99
2	S1	75.54	2	S42	92.85
3	S50	75.41	3	S1	92.32
4	S9	75.18	4	S46	91.95
5	S42	75.00	5	S4	87.78
6	S24	75.00	6	S32	79.99
7	S37	75.00	7	S24	78.38
8	S46	75.00	8	S37	78.38
9	S4	75.00	9	S50	76.75
10	S30	75.00	10	S44	75.28

#### 5. ZAKLJUČAK

Za kompanije je od velikog značaja kvalitetno odlučivanje o izboru snabdevača. U praksi se primenjuju različiti modeli a češće je to posledica ekspertske odlučivanja na osnovu njihovog znanja, ličnog osećaja i iskustva. U ovom radu je primenjena teorija fazi skupova i fazi logike kako bi se razvio fazi sistem za izbor snabdevača.

Osnovni cilj rada je bio razvoj postupka za definisanje fazi pravila koja na što bolji način modeliraju ponašanje eksperata u realnim uslovima. Kada postoji veći broj fazi pravila u sistemu odlučivanja potrebno je obezbediti konzistentnost izlazne veličine koja odgovara ekspertske razmišljanju i zaključivanju. U radu je definisan indeks težine svakog fazi pravila koji pomaže u definisanju fazi vrednosti izlazne veličine. Predloženi postupak je testiran na velikom skupu snabdevača i pokazao je dobre rezultate.

## NAPOMENA

Ovaj rad je finansiran od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije u okviru projekata TR 36005 i TR 36006 za period 2011-2022.

## LITERATURA

- [1] Çalik, A., (2020). A novel Pythagorean fuzzy AHP and fuzzy TOPSIS methodology for green supplier selection in the Industry 4.0 era. *Soft Computing* 25(3), 1-13.
- [2] Fagundes, M.V.C., Hellingrath, B., & Freires, F.G.M. (2021). Supplier Selection Risk: A New Computer-Based Decision-Making System with Fuzzy Extended AHP. *Logistics*. 5 (1), 1-17.
- [3] Haddad, A.N., Da Costa, B.F., De Andrade, L.S., Hammad, A., & Soares, C.A.P. (2021). Application of Fuzzy-TOPSIS Method in Supporting Supplier Selection with Focus on HSE Criteria: A Case Study in the Oil and Gas Industry. *Infrastructures*. 6(8), 1-16.
- [4] Jain, N., Singh, A.R., & Upadhyay, R.K., (2020). Sustainable supplier selection under attractive criteria through FIS and integrated fuzzy MCDM techniques. *International Journal of Sustainable Engineering*, 1-22.
- [5] Mitrović, M., Radivojević, G., & Popović, D. (2021). Fuzzy Model for Supplier Selection, *Proceedings of the XLVIII International Symposium on Operational Research – XLVIII SYM-OP-IS 2021*, Banja Koviljača 20-23 September 2021, Editors: Urošević, D., Dražić, M., Stanimirović, Z., pp. 355-360, ISBN 978-86-7589-151-2, Banja Koviljača, Serbia.
- [6] Puška, A., Korarević, S., Stević, Ž., & Stovrag, J. (2018). A new way of applying interval Fuzzy logic in group decision making for Supplier Selection. *Economic Computation and Economic Cybernetics Studies and Research*. 52, 217-234.
- [7] Zimmermann, H. J. (1996). *Fuzzy Set Theory and Its Applications*. London: Kluwer Academic Publishers.



## DINAMIČKI DISKRETNI MODEL ZA UPRAVLJANJE ZALIHAMA U REŽIMU MULTISTOP ISPORUKA ZA DVA SKLADIŠTA

### DYNAMIC DISCRETE INVENTORY CONTROL MODEL IN MULTISTOP DELIVERY MODE FOR TWO WAREHOUSES

SLOBODAN ANTIĆ<sup>1</sup>, LENA ĐORĐEVIĆ MILUTINOVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, slobodan.antic@fon.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, lena.djordjevic.milutinovic@fon.bg.ac.rs

**Rezime:** U radu je predstavljen upravljački model za kontrolu zaliha u industriji nameštaja. Cilj rada je razvoj dinamičkog diskretnog upravljačkog modela za kontrolu zaliha za praćenje isporuka na dva skladišta, od kojih je jedno skladište u proizvodnji, a drugo skladište van lokacije proizvodnje. Upravljački model proračunava količine pločaste iverice iz kamionskih isporuka za skladište u proizvodnji i za rezervno skladište. Zbog toga, razvijen je dinamički diskretni matematički model objekta diskretnog upravljanja s pripadajućim zakonom ponašanja posmatranog upravljačkog sistema. Posmatrani upravljački model treba da obezbedi unapređenje kontrole zaliha pločastih materijala, s ciljem da količina pločastih materijala u „proboju“ zaliha bude minimizirana.

**Ključne reči:** Upravljanje zalihama, Upravljački model, Diskretni objekat upravljanja, Isporuka, proizvodnja

**Abstract:** This paper presents the inventory control model in furniture industry. The purpose of this paper is to present a dynamic discrete inventory control model to track deliveries for two warehouses, where one warehouse is located in production site and the other warehouse is located outside of the production site. The control model enables the calculation of quantities for delivery of wooden plates from truck for production warehouse and for reserve warehouse. Therefore, a dynamic discrete mathematical model of a discrete control object with the corresponding law of behavior of the observed control system has been developed. The observed control model should improve of inventory control of wooden plates, in order to minimize the "breakthrough" of inventory of wooden plates..

**Keywords:** Inventory control, Control model, Discrete time control object, Delivery, Production.

## 1. UVOD

Značaj donošenja upravljačkih odluka u procesima upravljanja zalihama se ogleda u prepoznavanju nesklada između potreba i dostupnosti repromaterijala da bi se proizvodnja odvijala neprekidno [1]. Postoji veliki broj studija koji se odnose na modele upravljanja zalihama i transportne modele. Međutim, manji je broj studija koji podrazumevaju integraciju ove dve ključne aktivnosti u lancima snabdevanja. Ovakve studije podrazumevaju koncept ITI modela (Integrated Transportation-Inventory models), kao što su modeli: integrisani model rutiranja [3], integrisani model zaliha u globalnim lancima snabdevanja [6], integrisani modeli zelenih lanaca snabdevanja [17] i integrisani modeli upravljanja zalihama i transportom [12]. Generalno posmatrajući, tipovi modela analizirani u radovima autora se mogu klasifikovati u četiri grupe modela: uopšteni ITI modeli, ITI modeli sa promenljivim količinama za naručivanje (Lot sizing models), ITI modeli sa rutiranjem distribucije i ITI modeli s specijalnim primenama, navedeno u [15]. Autori Williams i Tokar [18] predstavili su studiju u kojoj se analiziraju modeli zaliha sa fokusom na kolaboraciju s drugim procesima u preduzeću (transport i skladištenje). U radu se razmatraju modeli s stohastičkom tražnjom u upravljanju zalihama oslanjajući se na upravljačke promenljive zasnovane na količini zaliha koja se transportuje. U radovima navedenim pod [5,16] predstavljeni su modeli u kojima se upravljanje zalihama vrši iz direktnih isporuka u transportu i iz isporuka u modelu distribucije putem pretovara (Cross docking). Model upravljanja zalihama u distribuciji putem pretovara s jednog vozila u više vozila (Cross docking) je analiziran u radu [11], dok je pretovar zaliha s jednog vozila na jedno vozilo analiziran u radovima od [7, 14]. Problem trgovačkog putnika, kao transportni problem uključuje više isporuka (multistop isporuke) na već planiranoj ruti distribucije u cilju minimiziranja pređenog puta. Problem isporuka na kružnoj ruti distribucije s više istovarnih mesta u multistop režimu isporuka (Milk run delivery route) je analiziran u radovima [4, 8, 9]. Bilo kako, ruta isporuke nije neophodno uslovljena razdaljinom ili pređenim putem, veći količinom prevezenih

zaliha od istovarnog mesta do istovarnog mesta, što je jedna od pretpostavki za analizu problema isporuka zaliha u ovom radu. U radu navedenom u [10], prikazan je integrisani model planiranja isporuka kupcima koji zavisi od količine zaliha u otpremnom skladištu i zaliha na lokaciji kupca, kao i od planiranih ruta isporuke za vozila. Modelom se definišu količine i intervali za obnavljanje zaliha u otpremnom skladištu, kao i količina za isporuku na planiranim rutama isporuke prema lokacijama kupaca [10]. Upravljački modeli za kontrolu zaliha predstavljaju uprošćenu predstavu sistema za upravljanje zalihama određene kompanije. Prema načinu za donošenje odluka, upravljački modeli mogu se podeliti, prema [13] na: optimizacioni modeli, simulacioni modeli; kontinualno promenljivi modeli; modeli fiksiranog vremena i modeli diskretnih događaja. Takođe, analitički spredšit modeli (ASM) predstavljaju primenu matematičkih modela u spredšitu, s ciljem primene istih u poslovanju [2].

Analizom literature navedene u uvodnom delu rada, može se zaključiti da je nemoguće izvršiti upoređivanje modela prikazanog u radu s modelima prikazanim u literaturi. Može se zaključiti da posmatrani model pripada grupi modela upravljanja zalihama u režimu multistop isporuka, ali kao takav ima određene specifičnosti koje nisu navedene u analiziranim radovima u literaturi. U narednom poglavlju biće predstavljen dinamički diskretni upravljački model za kontrolu zaliha u skladištima na više različitim lokacije u sistemu multi-stop isporuka sa matematičkim modelom objekta diskretnog upravljanja, Potrebno je doneti odluke o nivou dopunjavanja zalihama iverice dva skladišta na dve različite lokacije iz jedne šleperske isporuke. Multistop isporuke se ogleđaju u potrebi racionalizacije procesa isporuka, u kome se tokom isporuke kamion istovara na dva ili više istovarnih mesta u zavisnosti u kom skladištu je potrebna dopuna zaliha.

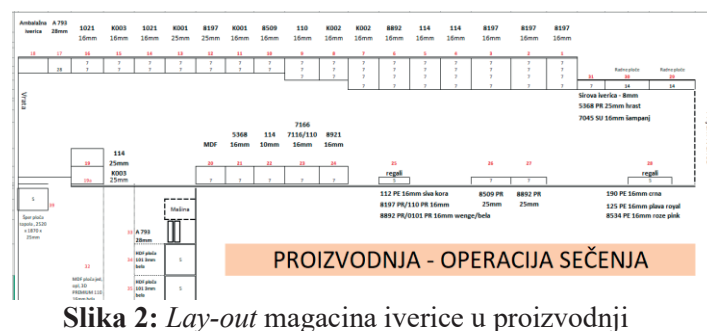
## 2. OPIS UPRAVLJAČKOG MODELA ZA KONTROLU ZALIHA U SKLADIŠTIMA NA VIŠE RAZLIČITIH LOKACIJA U REŽIMU MULTISTOP ISPORUKA

Proces proizvodnje u proizvodnji pločastog nameštaja se sastoji od četiri osnovne operacije: sečenje iverice, kantovanje ploča, bušenje ploča iverice za montiranje okova i pakovanje više elemenata pločaste iverice u kartonski paket. Kako ovaj niz operacija predstavlja jedan kontinualan niz, neophodno je obezbediti da se sve operacije u proizvodnji izvode bez kašnjenja. Međutim, u stvarnosti usled nedostatka pločaste iverice, dolazi do zastoja u operaciji sečenja na rezačima, a onda do zastoja na ostalim operacijama proizvodnje. Uzroci koji su doveli do pojave kašnjenja u sečenju pločaste iverice se mogu podeliti na: eksterne (kašnjenje nabavke u poručivanju iverice; kašnjenje proizvođača iverice; kašnjenje prevoznika u isporukama) i interne (greške u analizi potrebnih količina za skladišta; nedostatak analitičkih alata za upravljanje zalihama; nedostatak prostora u skladištu iverice u proizvodnji). Operacija sečenja u proizvodnji pločastog nameštaja ima obavezu da svaka 3 dana ( $T=3$ ) iseče u zadatoj dinamici, prema definisanom trodnevnom planu proizvodnje sledeće količine iverice, pogledati ispod sliku 1.

Plan sečenja za period:		OD	DO	ponedeljak			utorak	sreda		
		12.feb.18	14.feb.18	Secenje	Secenje	Secenje	Secenje			
Sifra materijala	materijal	Tip	12. feb.	13. feb.	14. feb.	Total sece	Averag			
31001772	Opl, iverica 8197 PR 16mm hrast sonoma	Iverica	22,67	26,10	27,00	331,73	24,69			
31001694	Opl, iverica 110 SM 16mm bela	Iverica	1,60			94,20	4,59			
31001657	HDF ploča 101 3mm bela	HDF	2,44	0,50	0,48	17,77	1,31			
31001708	Opl, iverica 8892 PR 16mm wenge	Iverica	5,93	16,09	16,43	111,69	7,89			
31005088	Opl, iverica K002 PW 16mm sivi zanatski hrast	Iverica	5,57	2,10	1,87	93,26	6,62			

Slika 1: Trodnevni plan sečenja iverice

Zalihe pločaste iverice koje se koriste prema zadatom planu sečenja se smeštaju u skladište proizvodnje, koje ima ograničen kapacitet izražen u broju buntova pločaste iverice (bunt sadrži 30 ploča iverice) na određenoj skladišnoj lokaciji za određeni artikal pločaste iverice, pogledati ispod sliku 2.



Slika 2: Lay-out magacina iverice u proizvodnji

Dotatna količina zaliha pločaste iverice se skladišti u eksternom rezervnom skladištu zbog ograničenog kapaciteta skladišta proizvodnje. Služba nabavke unapred dogovara isporuke pločaste iverice od proizvođača i dostavlja plan nabavke za narednih sedam dana, koji vremenski obuhvata dva plana proizvodnje. Problem nastaje u trenutku kada je potrebno odlučiti koja isporuka i količina iz određene isporuke se isporučuje u dva skladišta, jer dobavljač nema informaciju o deljenju količine za isporuku po skladištima, a često se dešava da proizvođač pošalje kamione s ivericom ne poštujući prosleđeni redosled plana isporuka po danima, pogledati ispod sliku 3.

Plan isporuka za period:		Od	Do								Total	Average
		12.feb.18	14.feb.18	ponedeljak	utorak	sreda	četvrtak	petak	subota	nedelja		
Br.	Opis	Tip	12.2.2018	13.2.2018	14.2.2018	15.2.2018	16.2.2018	17.2.2018	18.2.2018			
31001772	Opl, iverica 8197 PR 16mm hrast sonoma	Iverica	48,00	47,00				14,00		425,00	25,39	
31001694	Opl, iverica 110 SM 16mm bela	Iverica						5,00		294,00	19,42	
31001657	HDF ploča 101 3mm bela	HDF	1,00	3,00	10,00					28,00	1,75	
31001708	Opl, iverica 8892 PR 16mm wenge	Iverica				27,00		4,00		164,00	7,92	
31005088	Opl, iverica K002 PW 16mm sivi zanatski hrast	Iverica		14,00	17,00					12,00	171,00	9,29

Slika 3: Plan isporuka pločaste iverice za sedam dana

Potrebno je napraviti plan podele isporuke pločaste iverice sa ciljem da se ostvari plan sečenja iverice tako da razlika između potrebnih količina iverice i isporučениh količina iverice u magacinu proizvodnje bude minimalna u naredna tri vremenska perioda. Upravljački model se može predstaviti pomoću algoritma odlučivanja, prema kome se prvo popune zalihe iverice magacina proizvodnje s potrebnom pločastom ivericom prema planu sečenja mašina rezača i prema kapacitetima slobodnih skladišnih lokacija u magacinu, pogledati ispod slika 4.

ULAZNA KOLIČINA PLOČASTE IVERICE U MAGACIN PROIZVODNJE:	
IF (Kapacitet skladišne lokacije artikla $\geq$ Stanja zaliha artikla u magacinu proizvodnje) THEN	
IF (Količina za dopunu lokacije artikla (Razlika između kapaciteta lokacije i stanja zaliha na lokaciji) $\geq$ Količine iz najavljene isporuke artikla) THEN	
Količina iz najavljene isporuke artikla	
Else	
Količina za dopunu lokacije artikla (Razlika između kapaciteta lokacije i stanja zaliha na lokaciji)	
End IF.	
Else	
Nula	
End IF.	

Slika 4: Pseudokod algoritma popune skladišta proizvodnje iz isporuka dobavljača

Onda se vrši dopuna zaliha pločaste iverice na lokaciji rezervnog magacina iverice s preostalom ivericom s kamiona dobavljača, pogledati ispod sliku 5.

ULAZNA KOLIČINA PLOČASTE IVERICE U REZERVNI MAGACIN (RM):	
Količina iz najavljene isporuke artikla - ULAZNA KOLIČINA PLOČASTE IVERICE U MAGACIN PROIZVODNJE	

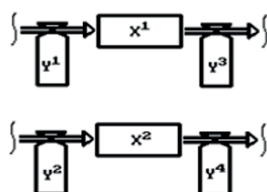
Slika 5: Pseudokod algoritma popune rezervnog skladišta iz isporuka dobavljača

Na kraju se vrši prebacivanje nedostajuće pločaste iverice s rezervnog magacina iverice na magacin proizvodnje, pogledati ispod sliku 6.

ULAZ PLOČASTE IVERICE REZERVNOG MAGACINA (RM) U MAGACIN PROIZVODNJE:	
IF (Stanje zaliha artikla u RM = 0) THEN	
Nula	
Else	
IF (ULAZNA KOLIČINA PLOČASTE IVERICE U REZERVNI MAGACIN (RM) > 0) THEN	
Nula	
Else	
IF (Zbir potrebnih količina za tri dana sečenja < Stanja zaliha artikla u magacinu proizvodnje + ULAZNA KOLIČINA PLOČASTE IVERICE U MAGACIN PROIZVODNJE) THEN	
Nula	
Else	
Zbir potrebnih količina za tri dana sečenja - (Stanja zaliha artikla u magacinu proizvodnje + ULAZNA KOLIČINA PLOČASTE IVERICE U MAGACIN PROIZVODNJE)	
End IF.	
End IF.	
End IF.	

Slika 6: Pseudokod algoritma popune skladišta proizvodnje iz rezervnog skladišta

Kada govorimo o identifikaciji tokova u sistemu, moguće je uočiti dva materijalna toka za dva skladišta, čiji su subjekti tokova pločasta iverica. Znači, iako je subjekat toka isti za oba toka, subjekat toka se nalazi na dve različite lokacije, pa se iz tog razloga formiraju dva materijalna toka. Prvi materijalni tok se odnosi na skladište iverice u proizvodnji, dok se drugi materijalni tok odnosi na rezervno skladište, pogledati sliku 7.



Slika 7: Materijalni tokovi iverice za magacin proizvodnje i rezervni magacin

Faze na prvom materijalnom toku za skladište iverice u proizvodnji su: ulazna količina zaliha iverice iz isporuke dobavljača, količina zaliha iverice na skladištu proizvodnje, lansirana količina zaliha iverice u proizvodnju. Respektivno, faze na drugom materijalnom toku su iste za rezervno skladište zaliha iverice.

Zakon ponašanja diskretnog objekta upravljanja predstavljen je u relaciji (1):

$X_t^1 =$  poznato,

$X_t^2 =$  poznato,

$$X_t^1 = X_{t-1}^1 + Y_t^1 + Y_t^4 - Y_t^3,$$

$$X_t^2 = X_{t-1}^2 + Y_t^2 - Y_t^4, \quad t=1,2,3 \quad (1)$$

Pri čemu su:

- $X_t^1$ - stanje zaliha u magacinu proizvodnje na kraju vremenskog perioda  $t$ ,
- $X_t^2$ - stanje zaliha u rezervnom magacinu na kraju vremenskog perioda  $t$ ,
- $Y_t^1$ - ulazna količina zaliha iverice iz isporuke dobavljača u magacin proizvodnje u periodu  $t$ , koja predstavlja složenu funkciju u zavisnosti od potrebne količine za dopunu zaliha,
- $Y_t^2$ - ulazna količina zaliha iverice iz isporuke dobavljača u rezervni magacin u periodu  $t$ , koja predstavlja složenu funkciju u zavisnosti od potrebne količine za dopunu zaliha,
- $Y_t^3$ - lansirana količina zaliha iverice u proizvodnju u periodu  $t$ ,
- $Y_t^4$ - izlazna količina zaliha iverice iz rezervnog magacina u vidu dopune u skladište proizvodnje u periodu  $t$ , koja predstavlja složenu funkciju u zavisnosti od potrebne količine za dopunu zaliha.

Potrebno je primetiti da u zakonu ponašanja za skladište proizvodnje ulaznu količinu u akumulaciju čine da toka  $Y_t^1 + Y_t^4$ .

Identifikovane su sledeće okolnosti:

- $p_t^1$ - planirana količina za sečenje iverice u toku vremenskog perioda  $t$ ,
- $p_t^2$ - planirana količina za najavljenju isporuku iverice u toku vremenskog perioda  $t$ ,
- $p_t^3$ - kapaciteti skladišnih lokacija pojedinačnih artikala iverice u magacinu proizvodnje u toku vremenskog perioda  $t$ .

Može se uvesti sledeća pomoćna pomenljiva:

- $S_t^1$ - prosečna dnevna količina iverice isečena u prethodnim vremenskim periodima ( $t-1$ ,  $t-2$ ,  $t-3$ ).

Vreme koje protekne između naručivanja i isporuke (Lead time - LT) je 1 dan, pa se kašnjenje u modelu može zanemariti, jer su diskretni vremenski periodi definisani u danima, pri čemu  $t=1,2,3,\dots,31$ . Ulazna količina zaliha iverice iz isporuke dobavljača u magacin proizvodnje u periodu  $t$  ( $Y_t^1$ ), predstavljena je pseudokodom algoritma popune skladišta proizvodnje iz isporuka dobavljača (Slika.4) i može se prikazati sledećom matematičkom relacijom, pogledati ispod relacija (2). Relacija (4) pokazuje da količinu za dopunu skladišta proizvodnje ( $Y_t^1$ ) dobijamo ukoliko je kapacitet skladišnih lokacija  $p_t^3$  veći od stanja zaliha magacina proizvodnje  $X_{t-1}^1$ , a količinu za dopunu dobijamo iz veličine iznosa razlike kapaciteta skladišnih lokacija  $p_t^3$  i stanja zaliha magacina proizvodnje  $X_{t-1}^1$ , tako da skladište u proizvodnji možemo dopuniti s celom isporukom iz kamiona  $p_t^2$  ili samo s delom isporuke iz kamiona.

$$Y_t^1 = \left\{ \left\{ \begin{array}{ll} p_t^2, & (p_t^3 - X_{t-1}^1) \geq p_t^2 \\ p_t^3 - X_{t-1}^1, & \text{u suprotnom} \end{array} \right\}, \quad p_t^3 \geq X_{t-1}^1 \right\}, t = 1, 2, 3 \quad (2)$$

Ulazna količina zaliha iverice iz isporuke dobavljača u rezervni magacin u periodu  $t$  ( $Y_t^2$ ), predstavljena je pseudokodom algoritma popune rezervnog skladišta iz isporuka dobavljača (Slika.5) i može se prikazati sledećom matematičkom relacijom, pogledati ispod relacija (3). Relacija (3) pokazuje da količinu za dopunu rezervnog skladišta dobijamo kada od planirane količine za najavljenju isporuku iverice u toku vremenskog perioda  $t$  ( $p_t^2$ ) oduzmemo količinu s kamiona kojom je dopunjeno skladište proizvodnje  $Y_t^1$ .

$$Y_t^2 = p_t^2 - Y_t^1, \quad t = 1, 2, 3 \quad (3)$$

Lansirana količina zaliha iverice u proizvodnju u periodu  $t$  ( $Y_t^3$ ) jednaka je planiranoj količini za sečenje iverice iz plana proizvodnje za operaciju sečenja u toku vremenskog perioda  $t$  ( $p_t^1$ ). Može se predstaviti sledećom relacijom (4):

$$Y_t^3 = p_t^1, \quad t = 1, 2, 3 \quad (4)$$

Izlazna količina zaliha iverice iz rezervnog skladišta u vidu dopune u skladišta proizvodnje u periodu  $t$  ( $Y_t^4$ ), uvećava stanje zaliha skladišta proizvodnje  $X_t^1$  i predstavljena je pseudokodom algoritma popune skladišta proizvodnje iz rezervnog skladišta (Slika.6) i može se prikazati sledećom matematičkom relacijom, pogledati ispod relacija (5). Relacija (5) pokazuje da količinu za dopunu skladišta proizvodnje iz rezervnog skladišta dobijamo ako je količina za trodnevni plan sečenja veća od stanja zaliha skladišta proizvodnje  $X_{t-1}^1$  uvećanog za ulaznu količinu zaliha iverice iz isporuke dobavljača u magacin proizvodnje u periodu  $t$  ( $Y_t^1$ ),

$$Y_t^4 = \left\{ \left\{ \begin{array}{ll} 0, & X_{t-1}^2 = 0 \\ 0, & Y_t^2 > 0 \end{array} \right\}, \text{u supr.} \right\}, t = 1, 2, 3 \quad (5)$$

Pri čemu je pomoćna promenljiva predstavljena kao prosečna dnevna količina iverice koja je isečena u prethodnim vremenskim intervalima ( $t-1$ ,  $t-2$ ,  $t-3$ ), pogledati relaciju (6) ispod.

$$S_t^1 = \frac{\sum_{-t} p_t^1}{3}, \quad t = 1, 2, 3 \quad (6)$$

Kada govorimo razvijanju ograničenja na materijalnim tokovima, iz verbalnog modela upravljačkog problema možemo izvesti više ograničenja za postavljene akumulacije (skladišta). Međutim, kako se određena ograničenja mogu smatrati prirodnim uslovima nenegativnosti, koja korisnik u toku profesionalnog rada neće narušiti, moguće je izvesti samo jedno važno ograničenje za kapacitet skladišta magacina proizvodnje, koje ima direktan uticaj na ostvarenje plana proizvodnje. Uslov nenegativnosti zaliha u vremenskom periodu  $t$  za skladište proizvodnje bi glasilo da u zadatom vremenskom intervalu  $T=3$  nije moguće u skladište proizvodnje uneti veću količinu sirovine nego što je slobodan kapacitet skladišta izražen u broju buntova iverice po skladišnim lokacijama na početku vremenskog perioda  $t$  (odnosno na kraju prethodnog vremenskog perioda  $t-1$ ), pogledati ispod relacija (7).

$$0 \leq p_t^3 - X_{t-1}^1, \quad t=1, 2, 3. \quad (7)$$

Uslov nenegativnosti zaliha za rezervno skladište iverice u matematičkom modelu upravljačkog problema možemo zanemariti, jer je kapacitet rezervnog skladišta znatno veći od projektovanih mesečnih potreba proizvodnje, pa se pomenuti uslov neće razvijati u matematičkom modelu posmatranog upravljačkog problema. Što se tiče posmatranog upravljačkog problema neophodno je utvrditi plan podele isporuke iverice za skladište proizvodnje i rezervno skladište, kako bi se obezbedila planirana dinamika operacije sečenja iverice u proizvodnji uz minimalnu razliku između potrebnih količina iverice iz plana sečenja iverice i

obezbeđenih količina iverice na kraju vremenskog intervala  $T$ . Ciljni funkcija se može prikazati relacijom (8) navedenoj ispod.

$$(\min J) = \sum_{t=0}^3 p_t^1 - X_t^1, t = 1,2,3 \quad (8)$$

### 3. ZAKLJUČAK

U budućim pravcima razvoja ovog upravljačkog modela neophodno je pomenuti upravljački model razviti za veći broj skladišta s dodatnim ograničenjima skladišnog prostora i pokušati primeniti određeni heuristički algoritam u cilju pronalaženja dovoljno dobrih dopustivih rešenja.

### LITERATURA

- [1] Antić, S. (2014). Modeli i metode upravljanja zalihama zasnovani na metaheuristikama, *Doktorska disertacija*, Fakultet organizacionih nauka, Beograd.
- [2] Antić, S., Đorđević, L. (2018). Upravljački modeli i aplikacije u spredšitovima. *Finansijski menadžment, kontrola i menadžersko računovodstvo*, 234-267, Fakultet organizacionih nauka, Beograd.
- [3] Baita, F., Ukovich, W., Pesenti, R., Favaretto, D. (1998). Dynamic routing-and-inventory problems: a review. *Transportation Res. Part A: Policy Practice 1998*, 32(8): 585–98.
- [4] Bertazzi, L., Bosco, A., Laganà, D. (2015). Managing stochastic demand in an inventory routing problem with transportation procurement. *Omega 2015*, 56: 112–21.
- [5] Berman, O., Wang, Q. (2006). Inbound logistic planning: minimizing transportation and inventory cost. *Transp Sci 2006*, 40(3): 287–99.
- [6] Bookbinder, H., Matuk, A. (2009). Logistics and transportation in global supply chains: review, critique, and prospects. *INFORMS, Deci Technol Appl 2009*, 182–211,
- [7] Carlsson, D., Flisberg, P., Rönnqvist, M. (2014). Using robust optimization for distribution and inventory planning for a large pulp producer. *Comput Oper Res 2014*, 44: 214–25.
- [8] Çetinkaya, S., Lee, CY. (2000). Stock replenishment and shipment scheduling for vendormanaged inventory systems. *Manag Sci 2000*, 46(2): 217–32.
- [9] Çetinkaya, S., Mutlu, F., Lee, CY. (2006). A comparison of outbound dispatch policies for integrated inventory and transportation decisions. *European Journal of Operation Research 2006*, 171(3): 1094–112.
- [10] Chandra, P. (2014). A Dynamic Distribution Model with Warehouse and Customer Replenishment Requirements, *Journal of the Operational Research Society*, 44(7): 681-692, DOI:10.1038/sj/jors/0440705
- [11] Chan, A., Muriel, A., Shen, M., Simchi-Levi, D., Teo, P. (2002). Effective zero-inventory ordering policies for the single-warehouse multi-retailer problem with piecewise linear cost structures. *Manag Sci 2002*, 48(11): 1446–60.
- [12] Engebretsen, E., Dauzere-Peres, S. (2018). Transportation mode selection in inventory models: a literature review. *European Journal of Operation Research*, <https://doi.org/10.1016/j.ejor.2018.11.067>.
- [13] Kostić, K. (2008). Izrada i korišćenje poslovnih modela. Fakultet organizacionih nauka, Beograd.
- [14] Mogale, G., Dolgui, A., Kandhway, R., Kumar, K., Tiwari, K. (2017). A multi-period inventory transportation model for tactical planning of food grain supply chain. *Comput Ind Eng 2017*, 110: 379–94.
- [15] Mosca, A., Vidyarthi, N., Satir, A. (2019). Integrated transportation – inventory models: A review, *Operations Research Perspectives*, Volume 6, 2019, 100101, <https://doi.org/10.1016/j.orp.2019.100101>
- [16] Musa, R., Arnaout, JP., Jung, H. (2010). Ant colony optimization algorithm to solve for the transportation problem of cross-docking network. *Comput Ind Eng 2010*, 59(1): 85–92.
- [17] Srivastava, K. (2007). Green supply chain management: a state-of-the-art literature review. *International Journal of Management Rev 2007*, 9(1): 53–80.
- [18] Williams, B., Tokar T., (2008). A review of inventory management research in major logistics journals: themes and future directions. *The International Journal of Logistics Management 2008*, 19(2): 212–32.

## EVOLVEMENT OF BLOCKCHAIN IN SUPPLY CHAINS

ANĐELA ĐORĐEVIĆ<sup>1</sup>

<sup>1</sup> University of Belgrade – Faculty of Organizational Sciences andjela.djordjevic@fon.bg.ac.rs

**Abstract:** *This paper thoroughly explains evolving Blockchain technology, focusing on its integration with supply chains. Many advantages and benefits of this integration, that are supported by real use cases are presented in this paper. Blockchain enables trust and full transparency, which are the two biggest supply chain problems. With the comparative analysis featured in this paper, growth in a number of companies that are actively developing Blockchain solutions is shown. Most common issues with logistics services during the COVID-19 pandemic, which were the result of the conducted research, can be solved with Blockchain. Some of the Blockchain solutions from different industries, along with the problems they solve, are shown.*

**Keywords:** *Blockchain, Supply Chains, Logistics Services, Transparency*

### 1. INTRODUCTION

In today's business environment, sellers and buyers are generally located at great distances, even in other countries. Delivery of products to distant customers includes various means of transport, a lot of paperwork, and many other participants in the supply chain. Given that participants in the supply chain generally do not know each other, the lack of trust is also a major problem. One of the most reoccurring technologies in Industry 4.0 is Blockchain. Industry 4.0 is a digital transformation of production, connected industries, and value-creating processes [8]. Blockchain is a distributed, digital ledger of transactions that can not be altered because of the use of cryptographic methods [12]. Using Blockchain, trust is enabled. Even though logistic companies use automated processes and operations, human errors are still occurring. To avoid that, an integrated digital footprint from beginning to end is a must for supply chains. Taking into consideration that demand is changing from day to day, customers are more demanding, the market is more complex. Thus, supply chains are larger and more complex. The larger the supply chain, i.e., the greater the number of members, management, and transparency are more difficult to achieve. Current solutions and software on the market are not enough for overcoming these issues. This is where the most modern solution on the market, i.e., Blockchain technology takes place.

Blockchain's first applications were in the finance sector. Blockchain applications in logistics and supply chains, related challenges, and advantages are going to be presented in this paper.

This paper is structured in the following manner. In Section 2, Blockchain technology with its advantages and challenges is thoroughly explained. How Blockchain can be used in supply chains and what are the limitations, is shown in Section 3. In the first part of Section 4, companies' stances from different sectors on Blockchain technology are presented. The results of the conducted research about individuals' experiences with logistics services during the COVID-19 pandemic are presented in the second part of Section 4. It is also shown how Blockchain can help in overcoming the problems that occurred during the pandemic with logistics services. Section 5 summarizes the applications of Blockchain in practice and the problems it solves. Finally, a conclusion and discussion about further research are given in Section 6.

### 2. BLOCKCHAIN

Blockchain technology is an efficient, cost-effective, reliable, safe, and secure system for conducting and recording financial transactions. It was developed in order to overcome current transaction systems' limitations. Blockchain can record any type of transaction and track the movement of any asset, whether tangible or intangible. Blockchain network has four key characteristics [18]: 1) consensus, 2) provenance, 3) immutability, and 4) finality. The transaction is only valid if all network members agree. Network members know where the asset came from, and how its ownership has changed over time. Once the transaction is recorded in the ledger, no one can change it or delete it. If a recorded transaction is in error, a new transaction must be made in order to delete the error. In that case, both transactions are visible. A unique, shared ledger is the only place for determining the ownership of an asset or the completion of a transaction.

Blockchain is a distributed database that contains records of transactions shared among members. The Blockchain system is decentralized because the network is run entirely by its members, without relying on a central authority or a centralized infrastructure that establishes the trust [15]. Blockchain technology got its name for the way that it stores data (in the blocks that are connected, forming a chain). Advantages of Blockchain are [9]:

- trust,
- openness,
- independence,
- speed and
- effectiveness.

Before any data is added to Blockchain, it must first be approved by the majority of members. A more reliable and trustworthy system is created when the majority of users take part in the writing, creation, and alteration of such data. Since the data is recorded on the same shared platform, all transactions are visible to all members in real-time. Blockchain design and architecture are constructed in a way that this technology is not dependent on any financial institution. Because of this, Blockchain is less prone to regulations.

Access to Blockchain is anonymous and free. Blockchain uses a jointly generated electronic time stamp that all members trust, even if they do not trust each other. In this manner, it is easy to verify the origin and accuracy of the information no matter the source [15].

Blockchain technology is based on a centralized, computer architecture and because of that, it won't fail due to one mistake (failure).

Transactions that are conducted using Blockchain are publicly available and can not be altered. According to this, their integrity, transparency, and immutability are guaranteed.

Regulatory rules, security and privacy concerns, software vulnerability, integration problems, understanding of the technology, decentralized nature of Blockchain cultural acceptance and high implementation costs are some of the challenges that might hinder the implementation of Blockchain [9].

One of the biggest concerns and operational risks that Blockchain triggers is that relatively few people understand its way of working.

### 3. BLOCKCHAIN FOR SUPPLY CHAINS

A huge number of interested parties that are involved in the supply chain lead to poor transparency, non-standardized processes, and different levels of technology acceptance. As a result, it is difficult to trace the origin of goods and the status of shipments moving through the chain.

Blockchain has great application and impact on supply chains, which will be shown further in the paper. This technology secures data transparency and access to all relevant supply chain parties, creating an only source of truth. Also, trust between all parties involved in order to exchange information is enhanced, thanks to internal security mechanisms. The application of Blockchain in supply chains can lead to a drastic reduction in delays, costs, and human errors. In this way, with all exchange of products between parties, each transaction is documented, creating a permanent movement history of a product, from production to sales, i.e., from suppliers to consumers. During the product life cycle, with its downstream movement through the chain (from production to consumption), data generated at each step is documented as a transaction, creating a permanent product history. Blockchain can effectively improve the exchange of information regarding process, shipment, maintenance, and product damage in the relation between suppliers and consumers through its decentralized nature. Blockchain and supply chain integration gains are [15]: increased sustainability, reduced errors and delays, minimized transport costs, faster issue identification, enhanced trust between consumers and partners, enhanced product transport and enhanced inventory management.

Blockchain provides complete supply chain visibility. The cryptographic and immutable nature of Blockchain provides safety. Benefits of Blockchain in Supply Chain Management are [18]: increased trust, because no authority owns the provenance information, and increased efficiency that leads to reduction of time needed for identification and correction of errors.

Dabić – Miletić et al., [1] stated that in Industry 4.0 is crucial to increase transparency and traceability in the supply chain in both production and customer service. Also, the use of modern technologies is important for successful supply chain management. Blockchain is the key. For example, incorrect product labeling can poorly affect a company's reputation and cause serious economic losses in the logistic industry. This can be avoided by using Blockchain. Transaction data can not be altered or falsified, resulting in very reliable transaction information. Using smart contracts, the speed of realization of certain activities can be decimated. The dominant role of Blockchain in global trade is also mentioned in [1]. A special problem occurs with the globalization of supply chains and that is the impossible communication between users and other entities in



the supply chain due to large spatial distance. Blockchain uses electronic time stamps that supply chain members trust. As mentioned above, data is immutable, and that is how trust is established.

In China, due to its centralized logistics system, agricultural food losses are up to 30% per year. Dabić - Miletić et al., [1] proposed the use of a decentralized traceability system based on RFID and Blockchain. With Blockchain integration, the credibility of information is improved and food safety is guaranteed.

Even though Blockchain might not be as popular in other sectors, it has great possibilities. It is mentioned in [6] how Blockchain is supporting the green supply chain initiatives. Blockchain allows the monitoring and tracking of hazardous wastes and holds responsible system participants for cleanup costs.

Kshetri in [5] stated that Blockchain would make transactions in the supply chain much faster, reducing administrative and order delivery times. It was also mentioned that some of the most promising non-finance applications of Blockchain are expected in supply chains, power, and food/agriculture. Furthermore, Blockchain has great use in determining who is performing which actions, when, and where.

Cost reduction is possible due to the elimination of intermediaries [3].

Blockchain comes with challenges. One of them is its degree of adoption in supply chain processes. Blockchain requires new roles, responsibilities, and skills to support different aspects of technology adoption. Limited technical knowledge on the use of Blockchain represents a barrier to its adoption [16].

Enterprises with smaller financial power are skeptical about the high implementation costs and the potential impact of Blockchain applications, while big companies do not have that problem. In order for trust to be established, all supply chain members must share their information. Some organizations may be reluctant to do this, because other companies may exploit their information as a competitive advantage [14].

#### 4. RESEARCH

In this section, results from the research Hackius & Petersen conducted in 2017, are shown. Also, the results from self-conducted research into logistics services during the COVID-19 pandemic are presented.

##### 4.1. Companies' stances toward Blockchain

In 2017, Hackius & Petersen conducted research that included 152 participants. Most of them are employed at consulting agencies, followed by logistics companies and science. First, they questioned employees about their company's stance on Blockchain [4]. The results are shown in Figure 1.

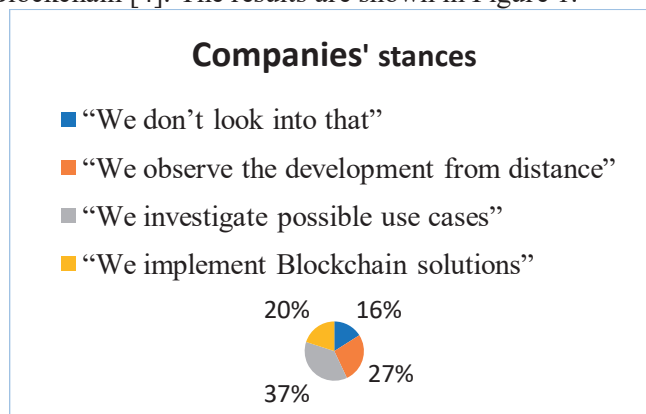


Figure 1: Companies' stances on Blockchain

In 2017, 65% of logistics companies are observing the development from a distance, only 2 of them are experimenting with Blockchain. What is the current situation? According to the website "forkast", 81 of the top 100 companies are using Blockchain technology. Blockdata research that was last updated in September 2021. shows that out of 81 companies, 65 are actively developing Blockchain solutions, while 16 remain in the research phase. Out of the mentioned 65, 14 are in the pilot phase, 24 are in development, and 27 are in production [2]. In Table 1, comparative results are presented.

Table 1: Blockchain development stages – a comparative analysis

Stances	Year	2017.	2021.
Active development of Blockchain solutions		20%	65%
Research phase		37%	16%
Zero-interest		43%	19%

Table 1 shows that in 4 years, the number of companies that are actively developing Blockchain solutions has increased by 40%.

#### 4.2. Logistics services and COVID-19 pandemic

As a contribution to this paper, research on logistics services during the COVID-19 pandemic was conducted. This research involved 100 (51 female and 49 male) participants from various backgrounds. The research took place between March and April 2022. Participants were questioned about their experience with using logistics services during the pandemic.

Out of 100 participants, 63% of them are in the age range between 19 and 25, followed by 17% in the age range between 26 and 34 and in 42 and older, and only 3% are in the age range between 35 and 41.

The majority of the participants (66%) have a full-time job, 16% have a part-time job or are an intern, while 18% are unemployed. Also, the majority of them have a college degree (42% - bachelor's, 31% - master's, and 9% - Ph.D) and 18% of them have only finished high school. Out of the employed participants, 30.5% of them work in the computer and technology sector, 18.3% in education, 9.8% in finance and economics, 9.8% in advertising and marketing, while the rest of them work in other sectors.

Out of 83% of them who claimed that they have been using logistics services during the pandemic, 73.6% of them said that the volume of logistics services they've been using has increased compared to before, while 20.7% claimed that the volume hasn't changed. Summarized answers to what extent participants have encountered certain problems are presented in table 2.

**Table 2:** Logistics services' problems during the pandemic

Problems	Frequency	1 - Never	2	3	4	5 – More than once
Order delays		28	22	32	13	5
Inability to track orders		27	36	15	13	9
One or more items are missing from the order		63	23	11	3	/
One or more items have arrived broken		65	23	8	2	2
Order stuck at the courier service warehouse		42	21	22	11	4
Difficult or impossible replacement or return of the order		47	22	19	8	4
Unreachable customer service		43	18	23	13	3
Major increase in shipment cost		42	19	20	14	5
With food ordering, spoiled food arrived		61	20	13	3	3
Spoiled food contagion		68	18	9	4	1

Some of the companies have gone above and beyond and have included new ways of doing their business to satisfy customers' needs and help them in the time of the pandemic. Percentages of participants that have recognized the innovations that companies introduced are presented in table 3.

**Table 3:** Companies' innovations during the pandemic

Innovation	Percentage
Home delivery	69.7
Contactless delivery	28.1
Payment via mobile app	51.7
Pickup delivery	33.7
Real - time customer support and problem – solving	21.3
The company has recognized its customers' problems and has offered genuine support	12.4

According to the results shown in table 2, problems like: 1) the inability to track orders, 2) order delays and, a 3) major increase in shipment costs are more frequent than others. Problems like spoiled food or spoiled food contagion are not that frequent, but should not be happening at all. The transparent nature of Blockchain and its key characteristic – provenance, can help in overcoming the first problem. Provenance provides the information of asset origin, and transparency enables all the Blockchain members to see the movement of an asset through the chain. With Blockchain, all members of the supply chain can access data at any time and from anywhere. Shipment management is easy and in real-time.

How can Blockchain prevent order delays and minimize shipment costs? A receipt of services, that includes information such as the type, quantity, and destination of goods in shipment is called the Bill of lading. Bill of lading documents are paper-based and get couriered to the shipper. This often results in delays and it's expensive. There are also detention fees the shipper has to pay when a delay occurs. Data encryption that Blockchain uses overcomes the fraud problems with paper documents. Delays are prevented because data is available in real-time and due to overall digitization. With Blockchain, manual labor, courier fees, and bank charges are reduced, resulting in a cut-down in shipment costs [19].

Problems regarding spoiled food or spoiled food contagion are a major issue and the likelihood of their occurring should be minimized. Since all product movements are invariably recorded in the distributed ledger, determining the origin of food of questionable quality can be done in a few hours, instead of a few days or even weeks. In this manner, the spread of contagion can be minimized.

**5. CASES**

Summarized applications of Blockchain in practice and the problems it solves are presented in table 4.

**Table 4:** Summarized Blockchain's applications

Reference	Blockchain solution	Industry	Problem that it solves	Blockchain key characteristic
[5]	Damco (Maersk & IBM)	Logistics	Tracking shipping containers and mountains of paperwork	Provenance and digitality
[5]	From shore to plate (Provenance)	Fishing	Traceability, lack of supervision, and corruption	Provenance and immutability
[5]	Track & trace (Chronicle, LinkLab)	Pharmaceutical	Tracking and falsification	Provenance and immutability
[5]	Stellar Blockchain (Bext360)	Coffee	Real – time payment	Immutability and provenance
[7]	Tracr	Jewelry	Traceability, trust and authenticity	Provenance, immutability, and finality
[7]	Trado	Beverage (tea)	Tracking, falsification, and transparency	Provenance, immutability, and consensus
[7]	Food Trust Blockchain	Food	Food poisoning	Provenance
[10]	Morpheus.Network	Multiple industries	Transparency and global trade inefficiencies	Smart contracts
[11]	Blockchain based oil solution (Adnoc & IBM)	Oil	Tracking, transactions automation, and transparency	Provenance, immutability, and digitality
[13]	aXedras	Precious metals	Transparency, regulatory compliance, and automation	Provenance, immutability, and digitality
[17]	VeChain – Agriculture solution	Agriculture	Tracking from planting to sales, digital data storage, and allowed data access to all participants	Provenance and digitality

**6. CONCLUSION**

Blockchain started as a technology for the finance sector, but it is continuously showing great potential for other sectors. In this paper, the advantages of using Blockchain in supply chains are highlighted. Based on various literature sources, the benefits of Blockchain and supply chain integration are featured in the paper. Blockchain secures transparency, which is the major supply chain problem. Due to the use of cryptographic methods, trust among supply chain members is enabled. Trust is crucial for all supply chain operations to run smoothly. Once a transaction has been recorded on to the ledger, it can not be changed. Blockchain represents the only source of truth, to which all supply chain members have access. Blockchain improves traceability and supports green supply chain initiatives.

The number of companies that are actively developing Blockchain solutions in 2017. was 20% of companies that were included in the research, while in 2021. that number increased to 65%.

As a contribution to this paper, research on logistics services during the COVID-19 pandemic was conducted. Major issues that individuals were dealing with during the pandemic were: the inability to track orders, order delays, and a major increase in shipment costs. The overall digital nature of Blockchain and its key characteristics can help overcome these problems. Other major issues that should be addressed are problems regarding spoiled food or spoiled food contagion. Many Blockchain solutions were developed to prevent these issues. Some of them are shown in Section 5 of this paper among a variety of use cases.

Major obstacles in adopting Blockchain are, not enough knowledge about this new technology and great implementation costs. Before adopting Blockchain technology, a cost-benefit analysis should be performed. If the benefits that Blockchain would potentially bring to the company are greater, or more important than implementation costs, then it is cost-effective to start developing a suitable Blockchain solution for the company. Further research could go in the direction of exploring possible Blockchain application advantages for smaller enterprises, considering major Blockchain implementation costs.

## REFERENCES

- [1] Dabić – Miletić, S., Božić, M., Pavlov, N. (2021). Importance of Blockchain technology for supply chain management. *E-business technologies conference proceedings*. 1, 1 (Sep. 2021), 151–156.
- [2] Forkast, <https://forkast.news/81-of-top-100-companies-use-blockchain-technology-blockdata/>, last access: 17.03.2022.
- [3] Hasan, H., AlHadhrani, E., AlDhaheri, A., Salah, K. and Jayaraman, R. (2019), Smart contract-based approach for efficient shipment management. *Computers and Industrial Engineering*, Vol. 136, pp. 149-159.
- [4] Hackius, N., Petersen, M. (2017). Blockchain in Logistics and Supply Chain: Trick or Treat? *Proceedings of the Hamburg International Conference of Logistics (HICL) – 23*.
- [5] Kshetri, N. (2018). 1 Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management*, Vol. 39, pp. 80-89.
- [6] Kim, J-S., Shin, N. (2019). The Impact of Blockchain Technology Application on Supply Chain Partnership and Performance. *Sustainability 2019*, 11, 6181; doi:10.3390/su11216181.
- [7] Logistics Bureau, <https://www.logisticsbureau.com/>, last access: 17.03.2022.
- [8] Lasi, H., Fettke, P., Feld, T., Hoffmann, M. (2014). Industry 4.0. *Business and Information Systems Engineering*, 6(4), 239–242, <https://doi.org/10.1007/s12599-014-0334-4/>.
- [9] Morabito, V. (2017). Business Innovation Through Blockchain: The B3 Perspective, Department of Management and Technology, Bocconi University, Milan, Italy.
- [10] Morpheus.Network, <https://morpheus.network/>, last access: 07.03.2022.
- [11] Openledger, <https://openledger.info/insights/blockchain-in-the-supply-chain-use-cases-examples/>, last access: 28.04.2022.
- [12] Pilkington, M. (2016). Blockchain Technology: Principles and Applications, Research Handbook on Digital Transformations. Ed. by F. X. Olleros and M. Zhegu. Edward Elgar Publishing, pp. 1–39.
- [13] R3 – aXedras, <https://www.r3.com/case-studies/axedras/>, last access: 28.04.2022.
- [14] Saberi, S., Kouhizadeh, M., Sarkis, J. and Shen, L. (2019), Blockchain technology and its relationships to sustainable supply chain management. *International Journal of Production Research*, Vol. 57 No. 7, pp. 2117-2135.
- [15] Tijan, E., Aksentijević, S., Ivanić, K., Jardas, M. (2019). Blockchain Technology Implementation in Logistics. *Sustainability*, 11(4), 1185; <https://doi.org/10.3390/su11041185>.
- [16] Varriale, V., Cammarano, A., Michelino, F., Caputo, M. (2020). New organizational changes with blockchain: a focus on the supply chain. *Journal of Organizational Change Management*, Vol. 34, No. 2, pp. 420-438, DOI 10.1108/JOCM-08-2020-0249.
- [17] VeChain, <https://www.vechain.com/solution/agriculture>, last access: 28.04.2022.
- [18] Wiley, J. (2017). Blockchain For Dummies, IBM Limited Edition, Hoboken.
- [19] XChange, <https://www.container-xchange.com/blog/blockchain-shipping/>, last access: 17.03.2022.

## IDENTIFICATION OF RISK IN SUPPLY CHAINS

TAMARA BOJANIĆ<sup>1</sup>, BRANISLAV STEVANOV<sup>2</sup>, NENAD NOVAKOVIĆ<sup>3</sup>

<sup>1</sup> University of Novi Sad – Faculty of Technical Sciences tamarabojanic@uns.ac.rs

<sup>2</sup> University of Novi Sad – Faculty of Technical Sciences branisha@uns.ac.rs

<sup>3</sup> University of Novi Sad – Faculty of Technical Sciences - doctoral academic studies novakovic.ftn@gmail.com

---

**Abstract:** *Supply chains in today's environment, which is rapidly changing, are more exposed to changes due to increased globalization, environmental instability, higher customer expectations in terms of sustainability and the occurrence of various internal and external risks. Due to the impact of the pandemic, there have been major changes in the business of supply chains. This work presents the exposure of the supply chain to risks due to the pandemic and changes in regulations. Companies are putting greater importance on risk identification and turn to better understanding and managing the risks in the supply chain. Companies have developed supply chain risk management in order to maintain business continuity and reduce the consequences of the risk.*

**Keywords:** *supply chain, risk, business process*



## IMPACT OF LOCATION ON DELIVERY DURATION IN URBAN AREAS

MLADEN KRSTIĆ<sup>1</sup>, MILOŠ VELJOVIĆ<sup>2</sup>, SNEŽANA TADIĆ<sup>3</sup>, SLOBODAN ZEČEVIĆ<sup>4</sup>

University of Belgrade - Faculty of Transport and Traffic Engineering, Belgrade,

<sup>1</sup> m.krstic@sf.bg.ac.rs <sup>2</sup> m.veljovic@sf.bg.ac.rs <sup>3</sup> s.tadic@sf.bg.ac.rs <sup>4</sup> s.zecevic@sf.bg.ac.rs

**Abstract:** *The duration of the delivery of goods to trade, catering, service and other commercial facilities, i.e. generators of logistic flows in urban areas, depends on numerous factors and attributes such as size of delivery, unloading technology, number of workers, place of stopping of the delivery vehicle, etc. Different architectural and civil engineering characteristics and conditions at different locations in the city area (infrastructure characteristics, intensity of pedestrian flows, etc.) can have a significant direct impact on the delivery duration. On the other hand, real estate prices, access restrictions to certain zones, traffic organization and street networks in different parts of the urban environment can indirectly affect the delivery duration. The aim of this paper was to determine whether and in what way the location of the flow generator affects the delivery duration. Fuzzy systems and data collected by field research in the area of the city of Belgrade was used to determine the duration of delivery to the flow generators in certain city zones.*

**Keywords:** *delivery duration, location, urban areas, city logistics, fuzzy systems*

### 1. INTRODUCTION

City is the place of the largest concentration of economic and social activities, and the delivery of goods is a prerequisite for the maintenance of urban life and business activities for achieving wealth and development of the city [13]. Logistics systems and processes that enable the realization of commodity flows support employment and generate income, therefore, logistics plays an important role in the competitiveness of urban areas and should be an integral part of the city's economy. City logistics is a field that attracts increasing attention of professionals and scientific community and international organizations. City logistics flows are characterized by partiality, spatial dispersion of generators, diversity in terms of the logistics chains structure, frequency of a large number of smaller shipments, dynamics, stochasticity etc. [14]. They are all influenced by various characteristics of the city.

The development of cities depends on numerous natural, historical, cultural, political, economic and other factors. In many cities, the impact of these factors led to spatial heterogeneity in terms of density of population, construction and street network, concentration of economic and other contents, intensity of flows of people, means of transport, goods, materials and freights. Therefore, in different zones of most cities there are also differences from the aspect of the attributes of the delivery of goods, materials and freights to facilities, i.e. the generators of logistics flows.

One of the most important attributes of the delivery is the duration and it implies the time from the arrival of the delivery vehicle to the place of delivery until its departure. It has tactical and operational importance for both delivery operators and receivers, i.e. the generators of logistics flows in the urban areas [11]. Based on the delivery duration, operators carry out operational distribution planning and vehicle routing, as well as dimensioning of the transport system (vehicle fleet, workforce, etc.). Delivery duration depends on a number of factors and attributes that describe logistics flow generators: delivery time (term), core business and size of the generator, frequency and size of deliveries, structure of goods, supply system, etc. [18].

According to the previous, the subject of this paper is the analysis of the delivery duration, and for solving this problem the fuzzy systems are developed in the paper. There are numerous applications of the fuzzy set theory in the field of city logistics. Most often it is used for developing different methods of multi-criteria decision-making in a fuzzy environment, e.g. for the assessment of city logistics initiatives [1], selection of a city logistics concept [8], selection of the city logistics terminal location [7,15,17], prioritizing the strategies to enhance smart city logistics [2], analyzing key influence factors of city logistics development [4], etc. There are also examples of fuzzy clustering, e.g. for modeling the demand-responsive city logistics distribution operations [8], analysis of regional city multi-level logistics distribution center [6], city logistics network optimization [16], determining city logistics center scale [5], etc. However, application of fuzzy systems in the field of city logistics is very limited. One such system is developed for determining delivery duration in urban areas [11]. Fuzzy systems have also been applied in combination with neural networks for vehicle routing in

urban areas [3]. By reviewing the existing literature, no research was found in which the impact of the location on the delivery duration was assessed using the fuzzy system.

The goal and contribution of this paper is reflected in determining the impact of the location of the flow generators on the delivery duration, based on the field research, interviews with the employees in the facilities and the results obtained by the established fuzzy system for determining the delivery duration. The applicability of the defined methodology was demonstrated on a real-life example for three different zones of the city of Belgrade.

The paper is organized as follows. After the introduction, the problem and the model based on the fuzzy systems, developed for solving it, are described in the second section. Values of the input variables for the model, established through the field research, are also presented within this section. In the third section, the results of the application of the model were presented and a comparative analysis for three zones was performed. At the end, concluding remarks and directions for future research are given.

## 2. PROBLEM STATEMENT, MODEL DEVELOPMENT AND INPUT VALUES

The problem discussed in this paper involves a comparative analysis of the delivery duration in three different zones of the city of Belgrade, in order to determine the dependence of the delivery duration on the location of the generator of logistics flows. At the same time, in order to obtain data on the delivery time in different zones, a comparative analysis of other delivery attributes and generators (delivery size, number of items, distance of the unloading place, etc.) was performed. The observed zones differ in terms of urban characteristics, density of population, construction and street network, concentration of economic contents, intensity of flows, etc. Zone 1 (Knez Mihajlova and the surrounding streets) represents the immediate center of the city, where there is the greatest concentration of cultural, historical, touristic, commercial, catering and other contents, and thus the flow of people, goods, materials and freights. The observed area is mostly pedestrian, and it is planned to be entirely dedicated to pedestrian traffic in the future. The zone is characterized by a high density of buildings. Zone 2 (Southern part of the municipality of Vračar) belongs to the wider city center. It is also characterized by a high density of buildings and a large concentration of various cultural and economic contents, but with a higher percentage of the area dedicated for residential purposes than in the case of Zone 1. Zone 3 (Vojvode Stepe Street with the surrounding, cross streets) connects the city center with the peripheral parts of the city with a lower degree of urbanity. It is characterized by a large concentration of various economic contents, but the share of residential space is higher than in the case of Zone 2.

A model based on fuzzy systems, defined in the paper [11], will be used to estimate the delivery duration in these three zones. The model is defined as follows. The delivery duration is obtained by summing up the time before unloading ( $T_1$ ), time of unloading ( $T_2$ ), time of goods control and delivery confirmation ( $T_3$ ) and other time ( $T_4$ ).  $T_2$  and  $T_3$  are obtained by applying the fuzzy system, while  $T_1$  and  $T_4$  are simulated based on the data obtained in the field survey. The input variables of the fuzzy system for determining  $T_2$  are the size of the delivery ( $I_1$ ), the distance of the unloading place ( $I_2$ ), the complexity of the unloading ( $I_3$ ) and the surrounding conditions ( $I_4$ ). Surrounding conditions are assessed using the second fuzzy system, whose input variables are the infrastructure characteristics ( $I_{4.1}$ ), the intensity of pedestrian/user flows ( $I_{4.2}$ ) and weather conditions ( $I_{4.3}$ ). The third fuzzy system estimates the time  $T_3$ , based on the size of the delivery ( $I_1$ ) and the number of items ( $I_5$ ).

The location of the flow generator directly or indirectly affects most of the listed factors/parameters. Considering the prices of space and the large concentration of various contents (cultural, historical, touristic, catering, trading, administrative, etc.), smaller facilities with smaller storage space are most often located in the inner city core. For this reason, deliveries to such facilities are usually of a smaller volume and take less time than in typical logistics systems (logistics, distribution centers, freight terminals etc.).

Through field research, monitoring deliveries and interviewing employees in facilities, the values of the specified input variables for 100 deliveries in all three zones were determined. The following will show the percentage share of certain ranges of values for all input variables, as well as the simulated values of the time before unloading and the other time.

Table 1 shows data on the *size of deliveries* ( $I_1$ ) in Zones 1, 2 and 3 (in  $m^3$ ). Through a comparative analysis, it can be concluded that there is a significant correlation in the structure of data on the size of deliveries in Zones 1 and 2. A small difference is reflected in a slightly higher percentage of small (up to  $0.066 m^3$ ) and large deliveries (over  $1 m^3$ ) in the case of Zone 2. Research has established that this difference is primarily influenced by the core business and size of the generators. Namely, compared to Zone 1, Zone 2 has a larger number of smaller service facilities (cosmetics, hairdressing, beauty salons, etc.) that generate a large number of smaller-scale deliveries. On the other hand, in Zone 2 there are a slightly larger number of large-area facilities, mainly grocery stores, which affects the increase of number of larger deliveries. From the aspect of



delivery size, the differences of Zone 3 compared to Zones 1 and 2 are somewhat larger. Namely, a significantly higher number of deliveries of a larger volume (over  $0.4 m^3$ ) were noticeable. This is primarily a consequence of the larger business and storage space of facilities in these zones, the price of which is lower than in the city center.

The *distance of the unloading place* ( $I_2$ ), i.e. the distance covered when unloading the goods, also has a significant impact on the delivery duration. The place where the delivery vehicle stops can be a roadway, a public parking lot, a parking lot owned by the facility, a sidewalk, a pedestrian zone, a public parking garage, etc. Given that a significant part of Zone 1 is a pedestrian zone, there are access restrictions in it, so delivery vehicles often stop in other streets or public parking garages, which increases the distance covered during delivery. In addition, the street profile, traffic organization, number of traffic lanes, positioning of the parking lot, tree line, dimensions of the sidewalk, etc., influence the distance of the delivery vehicle's stopping place from the facility. These factors also affect the differences between the observed zones in terms of the distance covered during delivery. Thus, in Zone 2, a large number of small deliveries are made by stopping the vehicle for a very short time on the road, often right next to the recipient's facility, so the distance covered is very small. On the other hand, in Zone 3, the number of such deliveries is lower, primarily because many facilities in Vojvode Stepe Street are not positioned directly next to the road or sidewalk.

Table 1 shows data on the *complexity of unloading* ( $I_3$ ) in the observed zones. This attribute includes specific (dangerous, bulky, easy breakable goods, etc.) and inhomogeneous logistic demands of goods (different shapes, dimensions, weight, unit of consolidation, etc.), compactness of consolidation, application of handling equipment (e.g. handcarts), vehicle characteristics (height from which the goods are picked up, vehicle equipment for unloading, complexity of vehicle preparation for unloading, etc.). This and all the input variables that are shown below, except for the number of items, can take values from the interval  $[0,10]$ , where the value 0 indicates the most favorable (e.g. very low complexity of unloading), and the value 10 the most unfavorable influence of the input variable on surrounding conditions (e.g. very high complexity of unloading).

**Table 1:** Values of input variables

$I_1$ ( $m^3$ )	Zone 1					Zone 2					Zone 3							
	<0.03	0.03-0.06	0.06-0.1	0.1-1	>1	<0.03	0.03-0.06	0.06-0.1	0.1-1	>1	<0.03	0.03-0.06	0.06-0.1	0.1-1	>1			
	25%	19%	17%	27%	12%	28%	20%	15%	22%	15%	12%	16%	15%	37%	20%			
$I_2$ (m)	<5	5-10	10-15	15-20	20-25	>25	<5	5-10	10-15	15-20	20-25	>25	<5	5-10	10-15	15-20	20-25	>25
	19%	17%	13%	12%	19%	20%	22%	16%	16%	14%	16%	16%	12%	17%	18%	16%	18%	19%
$I_3$	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10		
	24%	17%	31%	28%	21%	22%	30%	27%	22%	22%	25%	31%	22%	22%	25%	31%		
$I_{4.1}$	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10		
	19%	15%	21%	45%	22%	24%	28%	26%	22%	27%	28%	23%	22%	27%	28%	23%		
$I_{4.2}$	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10		
	4%	8%	19%	69%	15%	19%	22%	44%	21%	24%	23%	32%	21%	24%	23%	32%		
$I_{4.3}$	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10		
	21%	70%	9%	0%	24%	69%	7%	0%	30%	65%	5%	0%	30%	65%	5%	0%		
$I_4$	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10	<2.5	2.5-5	5-7.5	7.5-10		
	8%	18%	52%	22%	13%	23%	44%	20%	20%	42%	20%	18%	20%	42%	20%	18%		
$I_5$	<10	10-20	20-30	>30	<10	10-20	20-30	>30	<10	10-20	20-30	>30	<10	10-20	20-30	>30		
	27%	36%	22%	15%	29%	19%	19%	33%	23%	25%	27%	25%	23%	25%	27%	25%		

*Surrounding conditions* ( $I_4$ ) represent the fourth input variable in the fuzzy system for determining the time of unloading. Urban areas are characterized by the variability of surrounding conditions in space and time [7]. Logistics processes and activities are mostly carried out in open space. They depend on the characteristics, quality and adaptability of the infrastructure to logistics activities, traffic, pedestrian and user flows in the facilities, weather conditions, etc.

The first input variable in the fuzzy system for the assessing of surrounding conditions is *infrastructure characteristics* ( $I_{4.1}$ ). In general, the access and infrastructure of facilities in the city area is mostly insufficiently adapted to the performing logistics activities [9,10]. That is why the infrastructure used to perform the majority of deliveries in all three zones (66%, 57% and 51%) was evaluated with values between 5 and 10. The lack of adaptation of the infrastructure is particularly pronounced in Zone 1, which consists of narrow, mostly pedestrian and paved streets, with numerous cultural and historical contents, architectural and construction peculiarities, and which, from the aspect of functionality and attractiveness, is arranged above all in relation to the end users. The situation is somewhat more favorable in Zone 2. Finally, Zone 3 has a slightly higher degree of adaptation of the infrastructure to logistics activities compared to the other zones. Namely, a larger number of objects have a separate and adequate entrance to the object's storage space, a larger space for manipulation during unloading, etc. However, in this zone there is also a significant number of buildings whose

infrastructure is not at a sufficient level of quality due to obsolescence and the need for reconstruction, therefore a significant percentage of unfavorable evaluations are present in this zone as well.

The *intensity of pedestrian flows* ( $I_{4.2}$ ) is particularly high in Zone 1, which is mostly pedestrian and where the highest concentration of various economic, cultural, tourist and other contents is present. The intensity of pedestrian flows generally decreases with distance from the center to the outskirts of the city, so it is also lower in Zones 2 and 3.

The last input variable in the fuzzy system for the assessment of surrounding conditions is the *weather conditions* ( $I_{4.3}$ ). The research was conducted in relatively stable weather conditions, so mostly daily weather fluctuations affected the structure of the data, but there were no significant differences between the observed zones.

The *number of items* in the delivery ( $I_5$ ), i.e. their variety depends on numerous other attributes (core business and size of the facility, supply system, etc.). Increasing number of items in the delivery increases the time of goods control and delivery confirmation. Data on the number of items in deliveries in the observed zones are given in table 1.

### 3. RESULTS OF MODEL APPLICATION AND COMPARATIVE ANALYSIS

Table 2 shows the values of individual parts of the delivery duration, as well as the total delivery duration. Time  $T_1$  includes the time of searching for a place to stop, parking and possibly waiting for unloading (e.g. due to the receiver's unwillingness to receive the goods, delivery by other suppliers, etc.). Although the problems of finding a parking are usually most pronounced in the inner city core, the research shows that in Zone 2 there is a slightly higher percentage of deliveries in which the time before unloading is greater than 2.5 minutes than in Zone 1. Namely, in Zone 2 many delivery operators use public parking, so they depend more on other users, primarily residents. On the other hand, although there is often a problem of finding a parking space in Zone 1, access permits to the pedestrian zone for a certain period of time, strict compliance with these restrictions by operators and more serious planning of deliveries, contribute to greater efficiency in providing a place to stop. In Zone 3, on the other hand, the shorter average time of searching for parking is caused by a higher percentage of parking spaces in the possession of facilities to which goods are delivered, as well as lower traffic intensity. The data shown in Table 2, related to  $T_1$ , were simulated based on the data obtained in the field survey.

**Table 2:** Values of parts of delivery duration and total delivery duration (in minutes)

	Zone 1					Zone 2					Zone 3				
	< 1	1.5 - 2.5	2.5 - 5	> 5		< 1	1.5 - 2.5	2.5 - 5	> 5		< 1	1.5 - 2.5	2.5 - 5	> 5	
$T_1$	33%	43%	12%	12%		29%	21%	18%	32%		43%	25%	12%	20%	
$T_2$	< 5	5 - 10	10 - 15	> 15		< 5	5 - 10	10 - 15	> 15		< 5	5 - 10	10 - 15	> 15	
	11%	31%	29%	29%		14%	31%	26%	29%		8%	28%	33%	31%	
$T_3$	< 1	1 - 2	2 - 3	3 - 4	> 4	< 1	1 - 2	2 - 3	3 - 4	> 4	< 1	1 - 2	2 - 3	3 - 4	> 4
	18%	37%	17%	15%	13%	20%	34%	19%	13%	14%	19%	28%	22%	14%	17%
$T_4$	< 0.25	0.25 - 0.5	0.5 - 1.5	> 1.5		< 0.25	0.25 - 0.5	0.5 - 1.5	> 1.5		< 0.25	0.25 - 0.5	0.5 - 1.5	> 1.5	
	41%	31%	15%	13%		35%	29%	19%	17%		36%	27%	18%	19%	
$T$	< 5	5 - 10	10 - 15	15 - 20	> 20	< 5	5 - 10	10 - 15	15 - 20	> 20	< 5	5 - 10	10 - 15	15 - 20	> 20
	10%	31%	29%	13%	17%	12%	29%	26%	21%	12%	15%	28%	28%	13%	16%

Based on the values of the input variables, mentioned in the previous section, by applying the fuzzy systems, data on  $T_2$  and  $T_3$ , presented in table 2, were obtained. The example of applying the fuzzy system for deliveries in Zone 1 is as follows. For the input values of 9.2 (bad or normal) for the infrastructure characteristics, 9.1 (high or medium) for the intensity of pedestrian flows, and 3 (good or normal) for the weather conditions, the fuzzy system for assessing surrounding conditions resulted in the output value of 6.. Based on the input values for the second fuzzy system, namely the size of the delivery of  $0.015 m^3$ , the distance of the unloading place of  $7m$ , and the complexity of the unloading of 2 (low or medium), it was determined that the time of unloading is equal to 5.4 min. Based on the third fuzzy system, for the same size of delivery ( $0.015 m^3$ ) and the number of items equal to 5, the time of goods control and delivery confirmation was determined to be 0.35 min. Table 3 presents the results of the obtained values of  $T_2$  and  $T_3$  for this example case.

The other time in delivery ( $T_4$ ) refers to other activities, such as taking over the logistics units from the previous or current delivery, workers' rest etc. Data on this part of the delivery duration are given in Table 2.

By summing up the values of the four mentioned variables, the data on the delivery duration in all three zones ( $T$ ) were obtained.

Table 3 gives examples of the application of the model for deliveries in each of the observed zones. It can be seen that although they have different values of the input variables, the deliveries have similar durations.

For example, if deliveries in Zones 1 and 3 are observed, it is evident that there is an opposite effect of input variables. Namely, while the small delivery size ( $0.015 m^3$ ) contributes to the short delivery duration, unfavorable surrounding conditions (6) have the opposite effect. On the other hand, the larger delivery volume in Zone 3 ( $0.656 m^3$ ) makes the delivery time long, but the better surrounding conditions (2.36) contribute to making it shorter. Based on the analysis of field research data shown in table 1, it can be concluded that the examples in table 3 are representative, i.e. they reflect well the phenomenon of "neutralization" of the influence of the input variables in the entire sample. However, this does not mean that they reflect the absence of influence of the location of the logistics flow generator on the delivery duration in the city of Belgrade, nor in urban areas in general. On the contrary, the indirect influence of the location, i.e. its influence on the variables on which the delivery duration depends certainly exists and it is recorded in the description of the input variables, but based on what is shown it is possible to give only the following conclusion: the structure of the duration of deliveries in an area dominated by small-scale deliveries and unfavorable surrounding conditions has a high degree of correlation with the structure of the duration of deliveries in an area dominated by large-scale deliveries and more favorable surrounding conditions. The same conclusion can be drawn for all other combinations of input variables whose effects on the delivery duration "neutralize" each other.

**Table 3:** Examples of model application

	$T_1(\min)$	$I_1(m^3)$	$I_2(m)$	$I_3$	$I_{4,1}$	$I_{4,2}$	$I_{4,3}$	$I_4$	$T_2(\min)$	$I_5$	$T_3(\min)$	$T_4(\min)$	$T(\min)$
$Z_1$	2.43	0.015	7	2.0	9.2	9.1	3	6	5.4	5	0.35	0.45	8.63
$Z_2$	2.25	0.245	14	6.5	9.4	8.2	2.1	4.43	6.06	6	0.39	0.25	8.95
$Z_3$	2.33	0.656	7	2.0	0.4	7.4	3.3	2.36	5.57	10	0.40	0	8.30

#### 4. CONCLUSION

The duration of the delivery of goods to the generators of logistics flows in urban areas depends on numerous factors and other attributes of the delivery and generator. On the other side, it also affects numerous decisions concerning the delivery operator and receiver of goods at the tactical and operational level. Therefore, considering the structure of delivery duration in cities is a very important task, as well as analyzing the factors that influence it.

In this paper, the direct and indirect influence of the location of the logistics flow generator on the delivery duration is analyzed. The goal and contribution of the paper is reflected in determining the impact of the location of the generator on the delivery duration, based on field research, interviewing the employees in the facilities and applying the fuzzy system for determining the delivery duration. In addition, some of the most important city logistics parameters of parts of the city of Belgrade are presented.

The paper did not prove the direct influence of the location of the facility on the delivery duration, but rather on the variables upon which it depends. It has been established that the structure of the duration of deliveries in an area dominated by small-scale deliveries and unfavorable surrounding conditions has a high degree of correlation with the structure of the duration of deliveries in an area dominated by large-scale deliveries and more favorable surrounding conditions.

The paper represents the basis for numerous future researches. In one of the following, field research and surveys should be used to determine real data on the duration of delivery in the observed zones, and then compared to the data obtained by applying a model based on fuzzy systems, in order to identify their mutual correlation. Comparing the results of these two methods of determining the delivery duration can serve to verify the model based on fuzzy systems, that is, to observe its possible shortcomings and the possibility of modification or expansion. Defining a fuzzy system for determining delivery duration with different weights of input variables also represents one of the significant directions of future research.

#### Acknowledgement

This paper was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (projects TR36006 and TR36027).

#### REFERENCES

- [1] Awasthi, A., & Chauhan, S. S. (2012). A hybrid approach integrating Affinity Diagram, AHP and fuzzy TOPSIS for sustainable city logistics planning. *Applied Mathematical Modelling*, 36(2), 573–584.
- [2] Buyukozkan, G., & Göçer, F. (2019). Prioritizing the strategies to enhance smart city logistics by intuitionistic fuzzy CODAS. In *11th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2019)*, 805-811.
- [3] Ćirović, G., Pamučar, D., & Božanić, D. (2014). Green logistic vehicle routing problem: Routing light delivery vehicles in urban areas using a neuro-fuzzy model. *Expert Systems with Applications*, 41(9), 4245-4258.
- [4] Hongmei, H., & Haifang, C. (2012). Analyzing key influence factors of city logistics development using the fuzzy decision making trial and evaluation laboratory (DEMATEL) method. *African Journal of Business Management*, 6(45), 11281-11293.
- [5] Ming-bao, P. A. N. G., & Ling, X. I. E. (2006). Research into Merchant Logistics Center Scale Determining By Fuzzy Clustering. In *2006 IEEE International Conference on Management of Innovation and Technology*, 881-885.
- [6] Rao, C., Goh, M., Zhao, Y., & Zheng, J. (2015). Location selection of city logistics centers under sustainability. *Transportation Research Part D: Transport and Environment*, 36, 29–44.
- [7] Ren, Y. C., Xing, T., Quan, Q., & Zhao, G. Q. (2010). Fuzzy cluster analysis of regional city multi-level logistics distribution center location plan. In *Quantitative Logic and Soft Computing 2010*, 499-508.
- [8] Sheu, J. B. (2006). A novel dynamic resource allocation model for demand-responsive city logistics distribution operations. *Transportation Research Part E: Logistics and Transportation Review*, 42(6), 445–472.
- [9] Tadić, S., & Veljović, M. (2021). Differences between typical logistics systems and other city logistics generators, *Tehnika*, 76(4): 485-489.
- [10] Tadić, S., & Zečević, S. (2016). *Modelling city logistics concepts* (in Serbian). Belgrade: Faculty of transport and Traffic Engineering, University of Belgrade.
- [11] Tadić, S., Krstić, M., & Veljović, M. (2021). Determining delivery duration in urban areas using fuzzy system. In *Proceedings of XLVIII International Symposium on Operational Research, SYM-OP-IS 2021*, Banja Koviljača, Serbia, 20-23. september 2021, 398-403.
- [12] Tadić, S., Zečević, S., & Krstić, M. (2014). A novel hybrid MCDM model based on fuzzy DEMATEL, fuzzy ANP and fuzzy VIKOR for city logistics concept selection. *Expert Systems with Applications*, 41(18), 8112–8128.
- [13] Tadić, S. R., Zečević, S. M., & Krstić, M. D. (2014). City logistics initiatives aimed at improving sustainability within existing context of urban area. *Tehnika*, 69(3), 487-495.
- [14] Tadić, S. R., Zečević, S. M., & Krstić, M. D. (2014). City logistics initiatives aimed at improving sustainability by changing the context of urban area. *Tehnika*, 69(5), 834-843.
- [15] Tadić, S. R., Zečević, S. M., & Krstić, M. D. (2013). Locating city logistics terminal using fuzzy AHP analysis: Case of Belgrade. *Tehnika*, 68(4), 707-716.
- [16] Tadić, S., Zečević, S., & Krstić, M. (2012). City logistics terminal location selection using combined fuzzy AHP and fuzzy TOPSIS analysis. In *Proceedings of the 1st International conference on traffic and transport engineering, ICTTE*, 345-358.
- [17] Wang, Y., Ma, X., Lao, Y., & Wang, Y. (2014). A fuzzy-based customer clustering approach with hierarchical structure for logistics network optimization. *Expert systems with applications*, 41(2), 521-534.
- [18] Zečević, S., & Tadić, S. (2006). *City logistics* (in Serbian). Belgrade: Faculty of transport and Traffic Engineering, University of Belgrade.

## INTEGRATING BLOCKCHAIN INTO SUPPLY CHAIN MANAGEMENT

MILAN TODOROVIĆ<sup>1</sup>, LUKA MATIJEVIĆ<sup>2</sup>

<sup>1</sup> Mathematical Institute of the Serbian Academy of Sciences and Arts, mtodorovic@mi.sanu.ac.rs

<sup>2</sup> Mathematical Institute of the Serbian Academy of Sciences and Arts, luka@mi.sanu.ac.rs

**Abstract:** *Since its inception in 2008, blockchain technology has found its place in many different areas. One of such areas is supply chain management, as blockchain technology and smart contracts can be used to verify the origin of the goods and materials, share information in the supply chain and automatize some of the transactions. In this paper, we will present some of the applications and potential issues with incorporating blockchain into supply chain management.*

**Keywords:** *Smart contract, Distributed ledger, Traceability, Information sharing, Applications*

### 1. INTRODUCTION

In the era of globalization, almost every item that we can find in our homes had to be transported through a complex series of organizations before it reached its final destination. From the sourcing of raw materials, through the manufacturing of the parts and assembling the final product, all the way to the transportation of the product to the end-user, many organizations from all across the world have to be involved, each adding a value to the final product. Because of this, it is often impossible for the end-user to trace the origin of the product and to find reliable information on where the product was manufactured, where the materials came from, whether it was produced in an environmentally sustainable way or whether the companies involved in the process of manufacturing used unethical labor practices. Moreover, even the organizations involved in the supply chain often can not find this information about their high-tier suppliers. On the other hand, it is evident that the end-users and companies would benefit from the increased transparency in the supply chain, as it would allow them to make an informed decision while purchasing the products or materials. Furthermore, the amount of direct communication and paperwork necessary to transport the goods between participants in the supply chain can negatively affect the productivity of the participants and slow down the entire process of delivering goods, thus lowering the overall profit.

For these reasons, companies and researchers around the world have put a lot of effort into optimizing the supply chain in such a way that it minimizes the negative effects of the aforementioned issues. During the last several years, blockchain technology has been proposed as a potential solution to these problems, as it would provide a higher level of security and fraud prevention. In this paper, we explored such applications of blockchain to supply chain management, focusing mostly on tracing the goods, providing relevant information about them, sharing information between members of the supply chain, and automatizing transactions.

The paper is organized as follows. In Section 2 we present the overall overview of blockchain technology and smart contracts. Section 3 contains some of the applications of the blockchain to supply chain management, and in Section 4 we present some of the issues with adopting this technology, together with some possible solutions.

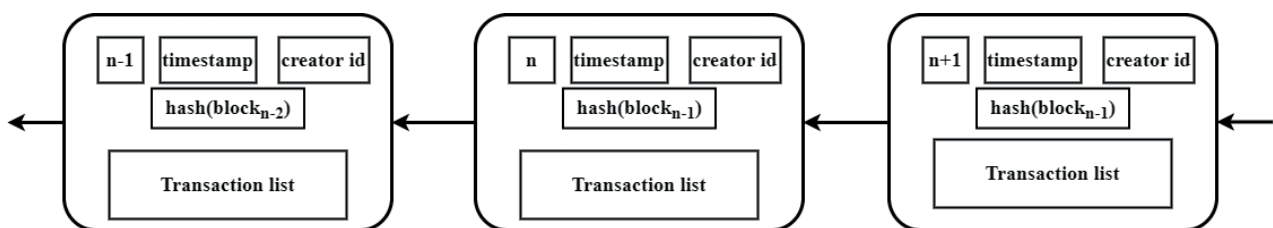
### 2. BACKGROUND

#### 2.1. Blockchain

Blockchain is relatively new technology that gained massive popularity by the appearance of Bitcoin cryptocurrency [14]. Soon, it became apparent that blockchain technology has potential for applications beyond cryptocurrency, and soon finds its uses in various areas such as healthcare [3], internet-of-things [10], digital rights management [17], supply chains [25], etc. Blockchain can be viewed as a distributed, decentralized append-only list of blocks, used in a peer-to-peer environment. Blockchain being distributed means that everyone who uses blockchain has a local copy of its whole content. Arguably, the most important characteristic of the blockchain is that it is decentralized, meaning that it doesn't need any centralized trusted authority to manage it. Instead, the whole management of the blockchain system is done by the users themselves [16]. Being

append-only means that once data is stored in blockchain, it can't be deleted or modified. Figure 1 illustrates part of a blockchain that consists of three consecutive blocks.

The users send data in form of transactions to the transaction pool. Special type of users, known as miners, select a number of transactions from the pool, verify them and group them in a block. Beside the transactions, the formed blocks contain data such as, block number, timestamp, the id of its creator, and cryptographic hash value of the previous block which is used to link all of the blocks in a type of linked list. However, because there is no centralized authority, the miner who forms a block must prove block's validity by executing consensus protocol [26] which is usually represented by a hard cryptographic puzzle, that a miner must solve. The solution of the puzzle is then embedded in the block, and the block is published. It must be noted that any user of the blockchain can decide to act as a miner. Once the block is published, the other users decide to accept it after they verify that the puzzle was solved correctly. However, it can happen that there are multiple valid blocks that are published simultaneously and in that case, every user may decide to accept any block that he wants. This can lead to creation of a fork in the list of blocks, which is usually resolved by users accepting the longest branch of the fork after certain period of time.



**Figure 1** Illustration showing an example of showing three consecutive blocks of a blockchain.

## 2.2. Smart contracts

The term "smart contract" was first created and used in the mid 1990s, by a computer scientist Szabo, who defined them as a group of promises in digital form, that also include defined protocols that enable parties to fulfil those promises [19]. However, only when the blockchain technology appeared smart contracts lived up to its full potential. The blockchain platform Ethereum [1] introduced support for customizable smart contracts, by implementing Turing-complete virtual machine, known as *Ethereum virtual Machine (EVM)*. Smart contracts represent a computer program that is executed on EVM, where every user of the Ethereum has an EVM implementation and executes the same code of the contract. Smart contracts are usually written using some high level programming language, most commonly the Solidity language. The code is then compiled into bytecode that is executed on EVM.

Smart contracts are submitted via transactions to the blockchain network, where they are verified by the miners and stored. [24]. The one who published the contract receives its address in the blockchain that can later be used to call methods that smart contracts implement. It should be noted, however, that because on the everyone must run the code of the called method, the execution can be rather costly. Blockchain mitigate this by having users pay a fee that is proportional to the number of instructions that are found in the method that was called.

## 3. APPLICATION OF BLOCKCHAIN TO SUPPLY CHAIN MANAGEMENT

A supply chain is a series of actions undertaken to deliver the goods from the initial supplier to the end customer through a series of organizations. It starts with the sourcing of materials and ends with the delivery of finished products. Along the way, each organization adds some value to the product. This process includes actions such as change of ownership, payments, transport of goods, security concerns, sharing documents and signing contracts between parties, customs clearances, and many others. Some of these actions can be automatized by using smart contracts and blockchain technology.

### 3.1. Traceability and verifiability

The most common application of blockchain to supply chain management is verifying the origin of a product and the way it was transported from its origin to the end customer. Currently, a customer usually has no way of

verifying the source of materials used in a product, how a product was transported, or whether the production is sustainable from the ecological point of view. Not knowing this information prevents the end customer from making an informed decision. Additionally, providing the customers with this information can potentially be a huge competitive advantage for any company, as transparency can increase the trust between the company and its customers.

One way of overcoming this problem and increasing transparency is the use of blockchain technology and *radio frequency identification technology (RFID)*. RFID is a wireless system composed of tags and readers. Tags are small devices consisting of integrated circuit and an antenna, which can store a certain amount of data. Readers can read or write data from/to tags by using radio waves.

During the product's life cycle, it will be in a position of different organizations, including suppliers, manufacturers, distributors, retailers, etc. Each of these organizations can log some important information about the product into the blockchain, using RFID as a unique identifier of the product. For example, an organization can store ownership data, time stamps, location data, product-specific data, data about a person responsible for handling the product, or environmental impact data on the blockchain [2]. Since the data is stored in blockchain, it can be verified in real-time. Another advantage is increased security, as the data can not be changed once it became a part of a block. Furthermore, in the case of a faulty product, it would be easy to determine the cause and time of the problem and to take the appropriate measures, simply by inspecting the data stored in blockchain.

This concept can be taken one step further by integrating blockchain with the *Internet of Things (IoT)* in order to enhance the quality control of products. For example, transporting perishable goods can have certain constraints on temperature, humidity, or exposure to light. By including sensors that can detect unfavorable conditions, we can prevent damaged products from being used in further production, by automatically adding the data about the damage into the blockchain. As manufacturers rarely have the data on how the product was handled by higher-tier suppliers, this can prevent the company from unknowingly buying an inferior product and potentially harm its reputation. Examples of blockchain being integrated into agricultural supply chains can be found in [7, 12, 13, 20, 21].

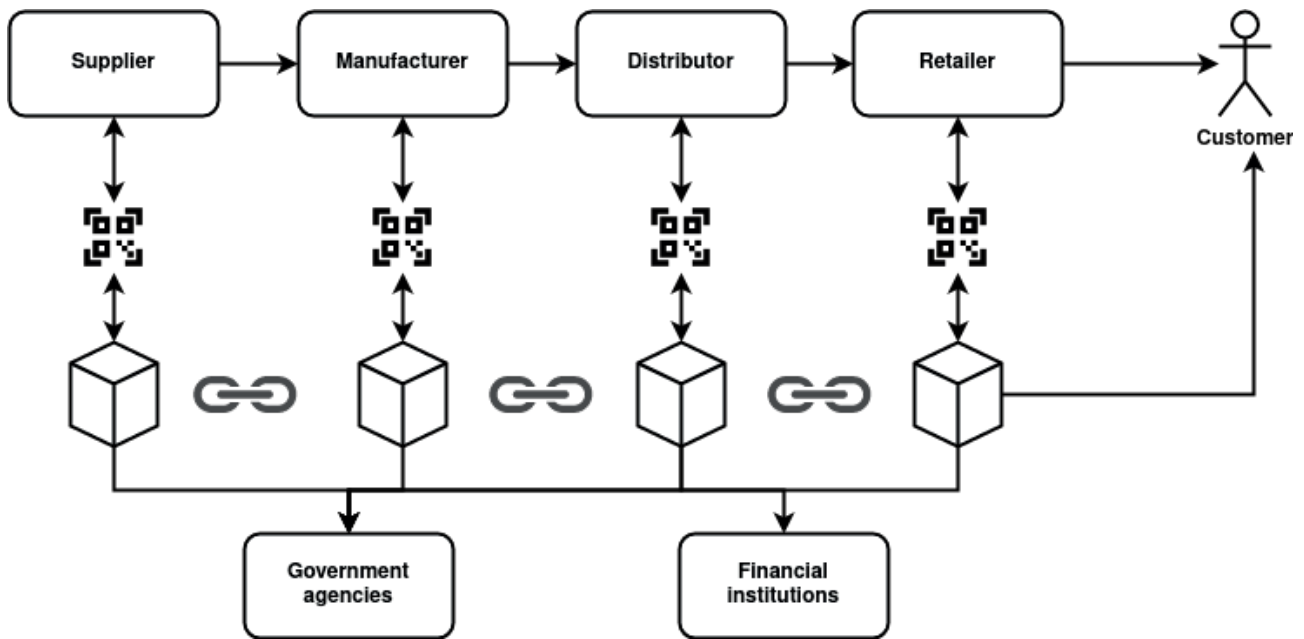
Additionally, being able to trace the origin of the product and materials may prevent product counterfeiting, as each customer can verify the manufacturer and the path the product took before it reached the customer. This is especially important in the pharmaceuticals industry, as counterfeit drugs can have a negative impact on the users health [5], and it is estimated that around 50% of the medications sold online are counterfeit [4]. In a paper by Tseng et al. [22], the authors proposed a blockchain-based method for the distribution of pharmaceutical drugs.

In Figure 2 we present an illustration of a simple supply chain that uses blockchain with RFID technology for tracing products. Each participant of the supply chain interacts with blockchain by verifying the origin of delivered goods and writing new information considering the product, using RFID as a unique identifier. A customer can access all the data about the product on the blockchain. Government agencies and financial institutions can also monitor the data written on the blockchain.

### 3.2. Information sharing

The modern supply chains are highly complex systems that require a lot of communication and paperwork. Because of this, these systems are prone to errors and delays, and potentially use more resources than necessary. For example, a shipment from East Africa to Europe can require over 200 interactions and involve over 30 individuals or organizations, and by removing some unnecessary interactions the global trade volume could increase up to 15% [23]. By using the blockchain and smart contracts to track shipments and share the information with all the parties involved in the supply chain, it is possible to avoid most of these unnecessary communications and paperwork. This approach has already been adopted by Maersk and IBM, who provide a platform for tracking cargo based on blockchain and IoT. All the information about the shipment, necessary documentation, and activities are stored in the blockchain, allowing all participants in the supply chain to effortlessly access all the data, eliminating the need for direct communication. As a result, there are fewer mistakes, no unnecessary delays, and the system is more resistant to fraud. Furthermore, as the information about shipment can be accessed in real time, better optimization of storage facilities might be possible.

Another advantage of information sharing comes in a form of improved demand forecasting. Demand forecasting is a process of estimating the future demand for products based on past and current demand. In the traditional supply chain, forecasting demand can be a challenging process, as much information can be inaccessible to a particular company. Therefore, a company can create a forecast only for itself, potentially



**Figure 2** Illustration of a simple supply chain that uses blockchain for tracing products

leading to overstocking its inventory as a safety measure. In contrast, if the information about demand is shared among all of the participants in a supply chain, collaborative demand management becomes much easier, as it allows the participants to create one shared demand forecast, based on all available data.

### 3.3. Payment Automation

Payment automation in a supply chain can speed up the entire process of delivery and eliminate the need for unnecessary paperwork. By creating a smart contract, participants in a supply chain can order materials or services, by specifying the conditions under which the contract is fulfilled. The contract can not be changed unless all of the involved parties agree on making a new contract. Once the conditions defined in a contract are met, the payment is automatically released, eliminating the need for a third-party service, like a financial institution. This approach has already been presented in [8, 9, 15].

Another advantage of using smart contracts for automating payments is to overcome the lack of trust between two participants in the supply chain. For example, in the construction industry, it is not uncommon for participants to delay payments, or even forfeit them entirely [18]. As smart contracts are self-executing, the payment-related issues can be avoided, because the payment is released as soon as the conditions specified in the contract are met.

## 4. Potential issues

### 4.1. Security

In the traditional blockchain, each node can propose a transaction which would then be used to create a new block. In a supply chain, this approach can potentially create some security issues, as not all nodes have to be trustworthy. To remedy this issue, Gao et al. [6] proposed separating nodes into several types, based on their role in the supply chain. *Ordinary users* are the direct participants in the supply chain who can submit new requests for material, track shipments, and analyze the data. *Third-party users*, usually government and insurance agencies, monitor the data in the blockchain. *Supporting entities* interact with blockchain to provide particular services, like payment processing. Only these three types of users are allowed to submit new records to the blockchain. Additionally, there is the fourth type of user called *helpers* who maintain the blockchain by performing Proof-of-work to create new blocks. Like in the traditional blockchain, there is no limitation on who can be a helper, anyone interested can contribute to the creation of a new block.

Another security issue is the unauthorized access to sensitive information. Even though the transparency and easy access to data in the supply chain is generally desirable, some data stored in the blockchain might contain sensitive information that shouldn't be accessed by companies' competition or other unfriendly parties. This information can be encrypted, preventing unauthorized users and helpers from reading it. On the other hand, the



user that created the encrypted data can make a list of authorized users who have a legitimate reason to read the data and grant them an access in cooperation with the *identity management component* [6].

## 4.2. Storage

The number of individual products that are being shipped daily is so high that it is virtually impossible to estimate. Considering that most of the products have to go through several organizations until they reach their final destination, the amount of data needed for tracking all of the products can be incredibly high. For example, tracking only one million products through ten organizations, where each organization adds only 200 bytes of information for each product to the blockchain, would result in 1.86GB of used storage space. Taking into account that the actual number of products is vastly higher, that products can travel through many more organizations, and that the amount of data written for each product can be considerably higher, it is easy to see that the storage of all the data can be problematic.

As a solution to this problem, Kumar and Tripathi [11] proposed using the *Interplanetary File System (IPFS)* as blockchain storage. IPFS is a decentralized, peer-to-peer network for storing and sharing data. After a miner chooses certain transactions for creating the next block, those transactions are stored in IPFS, and the miner receives a 46 bytes long hash value, that can be written into a block. Because the same data in the IPFS results in the same hash value, potential frauds on the blockchain can be easily detected.

## 5. CONCLUSION

In this paper, we presented some of the possibilities of applying blockchain technology to supply chain management. As the supply chains grow exceedingly complex, there is a need for reducing the amount of direct communication and for more transparency in a process of delivering goods. Potentially, blockchain technology could help with both of these problems, by allowing the involved parties to create smart contracts or to track materials and products in a more secure and fraud resistant way.

We presented several ways that blockchain technology could be incorporated into supply chain management in order to mend some of the existing problems. Furthermore, we presented some of the issues with the adoption of this technology, together with the possible solutions.

## Acknowledgement

This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, Agreement No. 451-03-9/2021-14/200029 and by the Science Fund of the Republic of Serbia, Grant "AI4TrustBC: Advanced Artificial Intelligence Techniques for Analysis and Design of System Components Based on Trustworthy BlockChain Technology".

## REFERENCES

- [1] "Ethereum Whitepaper", available at: <https://ethereum.org/en/whitepaper/> [Online], Accessed 07.06.2022.
- [2] Abeyratne, S. A., & Monfared, R. P. (2016). *Blockchain ready manufacturing supply chain using distributed ledger*. International Journal of Research in Engineering and Technology, 5(9), 1-10.
- [3] Agbo, C. C., Mahmoud, Q. H., & Eklund, J. M. (2019, June). Blockchain technology in healthcare: a systematic review. In *Healthcare* (Vol. 7, No. 2, p. 56). Multidisciplinary Digital Publishing Institute.
- [4] Behner, P., Hecht, M.-L., & Wahl, F. (2017). *Fighting counterfeit pharmaceuticals: New defenses for an underestimated—and growing—menace*. Retrieved June 4, 2022, from <https://www.strategyand.pwc.com/reports/counterfeit-pharmaceuticals>.
- [5] Dujak, D., & Sajter, D. (2019). *Blockchain applications in supply chain*. In *SMART supply network* (pp. 21-46). Springer, Cham.
- [6] Gao, Z., Xu, L., Chen, L., Zhao, X., Lu, Y., & Shi, W. (2018). CoC: A unified distributed ledger based supply chain management system. *Journal of Computer Science and Technology*, 33(2), 237-248.
- [7] George, R. V., Harsh, H. O., Ray, P., & Babu, A. K. (2019). Food quality traceability prototype for restaurants using blockchain and food quality data index. *Journal of Cleaner Production*, 240, 118021.
- [8] Kaid, D., & Eljazzar, M. M. (2018, December). Applying blockchain to automate installments payment between supply chain parties. In *2018 14th International Computer Engineering Conference (ICENCO)* (pp. 231-235). IEEE.

- [9] Kaid, D., & Eljazzar, M. M. (2018, December). Applying blockchain to automate installments payment between supply chain parties. In 2018 14th International Computer Engineering Conference (ICENCO) (pp. 231-235). IEEE.
- [10] Khan, M. A., & Salah, K. (2018). IoT security: Review, blockchain solutions, and open challenges. *Future generation computer systems*, 82, 395-411.
- [11] Kumar, R., & Tripathi, R. (2019, November). Implementation of distributed file storage and access framework using IPFS and blockchain. In 2019 Fifth International Conference on Image Information Processing (ICIIP) (pp. 246-251). IEEE.
- [12] Lin, Q., Wang, H., Pei, X., & Wang, J. (2019). Food safety traceability system based on blockchain and EPCIS. *IEEE access*, 7, 20698-20707.
- [13] Mondal, S., Wijewardena, K. P., Karuppuswami, S., Kriti, N., Kumar, D., & Chahal, P. (2019). Blockchain inspired RFID-based information architecture for food supply chain. *IEEE Internet of Things Journal*, 6(3), 5803-5813.
- [14] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*, 21260.
- [15] Nanayakkara, S., Perera, S., Senaratne, S., Weerasuriya, G. T., & Bandara, H. M. N. D. (2021, June). Blockchain and smart contracts: A solution for payment issues in construction supply chains. In *Informatics* (Vol. 8, No. 2, p. 36). Multidisciplinary Digital Publishing Institute.
- [16] Nofer, M., Gomber, P., Hinz, O., & Schiereck, D. (2017). Blockchain. *Business & Information Systems Engineering*, 59(3), 183-187.
- [17] Qureshi, A., & Megías Jiménez, D. (2020). Blockchain-based multimedia content protection: review and open challenges. *Applied Sciences*, 11(1), 1.
- [18] Ramachandra, T., & Rotimi, J. O. (2011). The nature of payment problems in the New Zealand construction industry. *Australasian Journal of Construction Economics and Building*, The, 11(2), 22-33.
- [19] Szabo, N. (1996). Smart contracts: building blocks for digital markets. *EXTROPY: The Journal of Transhumanist Thought*, (16), 18(2), 28.
- [20] Tian, F. (2016, June). An agri-food supply chain traceability system for China based on RFID & blockchain technology. In 2016 13th international conference on service systems and service management (ICSSSM) (pp. 1-6). IEEE.
- [21] Tian, F. (2017, June). A supply chain traceability system for food safety based on HACCP, blockchain & Internet of things. In 2017 International conference on service systems and service management (pp. 1-6). IEEE.
- [22] Tseng, J. H., Liao, Y. C., Chong, B., & Liao, S. W. (2018). Governance on the drug supply chain via gcoin blockchain. *International journal of environmental research and public health*, 15(6), 1055.
- [23] Van Kralingen, B. (2018). *IBM, Maersk joint blockchain venture to enhance global trade*. Retrieved June 4, 2022, from IBM.COM: <https://www.ibm.com/blogs/think/2018/01/maersk-blockchain/>.
- [24] Wang, S., Ouyang, L., Yuan, Y., Ni, X., Han, X., & Wang, F. Y. (2019). Blockchain-enabled smart contracts: architecture, applications, and future trends. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(11), 2266-2277.
- [25] Wang, Y., Han, J. H., & Beynon-Davies, P. (2018). Understanding blockchain technology for future supply chains: a systematic literature review and research agenda. *Supply Chain Management: An International Journal*.
- [26] Zhang, C., Wu, C., & Wang, X. (2020, May). Overview of Blockchain consensus mechanism. In *Proceedings of the 2020 2nd International Conference on Big Data Engineering* (pp. 7-12).

## ODRŽIVO UPRAVLJANJE LANCEM SNABDEVANJA

### SUSTAINABLE SUPPLY CHAIN MANAGEMENT

MILICA GRUJIĆ<sup>1</sup>, ALEKSANDAR ĐUKIĆ<sup>2</sup>

<sup>1</sup>Univerzitet u Beogradu – Fakultet organizacionih nauka, milicagrujic991@gmail.com

<sup>2</sup>Univerzitet u Beogradu – Beogradska akademija poslovnih i umetničkih strukovnih studija, aleksandardjukic1453@gmail.com

**Rezime:** Poslovne organizacije se sve više fokusiraju na kupovinu i nabavku proizvoda i usluga na način koji pozitivno utiče na smanjenje negativnih uticaja na životnu sredinu, društvo i privredu. U skladu sa navedenim, koncept održivog upravljanja lancem snabdevanja je privukao pažnju naučnika i praktičara zbog svog značaja za ekološku, društvenu i korporativnu odgovornost kroz ekonomske performanse. Cilj ovog rada je definisanje koncepta održivog upravljanja lancem snabdevanja i utvrđivanje značaja uvođenja principa održivosti u koncept upravljanja lancem snabdevanja

**Cljučne reči:** Održivo upravljanje lancem snabdevanja, Lanac snabdevanja, Održivost, Održivi razvoj

**Abstract:** The business organizations are increasingly focusing on buying and supplying of products and services in a manner to reduce the adverse impacts on the environment, society, and economy. In view of the above, the concept of sustainable supply chain management has received attention of the industry and academia due to its importance on environmental, social and corporate responsibility through economic performance. The aim of this paper is to define the concept of sustainable supply chain management and determine the importance of introducing the principle of sustainability in the concept of supply chain management

**Keywords:** Sustainable Supply Chain Management, Supply Chain, Sustainability, Sustainable Development

#### 1. UVOD

Jedan od ključnih pokretača promena u društvu i poslovanju je koncept održivog razvoja, koji sve više dobija na značaju, kako u poslovnom svetu, tako i u samom društvu. Pokretači koji su uticali na popularizaciju koncepta održivog razvoja uključuju snabdevanje i karakteristike potražnje koje se odnose na potrošnju energije, povećano razumevanje faktora i okolnosti koje su dovele do klimatskih promena, kao i veću transparentnost organizacija u vezi njihovih društvenih i ekoloških uticaja. Ova pitanja su postala veoma značajna za menadžere, jer stajholderi sve više zahtevaju da se organizacije bave i upravljaju društvenim i ekološkim pitanjima koja utiču na njihove operacije [2]. Stoga, savremeni uslovi poslovanja zahtevaju od organizacija da u svoje poslovne aktivnosti sve više uključuju koncept održivog razvoja, čiji je pozitivan uticaj na poslovanje dokazan kako teorijski, tako i praktično. Neke od koristi od usvajanja principa održivosti, odnosno društveno odgovornog poslovanja, podrazumevaju: uticaj na tržište (uticaj na korporativni imidž i reputaciju, diferencijacija i unapređenje konkurentnosti, privlačenje novih kupaca, uticaj na lojalnost i zadržavanje postojećih kupaca); uticaj na radno okruženje (uticaj na privlačenje kvalitetnih zaposlenih, uticaj na motivaciju i posvećenost zaposlenih, smanjen apsentizam i povrede na radu, povećanje produktivnosti, uticaj na zadržavanje kvalitetnih zaposlenih); uticaj na poslovanje (smanjenje operativnih troškova poslovanja, smanjenje rizika, povećanje prihoda usled rasta prodaje i/ili tržišnog učešća, unapređenje kvaliteta proizvoda i usluga, bolje finansijske performanse, uvećanje vrednosti akcija, pozitivan uticaj na pristup kapitalu – odgovorno investiranje, unapređenje efikasnosti upravljanja otpadom) [17].

Na osnovu navedenog, može se zaključiti da više nije dovoljno da cilj bude isključivo generisanje profita, već da organizacije treba i da vrate nešto zajednici u kojoj posluju, što se prvenstveno odnosi na minimizaciju negativnih uticaja na životnu sredinu i preuzimanje odgovornosti za ponašanje svojih dobavljača po pitanjima kao što su dečiji rad, zdravlje, bezbednost i zagađenje [18]. Poslednjih godina raste interesovanje naučnika i praktičara za uključivanje principa održivosti u lanac snabdevanja, odnosno za lanac snabdevanja koji uključuje tri stuba održivosti (ekonomija, društvo i ekologija). Menadžeri koji upravljaju lancem snabdevanja

su u povoljnoj poziciji da utiču kako pozitivno, tako i negativno na društvene i ekološke performanse kroz izbor i razvoj dobavljača, izbor prevoznika, određivanje putanje vozila, odluke o lokaciji izbor pakovanja [2]. Ispunjavanje ekoloških i društvenih standarda u svim fazama lanca snabdevanja, značajno utiče na poboljšanje „održivih“ performansi organizacije [15].

Cilj rada predstavlja definisanje pojma održivog upravljanja lancem snabdevanja i utvrđivanje značaja uvođenja principa održivosti u koncept upravljanja lancem snabdevanja.

## 2. POJAM ODRŽIVOSTI

Za popularizaciju koncepta održivog razvoja zaslužna je „Svetska komisija za životnu sredinu i razvoj“ (engl. *World Commission on Environment and Development*) poznatija i kao „Brundtland komisija“, koja je 1987. godine predstavila izveštaj „Naša zajednička budućnost“ (engl. *Our Common Future*) u kome je održivi razvoj definisan kao: „Razvoj koji zadovoljava potrebe sadašnjih generacija bez ugrožavanja sposobnosti budućih generacija da zadovolje svoje potrebe“ [19]. Iako postoji više desetina različitih definicija pojma održivosti/održivog razvoja [14], ova definicija je postala najzastupljenija i najšire prihvaćena u literaturi. Međutim, autori [9] navode da je navedena definicija prilično široka i kao takva predstavlja izazov za savremene organizacije, kao što je identifikacija pojedinačne uloge organizacije u široj makroekonomskoj perspektivi ili nedostatak uputstava o tome kako efikasno identifikovati sadašnje i buduće potrebe i prakse. Pored toga, postoje poteškoće u odabiru tehnologija i resursa za ispunjenje sadašnjih i budućih potreba. Stoga, izazovi u razvoju razumevanja značaja balansiranja odgovornosti stejkholdera u okviru lanca snabdevanja predstavlja primarni izazov za savremene organizacije.

Autor [11] navodi da se koncept održivog razvoja zasniva na sledećim idejama:

- Povezanost i međusobna zavisnost ekonomije i životne sredine;
- Uticaji aktivnosti sadašnjih generacija na buduće, odnosno orijentisanost na budućnost;
- Doprinos smanjenju štetnog uticaja na životnu sredinu;
- Jednakost između generacija;
- Poboljšanje kvaliteta života društva u celosti;
- Reformisanje i organizovanje institucija sa ciljem da se omogući da celokupno društvo učestvuje u procesu odlučivanja.

Sa druge strane, mikroekonomski pristup konceptu održivosti je istražen u oblasti menadžmenta, operacija i inženjerstva. Uloga organizacije u održivom razvoju uglavnom se pominje u kontekstu odgovornosti prema društvu, odnosno kao obaveza organizacija da smanje negativne uticaje poslovanja na životnu sredinu i lokalnu zajednicu u kojoj posluju. Autor [1] navodi da principe održivog razvoja treba posmatrati kao povoljne prilike (šanse) za unapređenje poslovanja, odnosno da usvajanje principa održivosti u poslovnu strategiju treba posmatrati kao šansu za stvaranje vrednosti – i za poslovnu organizaciju i za društvo. Poslovne organizacije imaju veliki uticaj na životnu sredinu, zaposlene i celokupno društvo izborom materijala i dobavljača, korišćenjem zemljišta, proizvodnim procesima (stvaranjem otpada i zagađenja), geografskom lokacijom, finansijskim ugovorima, sistemima upravljanja, zapošljavanjem i uslovima rada, odnosima sa kupcima, aktivnostima u lokalnoj zajednici, informisanjem i lobiranjem. Na osnovu navedenog, može se zaključiti da poslovne organizacije i njihove aktivnosti imaju značajnu ulogu u procesu održivog razvoja.

Održivost organizacije se uglavnom zasniva na tri komponente: prirodno okruženje, društvo i ekonomske performanse, što je u skladu sa istraživanjima autora [5], koje su definisali pojam trostrukog krajnjeg rezultata (engl. *Triple Bottom Line*) ili 3P (engl. *People, Planet, Profit*). Odnos između navedenih komponenata/dimenzija održivog razvoja podrazumeva njihovu međuzavisnost i ravnotežu (slika 1). Koncept trostrukog krajnjeg rezultata sve više dobija na značaju kao okvir za merenje performansi organizacije u kontekstu ekonomskog, društvenog i ekološkog uticaja, što znači da uspeh organizacije ne treba da se meri samo kroz finansijsku dobit, već i osnovu društvenih i ekoloških performansi, što znači da koncept podrazumeva upravljanje društvenim, ekološkim i finansijskim pitanjima, objedinjujući sve performanse poslovanja koje su se ranije merile zasebno [12].



Slika 1. Održivost: trostruki krajnji rezultat

### 3. OD TRADICIONALNOG DO ODRŽIVOG UPRAVLJANJA LANCEM SNABDEVANJA

Upravljanje lancem snabdevanja je definisano kao: „integracija ključnih poslovnih procesa koji obuhvataju sve aktivnosti od kupovine sirovina do isporuke krajnjeg proizvoda kupcima“ [8]. Na ovaj način se stejkholderi povezuju i postaju deo lanca. Upravljanje lancem snabdevanja predstavlja sredstvo za unapređenje efikasnosti aktivnosti organizacije od nabavke sirovina do isporuke krajnjeg proizvoda korisnicima i izazov za organizacije, koji postaje još zahtevniji proces kada se poveže sa ciljevima i politikama održivosti [9].

Prema [10] sledeće aktivnosti su neophodne za uspešno upravljanje lancem snabdevanja:

- Integrirano ponašanje – usklađivanje ponašanja i potreba svih partnera u lancu snabdevanja (dobavljača, proizvođača, prevoznika) u cilju adekvatnog odgovora na potrebe krajnjeg korisnika;
- Međusobna razmena informacija – otvoreno deljenje informacija (kao što su nivoi zaliha, predviđanja, strategije unapređenja prodaje i marketinške strategije) koje dovodi do smanjenja nesigurnosti između partnera u lancu snabdevanju i rezultira poboljšanim performansama;
- Međusobno deljenje rizika i nagrada – što je od velikog značaja za dugoročni fokus i saradnju među članovima lanca snabdevanja;
- Saradnja – odnosi se na slične ili komplementarne aktivnosti koje obavljaju organizacije u poslovnom odnosu kako bi proizvele superiorne zajedničke ili jedinstvene ishode koji se obostrano očekuju tokom vremena;
- Isti cilj i isti fokus na zadovoljavanje potreba kupaca – ukoliko svi akteri u lancu snabdevanja imaju isti fokus, lanac snabdevanja i aktivnosti unutar lanca će biti značajno unapređene;
- Integracija procesa – implementacija upravljanja lancem snabdevanja zahteva integraciju svih procesa od nabavke, preko proizvodnje do distribucije;
- Izgradnja i održavanje dugoročnih odnosa sa partnerima – formiranje strateških saveza sa partnerima u lancu snabdevanja kao što su dobavljači, kupci ili posrednici (npr. usluge transporta i/ili skladištenja) obezbeđuje konkurentsku prednost kroz stvaranje vrednosti za kupce.

Kao što je već rečeno, globalizacija zahteva od poslovnih organizacija takav pristup konceptu upravljanja lancem snabdevanja koji ne uzima u obzir samo ekonomska pitanja. Stoga, održivo upravljanje lancem snabdevanja postaje sve značajniji koncept za mnoge organizacije koje posluju u različitim industrijama [15]. Autori [3] definišu održivo upravljanje lancem snabdevanja kao: „strateško, transparentno integrisanje i ostvarivanje ekonomskih, ekoloških i društvenih ciljeva organizacije sistemskom koordinacijom ključnih poslovnih procesa, što ima za cilj unapređenje dugoročnih ekonomskih performansi organizacije i njenog lanca snabdevanja“. Autori [16] definišu održivo upravljanje lancem snabdevanja kao: „upravljanje tokovima materijala, informacija i kapitala i saradnju organizacija u lancu snabdevanja, uzimajući u obzir ciljeve sve tri dimenzije održivog razvoja koje proizilaze iz zahteva kupaca i ostalih stejkholdera.“ Prema [6] održivo upravljanje lancem snabdevanja definiše se kao skup menadžerskih praksi koje obuhvataju:

- Uticaj na životnu sredinu kao imperativ;
- Razmatranje svih faza lanca vrednosti za svaki proizvod;
- Multidisciplinarnu perspektivu koja obuhvata ceo životni ciklus proizvoda.

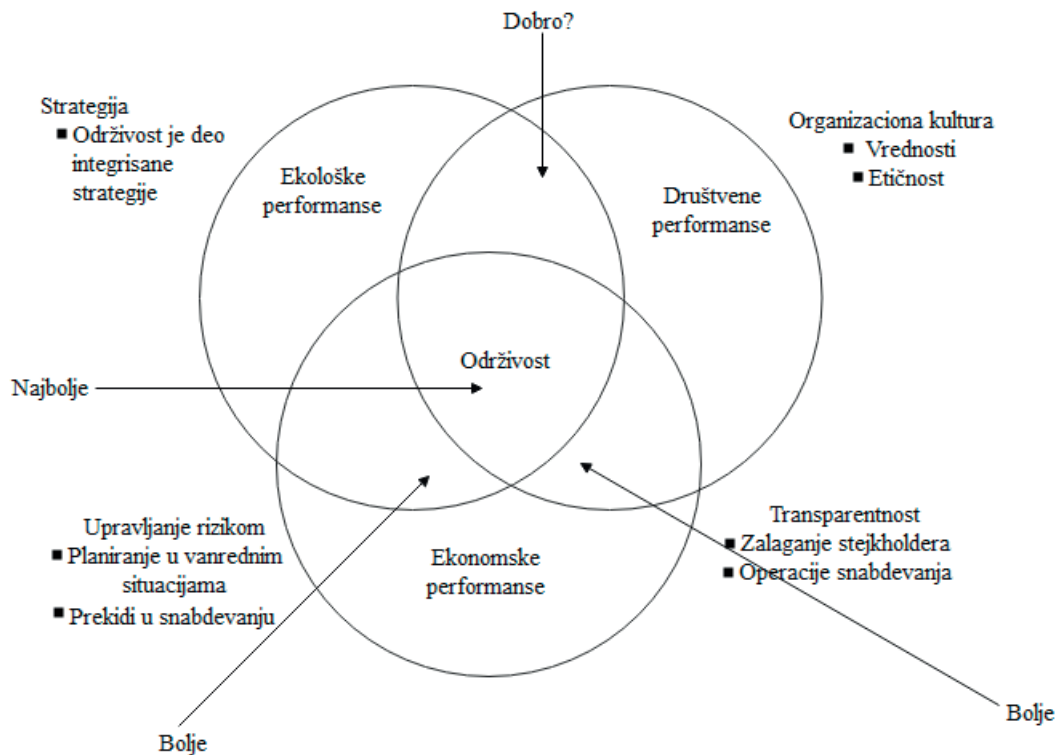
Organizacije moraju da posmatraju uticaj svojih aktivnosti na životnu sredinu kao sastavni deo donošenja odluka, a ne kao ograničenje koje je nametnuto zakonskim regulativama, različitim oblicima pritiska javnosti ili sredstva koje će doprineti kratkotrajnom poboljšanju imidža (sredstvo kojim će se organizacija predstaviti u javnosti kao „zelena“). Nadalje, organizacije moraju da obrate pažnju na ekološke aspekte celog lanca snabdevanja, uključujući dobavljače, distributere, partnere i kupce. Pored toga, pogled organizacije na održivost mora prevazići usku funkcionalnu perspektivu i obuhvatiti širi pogled koji integriše različite probleme i njihova potencijalna rešenja širom funkcionalnih granica [6].

Autori [16] kao glavne pokretače održivog upravljanja lancem snabdevanja navode eksterni pritisak i inicijative određenih stejkholdera, od kojih su posebno značajni kupci (upravljanje lancem snabdevanja postoji samo ako su proizvodi i usluge „prihvaćeni“ od strane kupaca) i vlada (bilo da se radi o lokalnim ili nacionalnim oblicima vlasti). Ukoliko je centralna organizacija u lancu snabdevanja izložena eksternom pritisku stejkholdera, to se obično odražava na dobavljače. Gledajući ceo lanac snabdevanja (ili životni ciklus) proizvoda, centralna organizacija često ima mnogo veću ulogu nego što je potrebno, uglavnom iz ekonomskih razloga, što može dovesti do pojave različitih faktora koji podržavaju ili ometaju saradnju sa dobavljačima (što se odnosi na posedovanje informacija ekološkim i društvenim performansama u različitim fazama proizvodnje, kao i poboljšanju performansi glavnih dobavljača). U skladu sa navedenim isti autori [16] navode dve strategije za održivo upravljanje lancem snabdevanja:

- Upravljanje dobavljačima za rizike i performanse – organizacije koje slede ovu strategiju obično imaju problem sa strahom od gubitka reputacije ukoliko se pojave problemi. Ukoliko dođe do problema, uzimaju se u obzir ekološki i društveni parametri koji dopunjuju ekonomski zasnovanu evaluaciju dobavljača;
- Upravljanje lancem snabdevanja za održive proizvode – koje zahteva definisanje standarda zasnovanih na životnom ciklusu za društvene i ekološke performanse, koji se dalje primenjuju u celom lancu snabdevanja.

Autori [3] su definisali teorijski okvir održivog upravljanja lancem snabdevanja (slika 2). U osnovi konceptualizacije je trostruki krajnji rezultat. Isti autori naglašavaju značaj održivog upravljanja lancem snabdevanja kao koncepta koji dugoročno utiče na poboljšanje ekonomskih performansi i dugoročnog napretka organizacije. Pored značaja trostrukog krajnjeg rezultata za efikasno održivo upravljanje lancem snabdevanja autori su identifikovali i četiri prateća aspekta (slika 2):

- Strategija – identifikacija pojedinačnih inicijativa održivog upravljanja lancem snabdevanja koje podržavaju u skladu sa opštom strategijom održivosti organizacije;
- Upravljanje rizikom – uključuje planiranje nepredviđenih aktivnosti u svim fazama lanca snabdevanja;
- Organizaciona kultura – koja uključuje visoke etičke standarde i očekivanja (što je polazna tačka za izgradnju održivog upravljanja lancem snabdevanja), poštovanje prema društvu (i unutar i izvan organizacije) i svest o očuvanju prirodnih resursa;
- Transparentnost – u smislu proaktivnog angažovanja i komunikacije sa ključnim stejkholderima i vidljivosti aktivnosti u svim fazama lanca snabdevanja.



Slika 2. Okvir za održivo upravljanje lancem snabdevanja

Na osnovu slike 2 autori [3] navode sledeće poslovne prednosti održivog upravljanja lancem snabdevanja:

- Ušteda troškova usled smanjenja otpada od pakovanja, odnosno pakovanja koja se mogu ponovo koristiti i sastaviti;
- Smanjeni troškovi koji se odnose na zdravlje i bezbednost, zapošljavanje i odlazak zaposlenih, usled bezbednijeg skladištenja i transporta, kao i boljih uslova rada;
- Niži troškovi rada – bolji uslovi rada pozitivno utiču na motivaciju i produktivnost i smanjuju odsutnost zaposlenih u lancu snabdevanja;
- Proaktivno oblikovanje budućih zakona i regulativa – organizacije koje se proaktivno bave društvenim i ekološkim problemima mogu da utiču na vladine regulative;
- Smanjeni troškovi, kraće vreme isporuke i bolji kvalitet proizvoda – povezan sa implementacijom standarda ISO 14000, koji pružaju okvir za sisteme upravljanja životnom sredinom;
- Poboljšanje korporativne reputacije – usvajanje principa održivog razvoja u poslovanje pozitivno utiče na privlačenje kupaca, kvalitetnih zaposlenih, dobavljača i akcionara.

Autori [16] sumiraju osnovne razlike između održivog upravljanja lancem snabdevanja u poređenju sa tradicionalnim:

- Održivo upravljanje lancem snabdevanja uzima u obzir širi opseg različitih problema;
- Održivo upravljanje lancem snabdevanja bavi se širim skupom ciljeva performansi, uzimajući u obzir ekološku i društvenu dimenziju održivog razvoja;
- Naglašena potreba za saradnjom sa partnerima u održivom lancu snabdevanja.

#### 4. ZAKLJUČAK

Mnoge poslovne organizacije širom sveta žele da uključe prakse poput ekološke odgovornosti i brige o lokalnoj zajednici u svoje lance snabdevanja. Veliki korak ka tome je usvajanje principa održivosti u poslovanje. Koncept održivosti je prilično širok i razumevanje njegove interakcije sa upravljanjem lancem snabdevanja je ključno za uspešno upravljanje lancem snabdevanja u promenljivom i dinamičnom poslovnom okruženju. Održivo upravljanje lancem snabdevanja obuhvata integrisane perspektive ekonomskih, ekoloških i društvenih razmatranja. Za uspešnu implementaciju održivog upravljanja lancem snabdevanja, poslovne organizacije treba da sarađuju sa svim svojim stejkholderima uz optimalno korišćenje resursa, informacija i sredstava za maksimiziranje ponude vrednosti kupcima i profitabilnost lanca snabdevanja [13]. U procesu se minimizira uticaj na životnu sredinu i poboljšava društvena jednakost i blagostanje za sve stejkholdere.

## 5. LITERATURA

- [1] Baumgartner, R. J. (2014). Managing Corporate Sustainability and CSR: A Conceptual Framework Combining Values, Strategies and Instruments Contributing to Sustainable Development. *Corporate Social Responsibility and Environmental Management*, 21, 258-271. doi:10.1002/csr.1336
- [2] Carter, C. R., & Easton, L. P. (2011). Sustainable supply chain management: evolution and future directions. *International Journal of Physical Distribution & Logistics Management*, 41(1), 46-62. doi:10.1108/09600031111101420
- [3] Carter, C. R., & Rogers, D. S. (2007). A framework of sustainable supply chain management: moving toward new theory. *International Journal of Physical Distribution & Logistics Management*, 38(5), 360-387. doi:10.1108/09600030810882816
- [4] Diesendorf, M. (2000). Sustainability and Sustainable Development. In D. Dunphy, J. Benveniste, A. Griffiths, & P. Sutton, *The corporate challenge of the 21st century* (pp. 19-37). Sydney: Allen & Unwin.
- [5] Elkington, J., & Rowlands, I. H. (1999). Partnerships from Cannibals with Forks: The Triple bottom line of 21st Century Business. *Alternatives Journal*, 25(4), 37-51.
- [6] Gupta, S., & Palsule-Desai, O. D. (2011). Sustainable supply chain management: Review and research opportunities. *IIMB Management Review*(23), 234-245.
- [7] Klarin, T. (2018). The Concept of Sustainable Development: From its Beginning to the Contemporary Issues. *Zagreb International Review of Economics & Business*, 21(1), 67-94. doi:10.2478/zireb-2018-0005
- [8] Lambert, D. M., & Cooper, M. C. (2000). Issues in Supply Chain Management. *Industrial Marketing Management*, 29(1), 65-83.
- [9] Laurin, F., & Fantazy, K. (2017). Sustainable supply chain management: a case study at IKEA. *Transnational Corporations Review*, 1-10. doi:DOI: 10.1080/19186444.2017.1401208
- [10] Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. *Journal of Business logistics*, 22(2), 1-25.
- [11] Milutinović, S. (2012). *Politike održivog razvoja*. Niš: Fakultet zaštite na radu u Nišu.
- [12] Norman, W., & MacDonald, C. (2004). Getting to the Bottom of "Tripple Bottom Line". *Business Ethics Quarterly*, 14(2), 243-262.
- [13] Panigrahi, S., Bahinipati, B., & Jain, V. (2018). Sustainable supply chain management: A review of literature and implications for future research. *Management of Environmental Quality: An International Journal*, 30(5), 1001-1049. doi:10.1108/MEQ-01-2018-0003
- [14] Rogers, P. P., Jalal, K. F., & Boyd, J. A. (2008). *An Introduction to Sustainable Development*. London: Earthscan.
- [15] Seuring, S. (2013). A review of modeling approaches for sustainable supply chain management. *Decision Support Systems*, 1513-1520. doi:10.1016/j.dss.2012.05.053
- [16] Seuring, S., & Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. *Journal of Cleaner Production*, 16, 1699-1710. doi:10.1016/j.jclepro.2008.04.020
- [17] Vlastelica Bakić, T., Krstović, J., & Cicvarić Kostić, S. (2012). Poslovna opravdanost društveno odgovornog poslovanja. *Marketing*, 43(3), 191-198.
- [18] Walker, H. (2012). Sustainable supply chain management across the UK private sector. *Supply Chain Management*, 17(1), 15-28. doi:10.1108/13598541211212177
- [19] World Commission on the Environment and Development (1987). *Our Common Future*. Oxford: Oxford University Press.





# MATEMATIČKO PROGRAMIRANJE

---

## MATHEMATICAL PROGRAMMING



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business





## RELEVANT REASONING

MIRJANA ILIĆ<sup>1</sup>

<sup>1</sup> University of Belgrade - Faculty of Economics and Business, mirjanailic@ekof.bg.ac.rs

**Abstract:** *The notion which, with no doubt, plays the most important role in our deductive inferences is "if ... then ...". In everyday discourse, it is related to natural language conditionals, whereas in classical logic, it corresponds to the logical connective implication, which we symbolically denote by  $\rightarrow$ . However, natural language conditionals may have very different properties from the classical implication. In this paper, we shall give a short overview, mostly based on [2] and [19], on how such differences have motivated the formulation of Relevant logics.*

**Keywords:** *Implication, Paradoxes, Relevant Logics*

### 1. INTRODUCTION

The notion of logical implication is variously expressed by "if ... then", "implies", "entails", as well as "therefore", "hence", "consequently", "it follows that", etc. in natural language sentences. However, the meaning of natural language sequences may be crucially different from the meaning of the logical implication. For example, the logical implication is transitive: we have that from the premise that  $P \rightarrow Q$  and the premise that  $Q \rightarrow R$ , we have that  $P \rightarrow R$ . On the other hand, "if ... then ..." in conditional sequences is not. Assuming that the following two sentences are correct:

*If I save up enough money, then I'll buy a car,*

*If I buy a car, then I'll be completely broke,*

if transitivity of conditionals were to hold, we would be able to infer:

*If I save up enough money, then I'll be completely broke,*

which doesn't make any sense.

Consider now, the following example.

*The sum of the squares on the legs of a right triangle is equal to the square on the hypotenuse. Therefore,  $2+2=4$  or  $2+2 \neq 4$ .*

This inference is valid, but it is not a good proof, since the conclusion does not, in any intuitive sense, follow from the premise. This is a type of so-called *non-sequiturs* (in Latin, non-sequitur means "it does not follow"). Some non-sequitur statements are logical truths, as above, however, they are counterintuitive, whereas some of them are logically invalid, although they may seem intuitively correct. For example:

*If my dog barks, then he is hungry. My dog doesn't bark. Therefore, he is not hungry.*

In familiar symbolic notation, this example corresponds to:

*If  $p \rightarrow q$  and  $\neg p$ , then  $\neg q$ ,*

which is not logically correct. Really, when  $p$  is false, then  $p \rightarrow q$  is true, regardless of the value of  $q$ , so the above conclusion is not correct.

In the next section, we shall give some more examples of logically true, but counterintuitive proofs, which are known as *paradoxes of implication*.

### 2. PARADOXES OF IMPLICATION

In classical logic, where each proposition takes exactly one of two truth values, *true*, denoted by 1, or *false*, denoted by 0, every logical connective is *truth-functional*. As for the implication, this means that the value of  $p \rightarrow q$  is defined with respect to the truth values of its "if" part  $p$ , which is called the *antecedent* and its "then" part  $q$ , which is called the *consequent*. Namely, the implication  $p \rightarrow q$  is defined by the following truth table:

$p$	$q$	$p \rightarrow q$
1	1	1
1	0	0
0	1	1
0	0	1

From this table it follows that  $p \rightarrow q$  is *false* only when  $p$  is true and  $q$  is false; in all other cases, the implication is *true*. Hence,  $p \rightarrow q$  means exactly that “not both  $p$  and  $\neg q$ ”, or symbolically, that  $\neg(p \wedge \neg q)$ . On the other hand, it also means that “either  $\neg p$  or  $q$ ”, symbolically  $\neg p \vee q$ . The fact that  $p \rightarrow q$  is exactly the same as  $\neg(p \wedge \neg q)$  and as  $\neg p \vee q$  can be easily verified by using the truth table and showing that for all possible truth values of  $p$  and  $q$ , values of the formulae  $p \rightarrow q$ ,  $\neg(p \wedge \neg q)$  and  $\neg p \vee q$  are always the same.

However, the above, validity based definition of the implication, may not agree with our logical intuitions about the connection between the antecedent and the consequent in a true implication.

Namely, we can see that from the first and the third row of the above table, it follows that a true proposition, is implied by any proposition whatever. Thus, since the tautology, which we denote by  $\top$ , is a true proposition, the formula

$$p \rightarrow \top \quad (1)$$

is a logical truth. Bearing in mind that  $q \vee \neg q$  is a tautology, the formula:

$$p \rightarrow (q \vee \neg q) \quad (2)$$

is also a logical truth.

To see why formulae (1) and (2) may be counterintuitive, we give some non-symbolic, natural language examples of the formula (1):

*If Socrates is a man, then a triangle has three sides.*

*If Socrates is not a man, then a triangle has three sides.*

Both of them are true, because a triangle has three sides, no matter whether Socrates is a man or not. The fact that the true statement “a triangle has three sides” is implied by contradictory statements that “Socrates is a man” and “Socrates is not a man” seems somehow paradoxical. That is why the formulae (1) and (2) are referred to in the literature as “paradoxes of implication”. It should be noted here that, the term “paradox” doesn’t mean that we deal with something logically unacceptable or self-contradictory, what is usually meant by a logical paradox, rather that we deal with statements which do not match with our intuition.

From the first and the third row of the above truth table of implication, we also have that *if  $q$  is true, then  $p \rightarrow q$  is also true*, whatever the value of  $q$  is, *true* or *false*, therefore the formula

$$q \rightarrow (p \rightarrow q) \quad (3)$$

also presents a logical truth, so-called *positive paradox*.

Formulae (1), (2) and (3) are not the only “paradoxes of implication”. For example, from the third and the fourth row of the above truth table, it follows that a false proposition implies any proposition whatever. Thus, since the contradiction, which we denote by  $\perp$ , is a false proposition, the formula

$$\perp \rightarrow p \quad (4)$$

is a logical truth. This formula may also produce some “paradoxical situations”. For example, in the following two true sentences, corresponding to (4):

*If  $2 + 2 = 5$  then Socrates is a man.*

*If  $2 + 2 = 5$  then Socrates is not a man,*

the false proposition “ $2 + 2 = 5$ ” implies, in the first one, the statement that “Socrates is a man” and, in the second sentence, it implies its contradictory statement that “Socrates is not a man”. That may seem senseless.

Moreover, bearing in mind that  $q \wedge \neg q$  is a contradiction, the formula (4) can be seen as:

$$(q \wedge \neg q) \rightarrow p \quad (5)$$

which is another logical truth among the “paradoxes of implication”, known as *ex falso quodlibet*.

For example, some sentences corresponding to (5) are:

*If I am in New York and I am not in New York, then London is in Israel,*

or

*If I am in New York and I am not in New York, then London is in England,*

or

*If I am in New York and I am not in New York, then the price of wheat raises.*

Although all of them look strange, they are all logically true.

From the third and the fourth row of the above truth table of implication, we also have that, *if  $p$  is false, then  $p \rightarrow q$  is true*, whatever the value of  $q$  is, *true* or *false*, therefore

$$\neg p \rightarrow (p \rightarrow q) \quad (6)$$

presents another logical truth, so-called *negative paradox*.

Besides formulae (1)–(6), there are other “paradoxes of implication”, but we shall not give any more of them here. Our point was only to emphasize that some logical truths of classical logic, may seem counterintuitive. To understand them properly we must rely, not on our intuition about the meaning of the word “implies”, but on the definition of the notion “ $p$  implies  $q$ ”, which, in classical logic, means no more than it is not the case that  $p$  is true and  $q$  is false.

Consider now, the following *non-sequitur*:

*Lara plays a piano.*  
*Then, dogs are mammals or dogs are not mammals.*

Although this proof is valid – it corresponds to the tautology:

$$p \rightarrow (q \vee \neg q)$$

no one would take it seriously, because the conclusion seems to have nothing to do with the premise.

Therefore, in classical logic, where the implication is defined via the notion of validity, some non-sequiturs correspond to tautologies and, consequently, they are classified as good proofs. But, they are not. So, the notion of validity cannot provide a proper distinction between bad and good inferences in classical logic. Another logical system of implication, different from the classical logic, is needed, to overcome this problem.

### 3. RELEVANT LOGICS

The common denominator of all our “paradoxes of implication” is that in their implications an antecedent is irrelevant to a consequent. Ackermann was among the first logicians to think in that way and he defined in [1], a rigorous implication, such that, whenever the implication  $P \rightarrow Q$  is true, then a logical connection must hold between the propositions  $P$  and  $Q$ . By this, he set the basis for the central idea of relevant logics where  $P \rightarrow Q$  means that  $P$  is *relevant* for  $Q$ , in a sense that in a derivation of  $Q$ ,  $P$  must be really used. Adopting this principle, the above “paradoxes of implication” become invalid.

It should be noted that Ackermann was not the first logician of the modern times who considered the relevant implication. Before him, I. E. Orlov [21] set up the first axiomatization of the implication–negation fragment of the relevant logic  $R$ , [5]. But, Orlov’s work seems to be forgotten, while Ackerman’s ideas have influenced many logicians of his time, especially A. R. Anderson and N. D. Belnap, who are considered pioneers within Relevant logics.

In order to illustrate some proofs, we shall use Anderson and Belnap’s natural deduction system for the most famous relevant logic  $R$ . For example, to prove that a formula  $p \rightarrow ((p \rightarrow q) \rightarrow q)$  is a theorem of  $R$ , we proceed as follows:

- |    |   |                       |
|----|---|-----------------------|
| 1. | $p_{\{1\}}$   | hyp                   |
| 2. | $(p \rightarrow q)_{\{2\}}$                               | hyp                   |
| 3. | $q_{\{1,2\}}$   | 1, 2, $\rightarrow E$ |
| 4. | $((p \rightarrow q) \rightarrow q)_{\{1\}}$               | 2, 3, $\rightarrow I$ |
| 5. | $p \rightarrow ((p \rightarrow q) \rightarrow q)_{\{1\}}$ | 1, 4, $\rightarrow I$ |

The numbers in set brackets are called indices. They indicate the hypotheses used to prove the formula. The index  $\{1, 2\}$  in the line 3 means that we have used hypotheses  $\{1\}$  and  $\{2\}$  in order to derive  $q$ , by the application of the implication elimination rule,  $\rightarrow E$ . When we apply the implication introduction rule,  $\rightarrow I$ , we discharge a hypothesis, as in lines 4 and 5. Then the number of the hypothesis must really occur in the index of the formula that is to become the consequent of the implication.

One of the axioms of the logic  $R$  is the axiom of contraction  $(p \rightarrow (p \rightarrow q)) \rightarrow (p \rightarrow q)$ . For the natural deduction proof system for  $R$ , it means that a hypothesis is allowed to be used in a derivation more than once. We give the proof:

- |    |   |                       |
|----|---|-----------------------|
| 1. | $p \rightarrow (p \rightarrow q)_{\{1\}}$                                   | hyp                   |
| 2. | $p_{\{2\}}$   | hyp                   |
| 3. | $(p \rightarrow q)_{\{1,2\}}$   | 1, 2, $\rightarrow E$ |
| 4. | $q_{\{1,2\}}$   | 2, 3, $\rightarrow E$ |
| 5. | $(p \rightarrow q)_{\{1\}}$   | 2, 4, $\rightarrow I$ |
| 6. | $((p \rightarrow (p \rightarrow q)) \rightarrow (p \rightarrow q))_{\{1\}}$ | 1, 5, $\rightarrow I$ |

where the hypothesis  $p_{\{2\}}$  is used twice, in lines 3 and 4. In  $R$ , indices are sets, therefore, the index  $\{1, 2\}$  at the line 4 is the set union of  $\{2\}$  and  $\{1, 2\}$ .

Indices keep track of using assumptions. To ensure that each hypothesis is really used in the derivation, an index of a hypothesis must appear somewhere in the deduction, and at some later point it must disappear. This strategy disallows the inference of “paradoxes of implication”. Consider the following attempt to prove the positive paradox  $p \rightarrow (q \rightarrow p)$ :

- |    |   |                       |
|----|---|-----------------------|
| 1. | $p_{\{1\}}$                                 | hyp                   |
| 2. | $q_{\{2\}}$                                 | hyp                   |
| 3. | $p_{\{1\}}$                                 | 1, reiteration        |
| 4. | $(q \rightarrow p)_{\{1\}}$                 | 2, 3, $\rightarrow I$ |
| 5. | $(p \rightarrow (q \rightarrow p))_{\{1\}}$ | 1, 4, $\rightarrow I$ |

The step 3 is legitimate, but 4 is not, because  $p$  is not derived from  $q$ , really, 2 does not belong to  $\{1\}$ . Therefore the hypothesis  $q_{\{2\}}$  cannot be discharged and this does not present the proof of the positive paradox.

In addition to “paradoxes of implication”, relevant logics also disallow inferences of all implications whose antecedent is irrelevant for the consequent. For example, they reject the statement “If  $2 + 2 = 4$  then a triangle has three sides”, because its antecedent is irrelevant for the consequent, although both, the antecedent and the consequent are true. So, the truth value of  $p \rightarrow q$  in relevant logics, does not depend only on the truth values of  $p$  and  $q$ . Consequently, relevant logics are not truth-functional. This means that we cannot use truth tables in order to determine a truth value of a statement. To do that, we need semantics for relevant implication.

For a long time relevant logics were without semantics. The first semantics defined was the semilattice semantics, defined by Urquhart in 1972 [24]. Soon, other semantics appeared. Without intention to discuss all of them here, we shall mention only the ternary relation semantics defined by Routley and Meyer [23], where truth of formulae is defined at worlds, with the help of a three-place accessibility relation  $R$  on worlds. The truth condition for  $\rightarrow$  is:

$P \rightarrow Q$  is true at a world  $a$  iff for all worlds  $b$  and  $c$  such that  $Rabc$   
either  $P$  is false at  $b$  or  $Q$  is true at  $c$ .

However, in order to give relevant implication a real meaning on this semantics, the ternary accessibility relation needs an interpretation. There are several interpretation defined, but we shall mention only one, based on theories of information, suggested by Barwise [3] and developed by Restall [22]. On this view, worlds can be seen as sites and channels. A site is a context in which information is received and a channel is a conduit through which information is transferred.  $Rabc$  means that  $a$  is a channel between sites  $b$  and  $c$  and  $P \rightarrow Q$  is true at  $a$  iff whenever  $a$  connects a site  $b$  at which  $P$  obtains to a site  $c$ ,  $Q$  obtains at  $c$ .

It is clear that the semantics of relevant logics is not as easy as the semantics of the classical logic – while in classical logic we deal only with the form, in relevant logics, we must deal with the content of statements, also.

#### 4. CONCLUSION

Relevant logics present the most suitable logical systems of implication. Today, they are widely used in philosophy and computer science. Linear logic, the well-known logic of computational resources, initiated by Jean-Yves Girard [9], is a relevant logic with exponentials ( $!$  and  $?$ ), which reject contraction and the distribution of conjunction over disjunction.

As for the proof theory, in addition to the natural deduction systems, mentioned above, there are other proving systems for relevant logics. For example, sequent calculi based on the work of Dunn [6] and Minc [20], as well as “Display Logic” developed by Belnap [4].

The most famous relevant logic  $R$  is not decidable [25], but some of its fragments and extensions are. Sometimes the decidability problem of a logic may be solved by finding the appropriate sequent calculus and proving some of its important properties. As for the relevant logics, some solutions of such problems, suggested by Ilić, are presented in [10]–[17].

#### Acknowledgement

This work is partially supported by the Ministry of Science and Technology of Serbia, grant number ON174026.

## REFERENCES

- [1] Ackermann, W. (1956). Begründung einer strengen Implikation (A foundation for a Rigorous implication), *The Journal of Symbolic Logic*, 21, pp. 113–128.
- [2] Anderson, A. R., Belnap N. D. Jr. (1975). *Entailment: the logic of relevance and necessity*, vol. 1, Princeton University Press, Princeton, New Jersey
- [3] Barwise, J. (1993). Constraints, channels and the flow of Information, in P. Aczel, et al. (eds.), *Situation Theory and Its Applications (Volume 3)*, Stanford: CSLI Publications, pp. 3–27.
- [4] Belnap Jr., N.D., (1982). Display logic, *Journal of Philosophical Logic*, 11, pp. 375–417.
- [5] Došen, K., (1992). The first axiomatization of Relevant logic, *Journal of Philosophical Logic* 21, pp. 339–356.
- [6] Dunn, J.M., (1973). A 'Gentzen system' for positive relevant implication, *The Journal of Symbolic Logic*, 38, pp. 356–357.
- [7] Dunn, J. M., (1986). Relevance logic and entailment, D. Gabbay and F. Guenther (eds.) *Handbook of Philosophical Logic*, vol. 3, D. Reidel Publishing Company, pp. 117–224,
- [8] Dunn, J. M., Restall, G., (2002). Relevance logic, D. Gabbay and F. Guenther (eds.) *Handbook of Philosophical Logic*, vol. 6, Kluwer Academic Publishers, pp. 1–128,
- [9] Girard, J–Y., (1987). Linear logic, *Theoretical Computer Science*, 50, pp. 1–102.
- [10] Ilić, M. (2008). Cut elimination and decidability for classical Lambek logic, *Journal of Logic and Computation*, 18, pp. 171–199.
- [11] Ilić, M., Boričić, B. (2014). A cut-free sequent calculus for relevant logic  $RW$ , *Logic Journal of the IGPL*, 22(4), pp. 673–695.
- [12] Ilić, M., (2016). The proof—theoretical analysis of contraction—less relevant logic, *Bulletin of the international mathematical virtual institute*, 6, pp. 1–11.
- [13] Ilić, M., (2016). An alternative Gentzenisation of  $RW_+^o$ , *Mathematical Logic Quarterly*, 6, pp. 465–480.
- [14] Ilić, M., (2017) A natural deduction and its corresponding sequent calculus for positive contraction-less relevant logics, *Reports on Mathematical logic*, 52, pp. 93–124.
- [15] Ilić, M., (2021). A cut–elimination proof in positive relevant logic with necessity, *Studia Logica* 109(3), pp. 607–638.
- [16] Ilić, M., Boričić, B. (2021). A note on the system  $GRW$  with the intensional contraction rule, *Logic Journal of the IGPL*, 29(3), pp. 333–339.
- [17] Ilić, M., (2021). A note on an alternative Gentzenization of  $RW_+^o$ , *Mathematical Logic Quarterly*, 67(2), pp. 186–192.
- [18] Mansur, M. N., (2005). *The paradoxes of material implication*, Master of Arts thesis, Memorial University of Newfoundland
- [19] Mares, E. (2004). *Relevant logics, A philosophical interpretation*, Cambridge University Press, Cambridge
- [20] Minc, G., (1976). Cut elimination theorem for relevant logics, *Journal of Soviet Mathematics*, 6, pp. 422–428.
- [21] Orlov, I. E., (1928). The calculus of compatibility of propositions (in Russian), *Matematicheskii Sbornik* 35, pp. 263–286
- [22] Restall, G. (1996). Information flow and relevant logics, in J. Seligman and D. Westerstahl (eds.), *Logic, Language and Computation (Volume 1)*, Stanford: CSLI Publications, pp. 463–478.
- [23] Routley, R., Meyer, R. (1973). The semantics of entailment, in H. Leblanc (ed.), *Truth, Syntax and Modality*, Proceedings of the Temple University conference on alternative semantics, North–Holland, Amsterdam, pp. 199–243
- [24] Urquhart, A. (1972). Semantics for relevant logics, *The Journal of Symbolic Logic*, 37, pp. 159–169.
- [25] Urquhart, A. (1984). The undecidability of entailment and relevant implication, *The Journal of Symbolic Logic*, 49(4) pp. 1059–1073.





## DIRP MODEL PREDIKCIJE SKUPA BROJEVA

### THE DIRP MODEL FOR PREDICTING A SET OF NUMBER

SINIŠA VLAJIĆ<sup>1</sup>, DUŠAN SAVIĆ<sup>1</sup>, ILIJA ANTOVIĆ<sup>1</sup>, MILOŠ MILIĆ<sup>1</sup>, VOJISLAV STANOJEVIĆ<sup>1</sup>

<sup>1</sup> Fakultet organizacionih nauka, Katedra za softversko inženjerstvo, Beograd, sinisa.vlajic@fon.bg.ac.rs, dusan.savic@fon.bg.ac.rs, ilija.antovic@fon.bg.ac.rs, milos.milic@fon.bg.ac.rs, vojislav.stanojevic@fon.bg.ac.rs

**Rezime:** U ovom radu je opisan sopstveni DIRP model predikcije skupa brojeva. DIRP model zasnovan je na polinomskoj interpolaciji i ekstrapolaciji skupova brojeva  $(X, Y)$ , nezavisne promenljive  $X$ , koja obrazuje ekvidistantan skup i zavisne promenljive  $Y$ , koja je uređen niz (aritmetički ili geometrijski niz), ili se se diferenciranjem svodi na neki od navedenih uređenih nizova. Detaljno je objašnjen proces generisanja formule (standardnog ili eksponencijalnog polinoma) na osnovu ulaznog skupa brojeva. Razvijena je Java aplikacija koja podržava DIRP model predikcije i dat je primer postupka predikcije skupa brojeva korišćenjem aplikacije. Na kraju su testirani različiti primeri (skupovi brojeva), kategorizovani po tipovima nizova, pomoću DIRP modela. Generisane formule (standardni i eksponencijalni polinomi) za testirane primere precizno interpoliraju i ekstrapoliraju definisani skup ulaznih brojeva.

**Ključne reči:** Deterministički modeli predikcije, DIRP model, Generisanje matematičkih formula

**Abstract:** In this paper, our own DIRP model for predicting a set of numbers is presented. The DIRP model is based on polynomial interpolation and extrapolation of sets of numbers  $(X, Y)$ , the independent variable  $X$ , which forms an equidistant set, and the dependent variable  $Y$ , which is an ordered array (arithmetic or geometric sequence), or is reduced to one of the above ordered arrays by differentiating. The process of generating a formula (standard or exponential polynomial) based on the input set of numbers is explained in detail. A Java application that supports the DIRP prediction model has been developed and an example of a process for predicting a set of numbers using an application is given. Finally, different examples (sets of numbers) were tested, categorized by sequence types, using the DIRP model. The generated formulas (standard and exponential polynomials) for the tested examples precisely interpolate and extrapolate the defined set of input numbers.

**Keywords:** Deterministic prediction models, DIRP model, Mathematical formula generator

## 1. UVOD

Kada se razvija sistem za predikciju skupa brojeva potrebno je napraviti odgovarajući model predikcije. Model predikcije se trenira sa skupom trening brojeva ili formula, kako bi mogao da se koristi u predviđanju budućih skupova brojeva. Modeli predikcije mogu biti deterministički i stohastički [4]. Deterministički modeli su zasnovani na metodama polinomske interpolacije (Lagranžova metoda, Njutnova metoda, Nevilova metoda, Spline metoda, ...) [1,2,7], interpolacije najbližih suseda (Nearest-Neighbor Interpolation), metodama zasnovanim na Furijeovoj teoriji, ..., itd. Stohastički modeli su zasnovani na regresionim i autoregresionim metodama (Moving Average, ARMA, ARIMA,...) [3], metodama mašinskog učenja [5], metodama zasnovanim na dinamičkim podacima, ..., itd.

U ovom radu je kreiran sopstveni DIRP model predikcije skupa brojeva. DIRP naziv je skraćenica od ključnih koncepata koji definišu navedeni model: Diferenciranje, Integraljenje, Rekurzija i Predikcija. DIRP model zasnovan je na polinomskoj interpolaciji i ekstrapolaciji skupova brojeva  $(X, Y)$ , nezavisne promenljive  $X$ , koja obrazuje ekvidistantan skup i zavisne promenljive  $Y$ , koja je uređen niz (aritmetički ili geometrijski niz), ili se se diferenciranjem svodi na neki od navedenih uređenih nizova. DIRP model predikcije predstavlja proširenje i nadogradnju UN modela predviđanja [6].

DIRP model predikcije se trenira sa opštim oblikom linearne funkcije, na osnovu čega se generiše standardni polinom za skup brojeva  $(X, Y)$ , kod kojih zavisna promenljiva  $Y$  obrazuje aritmetički niz ili se diferenciranjem svodi na aritmetički niz. Generisani polinom je isti kao polinomi koji se generišu na

osnovu Lagranžove, Nutnove ili Nevilove metode<sup>1</sup>, za navedeni skup podataka pod navedenim ograničenjima ( $X$  obrazuje ekvidistanatan skup i  $Y$  je aritmetički niz ili se diferenciranjem svodi na aritmetički niz). Generisani polinom precizno interpolira i ekstrapolira definisani skup podataka ( $X, Y$ ).

DIRP model predikcije se trenira i sa opštim oblikom eksponencijalne funkcije, na osnovu čega se generiše **eksponencijalni polinom** za skup podataka ( $X, Y$ ), kod kojih zavisna promenljiva  $Y$  obrazuje geometrijski niz ili se diferenciranjem svodi na geometrijski niz. Generisani polinom interpolira i ekstrapolira definisani skup brojeva ( $X, Y$ ). Preciznost ekstrapolacije i interpolacije zavisi od broja iteracija tokom procesa generisanja eksponencijalnog polinoma. Sa povećanjem broja iteracija povećava se preciznost interolacije i ekstrapolacije. **Generisani eksponencijalni polinom predstavlja jedan od ključnih doprinosa DIRP modela predikcije u odnosu na postojeće determinističke modele predikcije koji su zasnovani na metodama polinomske interpolacije.**

## 2. DIRP MODEL PREDIKCIJE SKUPA BROJEVA

DIRP model predikcije se može predstaviti preko Slike 1. DIRP model predikcije skupa brojeva sastoji se iz dve komponente, generatora formule i prediktora sekvence brojeva. Generator formule, na osnovu ulaznog skupa brojeva, generiše formulu pomoću funkcije *procesGenerisanjaFormule*( $X, Y_m$ ). Prediktor sekvence brojeva radi interpolaciju i ekstrapolaciju ulaznog skupa brojeva, na osnovu generisane formule. Za nas je posebno značajna ekstrapolacija ulaznog skupa brojeva, jer ona u suštini predstavlja predikciju skupa brojeva.

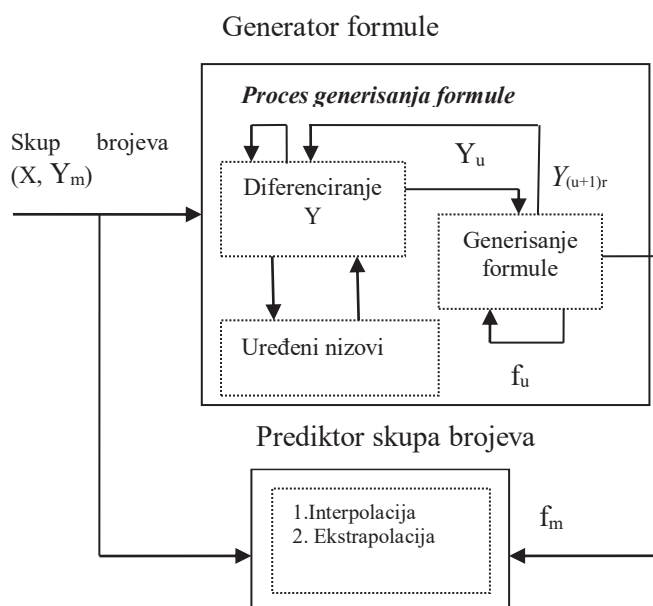
Generisanje formule je proces koji započinje ispitivanjem ulaznog skupa brojeva ( $X, Y_m$ ), odnosno zavisne promenljive  $Y_m$ , da li je ona uređen niz (aritmetički ili geometrijski). Ukoliko je  $Y_m$  uređen niz, generiše se formula  $f_m$ . Ukoliko  $Y$  nije uređen niz, radi se diferenciranje od  $Y_m$  kojim se generišu nove promenljive ( $Y_{m-1}, Y_{m-2}, \dots, Y_2, Y_1$ ), dok se ne dođe do promenljive ( $Y_u, u \in \{1, \dots, m-1\}$ ) koja je uređeni niz. Navedene promenljive obrazuju hijerarhijsku strukturu, gde je na vrhu hijerarhije  $Y_m$  a na dnu hijerarhije je promenljiva  $Y_u$  koja je uređen niz. Za promenljivu  $Y_u$  se generiše formula  $f_u$ . Ukoliko  $Y_u$  ima nadređenu promenljivu  $Y_{u+1}$  određuje se promenljiva  $Y_{(u+1)r}$  na osnovu sledeće formule:

$$Y_{(u+1)r} = Y_{(u+1)} - (Y_{(u+1)0} + \int f_u dx) \quad (1)$$

Zatim se ulazi u rekurziju generisanjem formule,  $f_{(u+1)r}$  za skup brojeva ( $X, Y_{(u+1)r}$ ). Nakon toga se određuje formula  $f_{(u+1)}$  za  $Y_{(u+1)}$ :

$$f_{(u+1)} = \int f_u dx + f_{(u+1)r} \quad (2)$$

Kao rezultat navedenog procesa generisaće se formula  $f_m$  za ulazni skup brojeva ( $X, Y_m$ ). Na osnovu formule  $f_m$  radi se interpolacija i ekstrapolacija ulaznog skupa brojeva ( $X, Y_m$ ).



Slika 1: DIRP model predikcije skupa brojeva

<sup>1</sup> Generisanje polinoma za navedene metode je urađeno korišćenjem online kalkulatora na adresi: <https://www.dcode.fr/>

### 3. PROCES GENERISANJA FORMULE KOD DIRP MODELA PREDIKCIJE

U ovom poglavlju biće objašnjen matematički model skupa brojeva  $(X, Y)$  za koji se generiše: a) standardni polinom, kada zavisna promenljiva  $Y$  obrazuje aritmetički niz ili se diferenciranjem svodi na aritmetički niz i b) eksponencijalni polinom, kada zavisna promenljiva  $Y$  obrazuje geometrijski niz ili se diferenciranjem svodi na geometrijski niz. U drugom delu poglavlja biće objašnjena funkcija *proces Generisanja Formule* kod DIRP modela predikcije.

#### 3.1 OPIS MATEMATIČKOG MODELA

**Definicija 1:** Za dati skup  $n + 1$  brojeva  $(X, Y)$ , odnosno  $(x_i, y_i)$  za koji je  $x_i \neq x_j$  i  $i \neq j$  (ne postoje dva ista  $x_i$ ), gde  $i, j \in \{0, 1, \dots, n\}$ , polinom  $P_n$  stepena manjeg ili jednakog  $n$ :

a)  $P_n(x_i) = a_n x_i^n + a_{n-1} x_i^{n-1} + \dots + a_1 x_i + a_0$  (standardni polinom) *interpolira* skup brojeva ako je  $P_n(x_i) = y_i$  i *ekstrapolira* skup brojeva ako je  $P_n(x_k) = y_k, k > n$  ili

b)  $P_n(x_i, b^{x_i}) = a_n x_i^n + a_{n-1} x_i^{n-1} + \dots + a_1 x_i + a_0 + c b^{x_i}$  (eksponencijalni polinom) *interpolira* skup brojeva ako je  $P_n(x_i, b^{x_i}) = y_i$  i *ekstrapolira* skup brojeva ako je  $P_n(x_i, b^{x_i}) = y_k, k > n$ .

Vrednosti nezavisne promenljive  $X$  obrazuju ekvidistantan skup:  $x_0, x_1, \dots, x_n$ , koji je reprezentovan preko formule:

$$x_{k-1} = x_0 + b(k-1) \quad (3)$$

gde  $k \in \{1, \dots, n+1\}$ , pri čemu je  $b$  je razlika između sukcesivnih vrednosti promenljive  $X$  ( $b = x_k - x_{k-1}$ ). Indeks  $k$  se dobija iz formule 3:

$$k = (x_{k-1} - x_0)/b + 1 \quad (4)$$

Zavisna promenljiva  $Y$  može biti reprezentovana preko sledećih promenljivih  $Y_1, Y_2, \dots, Y_m$  (Tabela 1).

**Tabela 1:** Skup brojeva  $(X, Y)$

K	X	Y <sub>1</sub>	Y <sub>2</sub>	Y <sub>3</sub>	...	Y <sub>m</sub>
1	x <sub>0</sub>	y <sub>01</sub>	y <sub>02</sub>	y <sub>03</sub>	...	y <sub>0m</sub>
2	x <sub>1</sub>	y <sub>11</sub>	y <sub>12</sub>	y <sub>13</sub>	...	y <sub>1m</sub>
3	x <sub>2</sub>	y <sub>21</sub>	y <sub>22</sub>	y <sub>23</sub>	...	y <sub>2m</sub>
4	x <sub>3</sub>	y <sub>31</sub>	y <sub>32</sub>	y <sub>33</sub>	...	y <sub>3m</sub>
5	x <sub>4</sub>	y <sub>41</sub>	y <sub>42</sub>	y <sub>43</sub>	...	y <sub>4m</sub>
...	...	...	...	...	...	...
n+1	x <sub>n</sub>	y <sub>n1</sub>	y <sub>n2</sub>	y <sub>n3</sub>	...	y <sub>nm</sub>

$Y_1$  može biti *aritmetička sekvenca*, koja je definisana preko formule:

$$f(k) = y_{(k-1)1} = y_{01} + b_1(k-1) \quad (5)$$

gde je  $b_1$  razlika između sukcesivnih vrednosti promenljive  $Y_1$  ( $b_1 = y_{k1} - y_{(k-1)1}$ ).

Ukoliko se zameni  $k$  u formuli 5, nezavisna promenljive  $Y_1$  može biti reprezentovana na sledeći način:

$$y_{(k-1)1} = y_{01} + b_1/b(x_{k-1} - x_0) \quad (6)$$

Promenljive  $Y_2, Y_3, \dots, Y_m$  su reprezentovane preko sledeće formule:

$$y_{(k-1)j} = y_{0j} + \sum_{i=0}^{k-2} y_{i(j-1)} \quad (7)$$

gde  $j \in \{2, \dots, m\}$

$Y_1$  može biti *geometrijska sekvenca*, koja je definisana preko formule:

$$f(k) = y_{(k-1)1} = y_{01} * b_2^{(k-1)} \quad (8)$$

gde je  $b_2$  količnik između sukcesivnih vrednosti promenljive  $Y_1$  ( $b_2 = y_{k1} / y_{(k-1)1}$ ). Promenljiva  $Y_1$  može biti reprezentovana na sledeći način:

$$y_{(k-1)1} = y_{01} * b_2^{(x_{k-1} - x_0)/b} \quad (9)$$

$Y_1$  mogu biti i druge sekvence (harmonijski niz, Fibonačijev niz, ...), koje u ovome radu nećemo razmatrati.

**Definicija 2:** Nezavisna promenljiva  $X$  je klasa koja ima atribut  $x$ , koji predstavlja niz realnih brojeva.

**Definicija 3:** Zavisna promenljiva  $Y$  je klasa koja ima sledeće atribute: *name*, *arranged*, *formula*,  $y$ ,  $d$ , *child* u *parent*. Atribut  $y$  je niz realnih brojeva. Atributi  $d$  (skraćenica od difference), *child* i *parent* pripadaju klasi  $Y$ . Atributi *name* i *formula* su tekstualnog tipa. Atribut *arranged* je logičkog tipa.

**Definicija 4:** Promenljive  $Y_1, Y_2, \dots, Y_m$  pripadaju klasi  $Y$ .

### 3.2 PROCES GENERISANJA FORMULE

Za skup podataka  $(X, Y_a)$ , gde je  $a \in \{0, 1, \dots, m\}$  potrebno je generisati formulu koja će da interpolira i ekstrapolira navedeni skup brojeva. Naziv nezavisne promenljive  $Y_a$  je  $Y_0$  ( $Y_{a.name} = "Y_0"$ ). Generisanje formule se radi pomoću funkcije **procesGenerisanjaFormule** ( $X, Y_a$ ). U daljem tekstu data je funkcija **procesGenerisanjaFormule** koja predstavlja glavnu funkciju DIRP modela predikcije.

#### Funkcija **procesGenerisanjaFormule** ( $X, Y_a$ )

**Korak 1:** Ispituje se da li je niz  $Y_{a.y}$  uređen niz (aritmetički ili geometrijski niz ili je broj članova niza jednak 2). Ukoliko je niz uređen atribut *arranged* promenljive  $Y_a$  dobija logičku vrednost true ( $Y_{a.arranged} = true$ ). Ukoliko niz nije uređen atribut *arranged* promenljive  $Y_a$  dobija logičku vrednost false ( $Y_{a.arranged} = false$ ).

**Korak 1.1:** Ukoliko je niz uređen ( $Y_{a.arranged} == true$ ) generiše se formula  $Y_{a.formula}$  za  $Y_{a.y}$ .

**Korak 1.1.1:** Ispituje se da li  $Y_a$  ima roditelja ( $Y_{a.parent}$ ).

**Korak 1.1.1.1:** Ukoliko  $Y_a$  nema roditelja prekida se izvršenje funkcije *formulaGeneration*.

**Korak 1.1.1.2:** Promenljiva  $Y_a$  postaje roditelj od  $Y_a$  ( $Y_a = Y_{a.parent}$ ). Za niz  $Y_{a.y}$  se generiše formula  $Y_{a.formula}$ . Kreira se promenljiva  $Y_{(Y_{a.name} + "_2")}$ . Naziv promenljive se dobija kada se nazivu promenljive  $Y_a$  doda nastavak "\_2" ( $Y_{(Y_{a.name} + "_2").name} = Y_{a.name} + "_2"$ ).

$$Y_{a.d} = Y_{(Y_{a.name} + "_2")}$$

$$Y_{a.y_i} = Y_{a.y_0} + \int Y_{a.child.formula} + Y_{a.d.y_i} \Rightarrow Y_{a.d.y_i} = Y_{a.y_i} - Y_{a.y_0} - \int Y_{a.child.formula}$$

Poziva se funkcija **procesGenerisanjaFormule** ( $X, Y_{a.d}$ ).

$$Y_{a.formula} = Y_{a.y_0} + \int Y_{a.child.formula} + Y_{a.d.formula}$$

$$Y_a = Y_{a.parent}$$

Povratak na korak 1.1.1.

**Korak 1.2:** Ukoliko niz nije uređen ( $Y_{a.y} == false$ ), kreira se nova promenljiva klase  $Y$ . Naziv promenljive se dobija kada se nazivu promenljive  $Y_a$  doda nastavak "\_1" ( $Y_{(Y_{a.name} + "_1").name} = Y_{a.name} + "_1"$ ). Radi se diferenciranje  $Y_{a.y}$ :

$$Y_{a.y_i}' = (Y_{a.y_{i+1}} - Y_{a.y_i})/b, i \in \{0, 1, 2, \dots, m-2\}$$

Kao rezultat diferenciranja dobija se novi niz  $Y_{(Y_{a.name} + "_1").y}$ :  $Y_{(Y_{a.name} + "_1").y} = \{y_1 - y_0, y_2 - y_1, \dots, y_{m-1} - y_{m-2}\}$

Atribut *child* promenljive  $Y_a$  postaje promenljiva  $Y_{(Y_{a.name} + "_1")}$ :  $Y_{a.child} = Y_{(Y_{a.name} + "_1")}$

Atribut *parent* promenljive  $Y_{(Y_{a.name} + "_1")}$  postaje promenljiva  $Y_a$ :  $Y_{(Y_{a.name} + "_1").parent} = Y_a$

Promenljiva  $Y_a$  postaje promenljiva  $Y_{(Y_{a.name} + "_1")}$ :  $Y_a = Y_{(Y_{a.name} + "_1")}$

Povratak na korak 1.

#### Kraj funkcije **procesGenerisanjaFormule**

### 4. PRIMER POSTUPKA PREDIKCIJE SKUPA BROJEVA NA OSNOVU DIRP MODELA

U Java programskom jeziku napravljena je desktop aplikacija koja podržava predikciju skupa brojeva na osnovu DIRP modela. Aplikacija generiše kompletan postupak (što se može videti u niže navedenom primeru) predikcije skupa brojeva na osnovu početnog skupa brojeva i ona obuhvata: a) generisanje formule na osnovu ulaznog skupa brojeva i b) interpolaciju i ekstrapolaciju ulaznog skupa brojeva.

*Izaberi tip niza: 1-AR 2-GM: 1 Unesi sifru niza: niz9251 Odredi broj parova ulaznog skupa brojeva: 4*

*Ulazni skup vrednosti: (X, Y0) = (1,3), (2,7), (3,12), (4,18)*

*Poziva se funkcija procesGenerisanjaFormule(X, Y0).*

*Korak 1: Ispituje se da li je niz (Y0.y) uređen (aritmetički): y[1]-y[0]=4.0, y[2]-y[1]=5.0, Niz nije uređen (aritmetički): 3 7 12 18*

*Korak 1.2: Kreira se promenljiva Y0\_1 na osnovu promenljive Y0. Radi se diferenciranje Y0.y.*

*Kao rezultat diferenciranja dobija se novi niz Y0\_1.y: 4 5 6, Y0.child=Y0\_1; Y0\_1.parent=Y0; Y0=Y0\_1*

*----Korak 1: Ispituje se da li je niz (Y0\_1.y) uređen (aritmetički): y[1]-y[0]=1.0, y[2]-y[1]=1.0, Niz je uređen (aritmetički): 4 5 6*

*----Korak 1.1: Generise se formula (Y0\_1.formula) za Y0\_1.y: Y0\_1.formula = +3+1\*x*

*----Korak 1.1.1: Ispituje se da li Y0\_1 ima roditelja. Korak 1.1.1.2: Promenljiva Y0\_1 postaje roditelj od Y0\_1(Y0\_1=Y0\_1.parent)*

*Za niz Y0.y generise se formula (Y0.formula).*

*Kreira se promenljiva Y0\_2, Y0.d=Y0\_2.*

*Y0.y(i) = Y0.y(0) + \int (+3+1\*x)dx + Y0.d.y(i), Y0.d.y(i) = Y0.y(i) - (Y0.y(0) + \int (+3+1\*x)dx), Y0.d.y(i) = Y0.y(i) - (+3+3\*x+0.5\*x^2)*

*Y0.d.y(0) = 3-(3)-(3)-(0.5) = -3.5, Y0.d.y(1) = 7-(3)-(6)-(2) = -4, Y0.d.y(2) = 12-(3)-(9)-(4.5) = -4.5, Y0.d.y(3) = 18-(3)-(12)-(8) = -5*

*Y0\_2: -3.5 -4 -4.5 -5*

*----Poziva se funkcija procesGenerisanjaFormule(X, Y0\_2).*

*----Korak 1: Ispituje se da li je niz (Y0\_2.y) uređen (aritmetički): y[1]-y[0]=-0.5, y[2]-y[1]=-0.5, y[3]-y[2]=-0.5, Niz je uređen (aritmetički): -3.5 -4 -4.5 -5*

*----Korak 1.1: Generise se formula (Y0\_2.formula) za Y0\_2.y: Y0\_2.formula = -3-0.5\*x*

*----Korak 1.1.1: Ispituje se da li Y0\_2 ima roditelja. Korak 1.1.1.1: Promenljiva Y0\_2 nema roditelja.*

*----Prekida se izvršenje funkcije procesGenerisanjaFormule(X, Y0\_2).*

$Y_0.formula = Y_0.y(0) + \int(Y_0\_1.formula)dx + Y_0\_2.formula$ ,  $Y_0.formula = +3+3*x-3-0.5*x+0.5*x^2$ ,  $Y_0.formula = +0.5*x^2+2.5*x$   
Prekida se izvršenje funkcije procesGenerisanjaFormule(X,Y0).

Sifra niza: niz9251 Broj generisanih formula: 3 Tip niza: AritmetičkiNiz

**INTERPOLACIJA:** Ulazni skup vrednosti:  $(X,Y_0) = (1,3),(2,7),(3,12),(4,18)$ , Generisan skup vrednosti  $(Y_g)$  na osnovu formule:  $(X,Y_g) = (1,3),(2,7),(3,12),(4,18)$  **Razlika između generisanih i početnih vrednosti  $(Y_g-Y_0) = (0,0),(0,0),(0,0),(0,0)$**

**EKSTRAPOLACIJA:** Očekivani skup vrednosti:  $(X,Y_0) = (5,25),(6,33),(7,42)$ , Generisan skup vrednosti  $(Y_g)$  na osnovu formule:  $(X,Y_g) = (5,25),(6,33),(7,42)$  **Razlika između generisanih i očekivanih vrednosti  $(Y_g-Y_0) = (0,0),(0,0),(0,0)$**

## 5. TESTIRANJE RAZLIČITIH SKUPOVA BROJEVA POMOĆU DIRP MODELA

Izvršeno je testiranje DIRP modela predviđenja, tako što su napravljene primeri koji su kategorizovani po tipovima nizova. Primeri u Tabeli 2 (od Niz1 do Niz8), testiraju skupove brojeva, kod kojih je zavisna promenljiva  $Y_0$  zasnovana na aritmetičkom nizu. Nezavisna promenljiva  $X$  sadrži sledeći skup vrednosti:  $(1,2,3,4,5,...)$ . Generisana formula (*standardni polinom*) precizno interpolira i ekstrapolira definisani skup ulaznih brojeva  $(X, Y_0)$ .

**Tabela 2:** Testiranje skupova brojeva koji su zasnovani na aritmetičkom nizu pomoću DIRP modela

Sifra niza	Zavisna promenljiva $Y_0$ (ulaz. vrednosti)	Zavisna promenljiva $Y_0$ (oček. vrednosti)	Generisana formula	Generisan skup brojeva $Y_g$ na osnovu formule
Niz1	1 2 3 4 5	6 7 8 9 10	$f(x) = +1*x$	1 2 3 4 5 6 7 8 9 10
Niz2	3 8 13 18 23	28 33 38 43 48	$f(x) = -2+5*x$	3 8 13 18 23 28 33 38 43 48
Niz3	4 7 10 13 16	19 22 25 28 31	$f(x) = +1+3*x$	4 7 10 13 16 19 22 25 28 31
Niz4	3 9 16 25	37 53 74 101 135 177	$f(x) = +0.1667*x^3-0.5*x^2 + 6.3333*x-3.0$	3 9 16 25 37 53 74 101 135 177
Niz5	13 24 48 98 192 353	609 993 1543	$f(x) = +0.2083*x^4+0.0833*x^3+0.7917*x^2+4.9167*x+7.0$	13 24 48 98 192 353 609 993 1543
Niz6	5 7 11 17 25	35 47 61 77 95 115	$f(x) = +1*x^2-1*x+5$	5 7 11 17 25 35 47 61 77 95 115
Niz7	3 7 12 18 25 33	42 52 63 75	$f(x) = +0.5*x^2+2.5*x$	3 7 12 18 25 33 42 52 63 75
Niz8	7 11 19 32 51 77	111 154 207 271	$f(x) = +0.1667*x^3+1*x^2-0.1667*x+6.0$	7 11 19 32 51 77 111 154 207 271

Primeri u tabeli Tabeli 3 (od Niz11 do Niz18), testiraju skupove brojeva, kod kojih je zavisna promenljiva  $Y_0$  zasnovana na geometrijskom nizu. Nezavisna promenljiva  $X$  sadrži sledeći skup vrednosti:  $(1,2,3,4,5,...)$ . Generisana formula (*eksponencijalni polinom*) precizno interpolira i ekstrapolira definisani skup ulaznih brojeva  $(X, Y_0)$ .

**Tabela 3:** Testiranje skupova brojeva koji su zasnovani na geometrijskom nizu pomoću DIRP modela

Sifra niza	Zavisna promenljiva $Y_0$ (ulaz. vrednosti)	Zavisna promenljiva $Y_0$ (oček. vrednosti)	Generisana formula	Generisan skup brojeva $Y_g$ na osnovu formule
Niz11	2 4 8 16 32	64 128 256	$f(x) = +1*(2)^x$	2 4 8 16 32 64 128 256
Niz12	2 6 18 54	162 486 1458	$f(x) = +0.6667*(3)^x$	2 6 18 54 162 486 1458
Niz13	1 4 16 64	256 1024	$f(x) = +0.25*(4)^x$	1 4 16 64 256 1024
Niz14	5 8 13 22 39	72 137 266 523	$f(x) = +1.0*(2)^x+1.0*x+2.0$	5 8 13 22 39 72 137 266 523
Niz15	1 3 7 15 31	63 127	$f(x) = +1.0*(2)^x-1.0$	1 3 7 15 31 63 127
Niz16	7 11 17 27 45	79 145 275 533	$f(x) = +1.0*(2)^x+2.0*x+3.0$	7 11 17 27 45 79 145 275 533
Niz17	12 20 33 55 94 166	303 569 1092	$f(x) = +0.5*x^2+ 2.0*(2)^x+ 2.5*x+ 5.0$	12 20 33 55 94 166 303 569 1092
Niz18	7 9 15 33	87 249 735 2193	$f(x) = +0.333333*(3)^x+6.0$	7 9 15 33 87 249 735.000011 2193.000033

Koeficijenti u formulama (od Niz1 do Niz17), kao i generisane vrednosti iz tih formula zaokruženi su na 4 decimale. Koeficijenti u formuli Niz18, kao i generisane vrednosti iz te formule zaokruženi su na 6 decimale.

## 6. ANALIZA DIRP MODELA I PRAVCI DALJEG ISTRAŽIVANJA

U ovom trenutku DIRP model predikcije daje dobre rezultate kod interpolacije i ekstrapolacije nizova brojeva koji se mogu svesti na aritmetički ili geometrijski niz. DIRP model omogućava interpolaciju nizova brojeva, slično kao i druge numeričke metode (Lagranžova metoda, Njutnova metoda, ...) za bilo koji ulazni niz brojeva. Međutim ukoliko se ulazni niz ne može svesti na aritmetički ili geometrijski niz *ekstrapolacija niza brojeva neće biti dobra*. Na primer ukoliko bi ulazni niz imao vrednosti  $(0.5\ 0.2\ 0.1\ 0.058823529\ 0.038461538\ 0.027027027)$  – ulazne vrednosti i  $(0.02\ 0.015384615\ 0.012195122\ 0.00990099)$  – očekivane vrednosti) koje su generisane iz funkcije:  $f(x) = 1/(1+x^2)$ , formula koja se dobija iz DIRP modela (koji je zasnovan na aritmetičkom nizu) ima sledeći oblik:  $f(x) = -6.0E-4*x^5.0+0.0139*x^4.0-0.1211*x^3.0+0.5361*x^2.0-1.2497*x+1.3214$ . Navedena formula generiše sledeće vrednosti  $(0.5\ 0.2\ 0.1\ 0.0588\ 0.0385\ 0.027)$  (interpolacija)  $(-0.0383\ -0.3483\ -1.2987\ -3.5674)$

(ekstrapolacija)), iz čega se može videti da je interpolacija ulaznog niza dobra, dok je ekstrapolacija loša. Formula koja se dobija iz DIRP modela (koji je zasnovan na geometrijskom nizu) ima sledeći oblik:  $f(x) = -5.0E-4 * x^3 + 0.0106 * x^2 + 0.0072 * (0.2533)^x + 1.3024 * (0.2532)^x - 0.0806 * x + 0.2389$ . Navedena formula generiše sledeće vrednosti (0.5 0.2 0.1 0.0588 0.0385 0.027 (interpolacija) 0.0191 0.0112 1.0E-4 , -0.0173 (ekstrapolacija)), koje predstavljaju dobru interpolaciju ulaznog niza i ekstrapolaciju koja je malo bolja od predhodne ekstrapolacije, ali je i dalje neprecizna u odnosu na očekivane vrednosti.

Trenutno naš sistem prepoznaje nizove koji se mogu svesti na funkcije: a)  $f(x) = a + b(x-1)$  – aritmetički niz i b)  $f(x) = a * b^{(x-1)}$  – geometrijski niz, gde je  $a$  prvi element niza, dok je  $b$  razlika bilo koje dve susedne vrednosti niza kod aritmetičkog niza, odnosno količnik bilo koje dve susedne vrednosti kod geometrijskog niza. Za funkcije aritmetičkog i geometrijskog niza implementiran je njihov integral, jer se integraljenje funkcija aritmetičkog ili geometrijskog niza koristi u procesu generisanja formule iz DIRP modela. To praktično znači da ukoliko želimo da u naš sistem ubacimo neku novu klasu funkcija, za koju ćemo raditi interpolaciju i ekstrapolaciju, potrebno je da implementiramo: a) funkciju koja će prepoznati da ulazni niz pripada toj novoj klasi funkcija i b) integral od te klase funkcija. *Naš sistem je proširiv sa svakom funkcijom za koju je moguće odrediti integral.* U planu je da u sistem ubacimo postojeća pravila i metode integraljenja kako bi naš sistem omogućio interpolaciju i ekstrapolaciju različitih klasa funkcija.

Jedan od pravaca našeg daljeg istraživanja jeste procena računске složenosti funkcije pomoću koje se generiše formula  $f_0$  za proizvoljan ulazni niz. Procena računске složenosti biće određena pomoću metrike koja će imati sledeće parametre: a) broj elemenata ulaznog niza brojeva b) broj operacija koje prate diferenciranje kod dobijanja novih nizova c) broj generisanih formula koji prate proces generisanje formule  $f_0$ , d) broj decimala na koji se zaokružuju koeficijenti svih formula koji prate proces generisanja formule  $f_0$  i generisane vrednosti iz formule  $f_0$  i e) vrednosti koje određuju granične tačke u kojima se smanjuje preciznost zaokruživanja vrednosti novih nizova. Pored desktop Java aplikacije koja podržava DIRP model, planiramo u sledećih nekoliko meseci da napravimo i Web Java aplikaciju, kako bi široj naučnoj zajednici omogućili da koriste i testiraju DIRP model.

## 6. ZAKLJUČAK

DIRP model predstavlja jedan od determinističkih modela predikcije skupa brojeva, koji je zasnovan na polinomskoj interpolaciji i ekstrapolaciji skupova brojeva  $(X, Y)$ , nezavisne promenljive  $X$ , koja obrazuje ekvidistantan skup i zavisne promenljive  $Y$ , koja je uređen niz (aritmetički ili geometrijski niz), ili se se diferenciranjem svodi na neki od navedenih uređenih nizova. DIRP model generiše isti (standardni) polinom, kao i Lagranžova, Njutnova i Nevilova metoda, kada je promenljiva  $Y$  zasnovana na aritmetičkom nizu. *DIRP model generiše i eksponencijalni polinom kada je promenljiva  $Y$  zasnovana na geometrijskom nizu, što predstavlja ključni doprinos DIRP modela u odnosu na postojeće determinističke modele.* Razvijena je Java aplikacija koja podržava DIRP model predikcije i dat je primer postupka predikcije skupa brojeva korišćenjem aplikacije. Na kraju su testirani različiti primeri (skupovi brojeva), kategorizovani po tipovima nizova, pomoću DIRP modela. Generisane formule (standardni i eksponencijalni polinomi) za testirane primere precizno interpoliraju i ekstrapoliraju definisani skup ulaznih brojeva.

## LITERATURA

- [1] Allan J Macleod, A comparison of algorithms for polynomial interpolation, Journal of Computational and Applied Mathematics, Volume 8, Issue 4, 1982, Pages 275-277, ISSN 0377-0427
- [2] Chapra, S. C., & Canale, R. P., *Numerical Methods for Engineers* (7th ed.). New York: McGraw-Hill Education, 2015.
- [3] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th Edition, ISBN: 978-1-118-61919-3 May 2013.
- [4] Lepot, M.; Aubin, J.-B.; Clemens, F.H.L.R. Interpolation in Time Series: An Introductory Overview of Existing Methods, Their Performance Criteria and Uncertainty Assessment. *Water* **2017**, *9*, 796. <https://doi.org/10.3390/w9100796>
- [5] Nam, H., Kim, S., & Jung, K., Number Sequence Prediction Problems for Evaluating Computational Powers of Neural Networks. *AAAI*, 2019.
- [6] Siniša Vlajić, Tatjana Stojanović, Miloš Milić, Vojislav Stanojević: Softverski sistem za predikciju sekvence brojeva zasnovan na UN modelu predviđanja, XXIV Međunarodna konferencija Informacione tehnologije (IT) Žabljak, 18 – 22. februar, 2020.
- [7] R. B. Srivastava and Purushottam Kumar Srivastava, Comparison of Lagrange's and Newton's interpolating polynomials, Journal of Experimental Sciences 2012, 3(1): 01-04 ISSN: 2218-1768

## A PARAMETRIC SOLUTION TO WEIGHTED SUM BI-CRITERIA MODELS

MOSHE EBEN-CHAIME

Ben Gurion University of the Negev – department of Industrial Engineering and Management, even@bgu.ac.il

---

**Abstract:** *Models with a weighted sum of a couple criteria are considered. A solution is provided for any weights' combination. The models can be constrained or not, the criteria may take any form and the weight may have any sign. The only requirement is that the each criterion has a finite value anywhere in the feasible region. An initial observation is that the solution depends not on the weights values but on their ratio. Consequently, the first step is to standardize the model such that the weights sum to 1 and solve while changing one weight from 0 to 1. Another insight is that independent of the specific criteria, their weighted sum is linear in the weight ratio. Accordingly, the parametric solution is piecewise linear in the weights' ration. Both observations facilitate the construction of a structured solution scheme in which the model is solved repeatedly with specific weight combinations. The number of repeated solutions is at most three times the number of segments in the parametric solution plus 1.*

**Keywords:** *Bi-Criteria, Parametric Solution, Weighted sum*







# MEKO RAČUNARSTVO

---

## SOFT COMPUTING



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business





## ANALYTIC DESCRIPTION OF FUZZY SET SOLUTIONS TO FUZZY "LP" PROBLEMS

BOGDANA STANOJEVIĆ<sup>1</sup>, MILAN STANOJEVIĆ<sup>2</sup>

<sup>1</sup> Mathematical Institute of the Serbian Academy of Sciences and Arts, bgdnpop@mi.sanu.ac.rs

<sup>2</sup> University of Belgrade – Faculty of Organizational Sciences, milan.stanojevic@fon.bg.ac.rs

**Abstract:** We aim to solve fuzzy linear programming problems with coefficients expressed by non-negative trapezoidal fuzzy numbers. Approaching the same class of problems, a recent work from the literature provided a methodology to derive numerical description of the membership function of the optimal fuzzy value of the objective function, and informed on its impossibility to disclose relevant characterization of the optimal fuzzy solutions. We expand that results by proposing analytic descriptions to the fuzzy sets representing the optimal values and solutions through a parametric analysis of the optimal solutions to crisp linear optimization problems. We illustrate our theoretic results through a numerical example recalled from the literature.

**Keywords:** Fuzzy Linear Programming, Fuzzy Numbers, Extension Principle, Analytic Approach

### 1. INTRODUCTION

In the recent literature (see for instance [1], [2] and [5]) the importance of reinstating the extension principle ([7], [8]) was emphasized, and novel research directions were suggested.

Fuzzy number LP problems as they are called in [6], or full fuzzy LP problems as they are called in [4] and [5] are formally linear programming problems with uncertain coefficients expressed by fuzzy numbers that implies fuzzy decision variables as well. Approaching this class of problems, Wang and Peng [6] provided a methodology to derive numerical description to the membership function of the optimal fuzzy value of the objective function, and informed on its impossibility to disclose relevant characterization of the optimal fuzzy solutions. Expanding their results we propose an analytic descriptions to the fuzzy sets representing the optimal values and solutions via a parametric analysis of the optimal solutions to crisp linear optimization problems.

The rest of the paper is organized as follows. Section 2 describes the notation and terminology, Section 3 briefly presents our novel optimization approach based on the extension principle, Section 4 reports our numerical results, and Section 5 concludes the paper with our final remarks and ideas for a future research.

### 2. PRELIMINARIES

Zadeh [7] introduced the concept of fuzzy set  $\tilde{A}$  over the universe  $X$  as a collection of pairs  $(x, \mu_{\tilde{A}}(x))$ , where  $x \in X$  and  $\mu_{\tilde{A}}(x) \in [0, 1]$ . Function  $\mu_{\tilde{A}} : X \rightarrow [0, 1]$  is the membership function that describes the fuzzy set  $\tilde{A}$ , and  $\mu_{\tilde{A}}(x)$  is the membership degree of  $x$  in  $\tilde{A}$ .

The support of a fuzzy set  $\tilde{A}$  is the set of values with strictly positive membership degree, i.e.  $\text{Supp}(\tilde{A}) = \{x \in X | \mu_{\tilde{A}}(x) > 0\}$ . The  $\alpha$ -cut of a fuzzy set  $\tilde{A}$  is denoted by  $[\tilde{A}]_{\alpha}$  and it is defined as the set of values with a membership degree greater or equal to  $\alpha$ ,  $\alpha > 0$ , i.e.  $[\tilde{A}]_{\alpha} = \{x \in X | \mu_{\tilde{A}}(x) \geq \alpha\}$ .

Fuzzy numbers are special cases of fuzzy sets. A fuzzy set  $\tilde{A}$  of the universe  $R$  of real numbers is called fuzzy number if and only if: (i) it is fuzzy normal and fuzzy convex; (ii) its membership function  $\mu_{\tilde{A}}$  is upper semi-continuous; and (iii) its support  $\{x \in R | \mu_{\tilde{A}}(x) > 0\}$  is bounded. The  $\alpha$ -cuts of the fuzzy numbers are always intervals.

We use trapezoidal fuzzy numbers in our presentation. The non-zero piece of the membership function of a trapezoidal fuzzy number forms a trapezoid with the abscissa when it is graphically represented. Generally, a trapezoidal fuzzy number  $\tilde{A}$  is expressed by a quadruple  $(a_1, a_2, a_3, a_4)$ ,  $a_1 \leq a_2 \leq a_3 \leq a_4$ . The interval  $(a_1, a_4)$  is  $\text{Supp}(\tilde{A})$ , and the interval  $[a_2, a_3]$  describes the set of values with the membership degree equal to 1. The  $\alpha$ -cut of a trapezoidal fuzzy number  $\tilde{A} = (a_1, a_2, a_3, a_4)$  is the interval  $[\tilde{A}]_{\alpha} = [(1 - \alpha)a_1 + \alpha a_2, \alpha a_3 + (1 - \alpha)a_4]$ . The inequality  $\mu_{\tilde{A}}(x) \geq \alpha$  is then equivalent to the double inequality  $(1 - \alpha)a_1 + \alpha a_2 \leq x \leq \alpha a_3 + (1 - \alpha)a_4$ .

Whenever  $a_2 = a_3$  in a trapezoidal fuzzy number, that fuzzy number is reduced to a so called triangular fuzzy number.

Zadeh [7] proposed the extension principle – as one of the basic ideas in fuzzy set theory – and applied it to develop the fuzzy arithmetic on fuzzy quantities. Ross [3] presented several methods to convert the extended fuzzy operations into efficient computational algorithms.

According to the extension principle, the fuzzy set  $\tilde{B}$  over the universe  $Y$  that is the result of evaluating the function  $f$  at the fuzzy sets  $\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_r$  over their universes  $X_1, X_2, \dots, X_r$  is defined through its membership function

$$\mu_{\tilde{B}}(y) = \begin{cases} \sup_{(x_1, \dots, x_r) \in f^{-1}(y)} \left( \min \left\{ \mu_{\tilde{A}_1}(x_1), \dots, \mu_{\tilde{A}_r}(x_r) \right\} \right), & f^{-1}(y) \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

See Ross [3] for more details on fuzzy arithmetic and extension principle.

### 3. THE NEW METHODOLOGY

The general form of the fuzzy LP problem we aim to solve is

$$\begin{aligned} \min \quad & \tilde{c}^T x \\ \text{s.t.} \quad & \tilde{A}x \geq \tilde{b}, \\ & x \geq 0, \end{aligned} \tag{2}$$

where

$$\begin{aligned} \tilde{c} &= (\tilde{c}_j)_{j=\overline{1,n}}, & \tilde{b} &= (\tilde{b}_i)_{i=\overline{1,m}}, & \tilde{A} &= (\tilde{a}_{ij})_{i=\overline{1,m}, j=\overline{1,n}}, \\ \tilde{c}_i &= (c_i^1, c_i^2, c_i^3), \quad i = \overline{1,n}, & \tilde{b}_i &= (b_i^1, b_i^2, b_i^3), \quad i = \overline{1,m}, & \tilde{a}_{ij} &= (a_{ij}^1, a_{ij}^2, a_{ij}^3), \quad i = \overline{1,m}, j = \overline{1,n}. \end{aligned}$$

Wang and Peng [6] employed the  $\alpha$ -cut formulas for triangular fuzzy numbers, and the following parametric LP problems (3) and (4) to derive the increasing and decreasing parts of the optimal objective fuzzy value membership function, respectively.

$$\min \left\{ c_L(\alpha)^T x \mid A_U(\alpha)x \geq b_L(\alpha), x \geq 0 \right\}, \tag{3}$$

$$\min \left\{ c_U(\alpha)^T x \mid A_L(\alpha)x \geq b_U(\alpha), x \geq 0 \right\}, \tag{4}$$

where

$$\begin{aligned} c_L(\alpha) &= ((1 - \alpha)c_j^1 + \alpha c_j^2)_{j=\overline{1,n}}, & b_L(\alpha) &= ((1 - \alpha)b_i^1 + \alpha b_i^2)_{i=\overline{1,m}}, & A_L(\alpha) &= ((1 - \alpha)a_{ij}^1 + \alpha a_{ij}^2)_{i=\overline{1,m}, j=\overline{1,n}}, \\ c_U(\alpha) &= ((1 - \alpha)c_j^3 + \alpha c_j^2)_{j=\overline{1,n}}, & b_U(\alpha) &= ((1 - \alpha)b_i^3 + \alpha b_i^2)_{i=\overline{1,m}}, & A_U(\alpha) &= ((1 - \alpha)a_{ij}^3 + \alpha a_{ij}^2)_{i=\overline{1,m}, j=\overline{1,n}}. \end{aligned}$$

They solved Problems (3) and (4) for several values of the parameter  $\alpha$  and derived numerically the corresponding  $\alpha$ -cut intervals of the optimal objective fuzzy value. They also collected the optimal solutions and tried to construct the membership functions of the fuzzy sets of optimal solution values. However, Wang and Peng [6] noticed that the so collected membership degrees generally can not describe a fuzzy number, thus finding the exact description of the fuzzy sets of the optimal solution values remained an open question.

The novelty of our approach is twofold: (i) we propose an analytic description of the fuzzy set optimal objective value, using the analytic analysis of Problems (3) and (4); and introduce two new quadratic optimization models, namely (5) and (6), to derive the membership functions of the optimal solution values.

$$\begin{aligned} \min \quad & x_i \\ \text{s.t.} \quad & A_U(\alpha)x \geq b_L(\alpha), \\ & c_L(\alpha)x \leq z_U(\alpha), \\ & c_U(\alpha)x \geq z_L(\alpha), \\ & x \geq 0. \end{aligned} \tag{5}$$

**Table 1:** The fuzzy number coefficients of Problem (7)

Coefficients	TrFNs	Quadruples	$\alpha$ -cut intervals
$\tilde{c}_1$	8.5	(4, 8, 9, 12)	$[4 + 4\alpha, 12 - 3\alpha]$
$\tilde{c}_2$	3.5	(2, 3, 4, 6)	$[2 + \alpha, 6 - 2\alpha]$
$\tilde{a}_{11}$	$\tilde{23}$	(19, 21, 25, 26)	$[19 + 2\alpha, 26 - \alpha]$
$\tilde{a}_{12}$	$\tilde{6}$	(2, 4, 8, 10)	$[2 + 2\alpha, 10 - 2\alpha]$
$\tilde{a}_{21}$	14	(9, 12, 16, 19)	$[9 + 3\alpha, 19 - 3\alpha]$
$\tilde{a}_{22}$	$\tilde{8}$	(6, 7, 9, 12)	$[6 + \alpha, 12 - 3\alpha]$
$\tilde{b}_1$	1100	(600, 1000, 1200, 1700)	$[600 + 400\alpha, 1700 - 500\alpha]$
$\tilde{b}_2$	900	(500, 700, 1100, 1300)	$[500 + 200\alpha, 1300 - 200\alpha]$

$$\begin{aligned}
 &\max \quad x_i \\
 &\text{s.t.} \quad A_U(\alpha)x \geq b_L(\alpha), \\
 &\quad \quad c_L(\alpha)x \leq z_U(\alpha), \\
 &\quad \quad c_U(\alpha)x \geq z_L(\alpha), \\
 &\quad \quad x \geq 0.
 \end{aligned} \tag{6}$$

Within Models (5) and (6), the expressions of  $z_L(\alpha)$  and  $z_U(\alpha)$  represent the endpoints of the interval of the minimal values of the original objective function. The endpoints are computed with respect to  $\alpha$ , and all combination of coefficients within their  $\alpha$ -cut intervals in Problem (2). The expressions of  $z_L(\alpha)$  and  $z_U(\alpha)$  are obtained by solving Problems (3) and (4), respectively.

#### 4. COMPUTATION RESULTS

To illustrate our theoretical statements we recall the first numerical example used in [6]. Thus we solve the following fuzzy number valued linear programming problem

$$\begin{aligned}
 &\min \quad \tilde{8.5}x_1 + \tilde{3.5}x_2 \\
 &\text{s.t.} \quad \tilde{23}x_1 + \tilde{6}x_2 \geq \tilde{1100}, \\
 &\quad \quad \tilde{14}x_1 + \tilde{8}x_2 \geq \tilde{900}, \\
 &\quad \quad x_1, x_2 \geq 0,
 \end{aligned} \tag{7}$$

where the fuzzy coefficients are expressed by the trapezoidal fuzzy numbers given in Table 1:.

Solving the parametric problem

$$\begin{aligned}
 &\min \quad (4 + 4\alpha)x_1 + (2 + \alpha)x_2 \\
 &\text{s.t.} \quad (26 - \alpha)x_1 + (10 - 2\alpha)x_2 \geq 600 + 400\alpha, \\
 &\quad \quad (19 - 3\alpha)x_1 + (12 - 3\alpha)x_2 \geq 500 + 200\alpha, \\
 &\quad \quad x \geq 0,
 \end{aligned}$$

that is the particular form of (3), we obtain

$$\begin{aligned}
 x_1^{*L}(\alpha) &= \frac{2200 + 2000\alpha + 800\alpha^2}{122 + 22\alpha + 3\alpha^2} \\
 x_2^{*L}(\alpha) &= \frac{1600 - 1100\alpha + 1000\alpha^2}{122 + 22\alpha + 3\alpha^2}
 \end{aligned} \tag{8}$$

$$z_L(\alpha) = \frac{12000 + 16200\alpha + 5700\alpha^2 - 2200\alpha^3}{122 + 22\alpha + 3\alpha^2}. \tag{9}$$

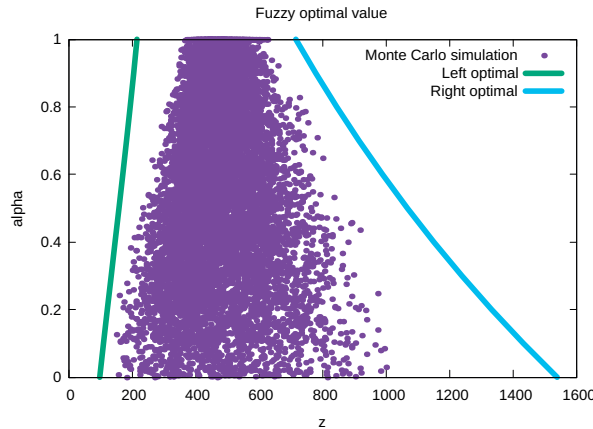


Figure 1 The graphic representation of optimal  $\tilde{z}$

Solving

$$\begin{aligned} \min \quad & (12 - 3\alpha)x_1 + (6 - 2\alpha)x_2 \\ \text{s.t.} \quad & (19 + 2\alpha)x_1 + (2 + 2\alpha)x_2 \geq 1700 - 500\alpha, \\ & (9 + 3\alpha)x_1 + (6 + \alpha)x_2 \geq 1300 - 200\alpha, \\ & x \geq 0, \end{aligned}$$

that is the particular form of (4), we obtain

$$\begin{aligned} x_1^{*U}(\alpha) &= \frac{1900 - 875\alpha - 25\alpha^2}{24 + 1.75\alpha - \alpha^2} \\ x_2^{*U}(\alpha) &= \frac{2350 - 450\alpha + 275\alpha^2}{24 + 1.75\alpha - \alpha^2} \end{aligned} \tag{10}$$

$$z_U(\alpha) = \frac{36900 - 23600\alpha + 4875\alpha^2 - 475\alpha^3}{24 + 1.75\alpha - \alpha^2}. \tag{11}$$

We report our numerical results and graphically compare them to the results from the literature in Figures 1 and 2. Figure 1 shows the Monte Carlo simulation of the optimal objective values derived in accordance to the extension principle, and the exact membership function of the fuzzy set of the optimal objective value obtained by our parametric approach. The results obtained in [6] fully match with our results.

Figure 2 shows: (i) the Monte Carlo simulation of the optimal solution values derived in accordance to the extension principle; (ii) the exact membership function of the fuzzy set of the optimal solutions obtained with the help of Models (5) and (6); and (iii) the optimal solutions obtained in [6] and confirmed analytically by the expressions given in (8) and (10) that obviously mislead since they are not derived in accordance to the extension principle.

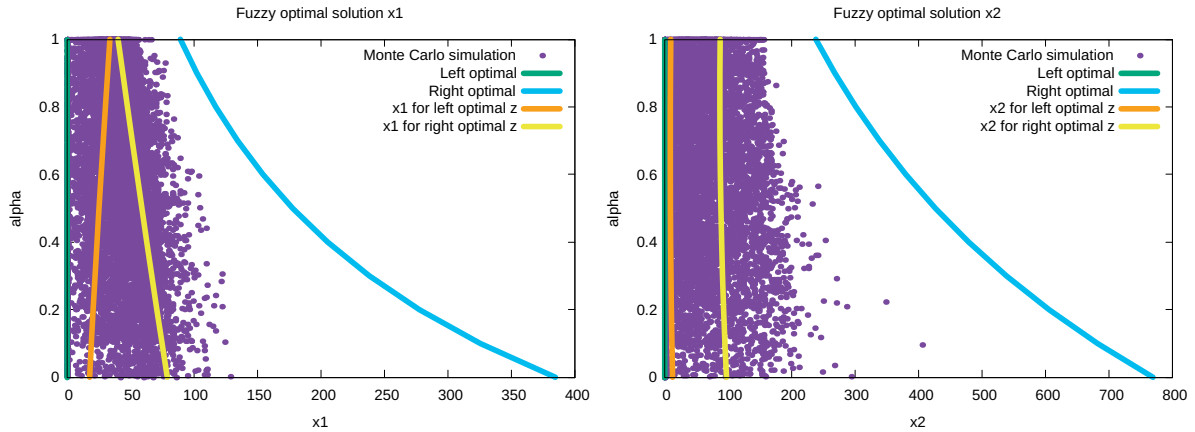
The optimal fuzzy solutions  $\tilde{x}^L$  and  $\tilde{x}^U$  were obtained by quadratic optimizations using Models (5) and (6).

### 5. FINAL REMARKS

In this paper we approached the fuzzy LP problems relying on the  $\alpha$ -cuts of the fuzzy parameters, and the parametric mathematical models recalled from the literature; advanced a new perspective; and succeeded to provide an analytical alternative to the numerical methods used so far in the literature.

We also derived the optimal fuzzy set solutions to the original problem, thus extended the procedure from the literature.

The proposed approach works on fuzzy LP problems with non-negativity constraints on its decision variables. Finding the needed modifications able to make it work on a larger class of problems is a direction for further researches.



**Figure 2** The graphic representation of misleading vs optimal  $\tilde{x}_1$  and  $\tilde{x}_2$

### Acknowledgement

This work was supported by the Serbian Ministry of Education, Science and Technological Development through Mathematical Institute of the Serbian Academy of Sciences and Arts and Faculty of Organizational Sciences of the University of Belgrade.

### REFERENCES

- [1] Diniz, M.M., Gomes, L.T. & Bassanezi, R.C. (2021). Optimization of fuzzy-valued functions using Zadeh's extension principle. *Fuzzy Sets and Systems*, 404:23 – 37.
- [2] Kupka.J. On approximations of zadeh's extension principle. *Fuzzy Sets and Systems*, 283:26 – 39, 2016. Theme: Fuzzy Intervals.
- [3] Ross. T.J. (2010). Fuzzy Arithmetic and the Extension Principle, chapter 12, pages 408–436. John Wiley & Sons, Ltd.
- [4] Stanojević, B. & Stanojević. M. (2021). Empirical versus analytical solutions to full fuzzy linear programming, volume 1243 of *Advances in Intelligent Systems and Computing*. Springer, Cham.
- [5] Stanojević, B., Stanojević. M. & Nădăban, S. (2021). Reinstatement of the extension principle in approaching mathematical programming with fuzzy numbers. *Mathematics*, 9(11).
- [6] Wang, G. & Peng, J. (2019). Fuzzy optimal solution of fuzzy number linear programming problems. *Int. J. Fuzzy Syst.*, 21:865–881.
- [7] Zadeh, L.A. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353.
- [8] Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning I. *Information Sciences*, 8(3):199–249.





## ANALIZA I PREDVIĐANJE BROJA NOVOZARAŽENIH TOKOM COVID-19 PANDEMIJE PRIMENOM ANSAMBL ALGORITAMA

### ANALYSIS AND PREDICTION OF NEW INFECTIONS DURING COVID-19 PANDEMIC USING ENSEMBLE ALGORITHMS

NIKOLA ĐORĐEVIĆ<sup>1</sup>, ALEKSANDAR RAKIĆEVIĆ<sup>3</sup>, NEMANJA MILENKOVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, nd20213514@student.fon.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, aleksandar.rakicevic@fon.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, nemanja.milenkovic@fon.bg.ac.rs

---

**Rezime:** *Pandemija COVID-19 virusa odnela je veliki broj života i, čini se, trajno izmenila društvene, ekonomske i političke tokove širom sveta. Pravovremeno korišćenje metoda mašinskog učenja moglo je pomoći u boljem upravljanju pandemijom. Cilj ovog rada je da istraži mogućnost primene ansambl algoritama za predviđanje broja novozaraženih COVID-19 virusom. Analize i predikcije širenja COVID-19 izvršene su na osnovu skupa podataka koji prati različite pokazatelje za 209 zemalja u periodu od dve godine. Za predviđanje broja novozaraženih u okviru ansambl algoritma primenjene su veštačke neuronske mreže. Kao ulazni podaci koriste se indikatori formirani na osnovu pokretnog proseka i pokretne standardne devijacije. Za merenje performansi korišćene su srednja apsolutna greška, kao i koren srednje kvadratne greške.*

**Ključne reči:** *COVID-19 pandemija, ansambl algoritmi, mašinsko učenje, neuronska mreža.*

**Abstract:** *COVID-19 pandemic has taken a lot of lives, and permanently impacted economic and political streams across the world. Timely usage of machine learning methods could have been of help for a better response to the pandemic. The goal of this paper was to create a machine learning model able to predict the future number of infections with COVID-19. Analysis and prediction of infection trends are done on data collected from 209 countries across the globe in a period of two years. With the focus on ensemble algorithms, artificial neural networks were used to predict the number of newly infected. The model was trained on variables created from moving averages and moving standard deviations. Measuring model performance was done using metrics like mean absolute error and root mean squared error.*

**Keywords:** *COVID-19 pandemic, ensemble algorithm, machine learning, neural network.*



## FAZI SISTEM ZA ALGORITAMSKO TRGOVANJE ZASNOVAN NA TEHNIČKOJ ANALIZI

### FUZZY SYSTEM FOR ALGORITHMIC TRADING BASED ON TECHNICAL ANALYSIS

ANDRIJA MATIĆ<sup>1</sup>, MILICA ZUKANOVIĆ<sup>2</sup>, ALEKSANDAR RAKIĆEVIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, andrija.matic42@gmail.com

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, milica.zukanovic@fon.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, aleksandar.rakicevic@fon.bg.ac.rs

---

**Rezime:** Cilj ovog rada je projektovanje sistema za algoritamsko trgovanje zasnovanom na tehničkoj analizi cenovnih podataka i njegovo testiranje na međunarodnom deviznom tržištu. Svrha sistema jeste da istraži mogućnosti primene fazi zaključivanja u automatizovanom trgovanju finansijskim instrumentima. Sistem se sastoji više komponenti (tehničkih indikatora) koji na osnovu cenovnih podataka izvlače informacije o trendu, volatiliteti i sentimentu tržišnih učesnika. Na osnovu unapred definisanih logičkih pravila zaključivanja (trgovačke strategije), sistem samostalno ispostavlja kupovne/prodajne naloge i nadgleda otvorene pozicije. Trgovačka logika implementirana je korišćenjem klasičnog binarnog (Bulovog) i fazi kontrolera. Implementacija fazi logike pruža mogućnost da se sistem algoritamskog trgovanja nosi sa nepreciznošću i nelinearnošću u složenom okruženju kao što je devizno tržište. Predloženi sistem je testiran nad višegodišnjim istorijskim podacima o kretanju više valutnih parova. Upoređene su performanse klasične binarnog i fazi kontrolera u upravljanju trgovanjem kroz više vremenskih intervala.

**Ključne reči:** Algoritamsko trgovanje, trgovačka strategija, fazi zaključivanje, tehnička analiza, međunarodno devizno tržište.

**Abstract:** The purpose of this paper is to propose a system for algorithmic trading based on technical analysis of price data and to test it on the Foreign exchange market. The goal of this system is to investigate whether the fuzzy inference could be successfully applied for algorithmic trading of financial instruments. The proposed system is consisted of few components (technical indicators) with the task to extract information on trend, volatility and market sentiment from price data. Based on the predefined inference rules (trading logic), the system is able to place buy/sell orders, and monitor open positions. Trading logic is implemented using both classical binary (Boolean) and fuzzy controller. Implementation of fuzzy logic provides the algorithmic trading system with the capability to deal with imprecision and nonlinearity in complex environment such is the foreign exchange market. The proposed system was backtested on multi-year historical data using several currency pairs. We compared performance of classical binary and fuzzy logic in trading control accross several time intervals.

**Keywords:** Algorithmic trading, trading strategy, fuzzy inference, technical analysis, foreign exchange market.



## PRIMENA FAZI KLASTEROVANJA ZA KATEGORISANJE TENISKIH IGRAČA

### CATEGORIZING TENNIS PLAYERS USING FUZZY CLUSTERING TECHNIQUE

LUKA TASIĆ<sup>1</sup>, ANA POLEDICA<sup>2</sup>, MILICA ZUKANOVIĆ<sup>3</sup>

<sup>1</sup> Connect The Dots, tasiicluka123@gmail.com

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, ana.poledica@fon.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, zukanovic.milica@gmail.com

**Rezime:** *Primena tehnika računarske inteligencije je sve više zastupljena u sportskoj analitici. Prilikom analize teniskih mečeva i igrača, postoji mnogo faktora koji se mogu razmatrati kako bi se utvrdilo koji će teniseri biti uspešniji. Osnovne statistike kao što su procenat prvog servisa, procenat osvojenih brejk lopti, procenat drugog servisa i druge daju uvid u stil igre i prednosti svakog igrača. Osnovni cilj ovog rada je da se na osnovu podataka o teniskim mečevima od 2000. do 2022. godine grupišu teniseri, utvrde ključne karakteristike tenisera i identifikuju klasteri prema stilu igre. Primenjen je algoritam fazi C-srednjih vrednosti koji je uspešno izdvojio 3 karakteristične grupe tenisera. Takođe, pokazalo se da postoje igrači koji prirodno imaju značajnu pripadnost više klastera koji su posebno razmatrani.*

**Ključne reči:** *fazi klasterovanje, fazi C-srednjih vrednosti, kategorisanje, tenis*

**Abstract:** *The application of computer intelligence techniques is increasingly present in sports analytics. When analyzing tennis matches and players, there are many factors that can be considered in order to determine which tennis players will be more successful. Basic statistics such as the percentage of the first serve, the percentage of break points won, the percentage of the second serve and others give an insight into the style of play and the advantages of each player. The main goal of this paper is to group tennis players based on data of tennis matches from 2000 to 2022, determine the key characteristics of tennis players and identify clusters according to the style of play. The fuzzy C-mean algorithm was applied, which successfully singled out 3 characteristic groups of tennis players. Also, it has been shown that there are players who naturally have a significant affiliation to multiple clusters that have been specifically considered.*

**Keywords:** *fuzzy clustering, fuzzy C-means, categorizing, tennis*

## 1. UVOD

Poslednjih godina metode računarske inteligencije i analize podataka postale su značajni alati u sportskoj analitici i zauzimaju značajnu ulogu kod sve većeg broja različitih sportskih disciplina npr. [3,6].

I u tenisu napredna analitika dobija sve veći značaj. Postoji veliki broj faktora koji mogu uticati na sam ishod meča na šta ukazuju i mnoge studije koje su se time bavile:

- Najbolji profesionalni rangovi igrača rođenih 1982. ili ranije bili su pozitivno povezani sa uzrastom u kojem su igrači zaradili svoj prvi ATP poen, a zatim ušli u prvih 100, što sugeriše da godine povezane sa ovim prekretnicama u rangiranju mogu imati određeni potencijal za predviđanje [14].
- Sila tokom teniskog udarca forhendom korelira sa veličinom hvata [12]. Autori su takođe otkrili korelaciju pola i iskustva sa udarnom silom.
- Tenis na travi zahteva specifične veštine za postizanje visokog ranga u svetu, kao što su veštine servisa i vraćanja servisa. Servis je najvažnija veština u tenisu i važniji je kod muškaraca i uspešniji je u dublu [7].
- Postoji veza između količine i nivoa konkurencije i kognitivna stručnosti nivo takmičenja mogu biti od ključnog značaja za razvoj stručnosti tenisera [8].

- U radu [17] su ispitani atributi i performanse 100 najboljih tenisera i izvršena je klaster analiza. Jedan klaster su igrači koji preuzimaju rizik, a drugi su igrači koji izbegavaju rizik.  $K$ -srednjih vrednosti (eng.  $k$ -Means) tehnika grupisanja je korišćena da bi se razumeli fizički atributi igrača tokom igre. Razmatrajući klastere otkriveno je da su duple greške i asovi važne ulazne varijable u predviđanju rizične prirode sportista na teniskoj travi.

Ovaj rad je nastao kao deo šireg istraživanja koje se bavi poređenjem različitih algoritama klasterovanja pri kategorizaciji teniskih igrača. Jedan od najčešćih algoritama koji se primenjuje u kategorizaciji sportista je algoritam  $k$ -srednjih vrednosti i uspešno je primenjen u klasterovanju igrača u košarci [15], američkom fudbalu [10], tenisu [16] itd. Nakon polaznih rezultata istraživanja za algoritam  $k$ -srednjih vrednosti, uvideli smo da postoji više karakterističnih igrača koji ne pripadaju strogo jednom klasteru, već poseduju karakteristike tenisera drugog klastera. Iako su tradicionalne tehnike tvrdog klasterovanja (eng. hard clustering) uspešne kada se klasteri mogu jasno razdvojiti i definisati, manje su pogodne u situacijama kada ne se ne mogu uspostaviti jasne granice i kada nam je važno da posebno pažnju posvetimo graničnim slučajevima. Tada je pogodnije koristiti metode mekog klasterovanja (eng. soft clustering) gde objekat može pripadati više klastera istovremeno.

U okviru ovog rada primenjen je algoritam fazi  $C$ -srednjih vrednosti, kako bismo obezbedili da za svaku instancu postoji nekoliko deskriptora pripadnosti klasterima. Podaci koji su korišćeni za analizu preuzeti su sa sajta Tennis Abstract<sup>1</sup> koji održava teniski fan Džef Sakman gde su javno dostupne statistike teniskih mečeva koje datiraju od 1968. godine. Za analizu posmatran je period od 2000. do 2022. godine. U okviru istraživanja analizirane su osnovne statistike koje daju uvid u stil igre i relativne prednosti svakog igrača kao što su procenat prvog servisa, procenat osvojenih poena na prvi servis, procenat osvojenih poena na drugi servis i druge. Skup podataka ne sadrži metrike nadigravanja kao što je dužina nadigravanja, da li je poen direktan (viner), osvojen iznuđenom, neiznuđenom greškom ili na mreži. Osnovna pretpostavka je da se primenom algoritma fazi  $C$  srednjih vrednosti mogu grupisati teniseri i pronaći obrasci različitih stilova igre koji karakterišu svaki klaster.

U nastavku rada, u okviru poglavlja 2 data je teorijska osnova algoritma fazi  $C$ -srednjih vrednosti. Zatim su opisani podaci i odabrani atributi koji su korišćeni kao ulaz u klasterovanje. U poglavlju 4 prikazani su rezultati fazi klasterovanja i analizirani dobijeni klasteri. Na kraju dat je zaključak istraživanja i predlozi za dalji rad.

## 2. KLASTEROVANJE METODOM FAZI C-SREDNJIH VREDNOSTI

Klasterovanje je metod za grupisanje opservacija na osnovu njihove sličnosti. Klaster analiza je značajna za mnoštvo oblasti poput prepoznavanja patern, analize podataka, obrade slika i istraživanja tržišta [9]. Uopšteno govoreći, klasterovanje se može podeliti u dve podgrupe, tvrdo klasterovanje i meko klasterovanje [13]. Algoritmi tvrdog klasterovanja, svaku opservaciju dodeljuju samo jednom klasteru i jedan od široko korišćenih jeste algoritam  $K$ -srednjih vrednosti [2].

Od nedavno, meko klasterovanje pokazalo se kao superiornije u odnosu na tradicionalne tvrde algoritme za klasterovanje [18]. Kod mekog klasterovanja, samim tim i kod algoritma fazi  $C$ -srednjih vrednosti (eng. Fuzzy  $C$ -Means, FMC), svaka opservacija pripada svakom klasteru, iako su funkcije pripadnosti drugačije. Nivo pripadnosti kod pomenutih algoritma zavisi od blizine opservacije od centroida, odnosno od centra klastera. Što je opservacija bliža centroidu, veća je vrednost pripadanja i obrnuto. Svrha ovakvog načina klasterovanja je minimizacija funkcije cilja kako bi se dokazala sličnost između opservacija u klasteru i smanjila sličnost između opservacija u poređenju sa drugim klasterima [1]. Funkcija cilja može da se definiše na sledeći način:

$$J = \sum_{i=1}^n \sum_{j=1}^c \mu_{ij}^m \|x_i - r_j\|^2 \quad (1)$$

gde je  $n$  broj instanci,  $c$  broj klastera,  $x_i$   $i$ -ta opservacija,  $r_j$   $j$ -ti centroid,  $\mu_{ij}$  funkcija pripadnosti  $i$ -te opservacije  $j$ -tom centroidu, gde funkcija ima vrednost od 0 do 1, uključujući te vrednosti za svako  $i=1,2,\dots,n$  i svako  $j=1,2,\dots,c$ ,  $m$  fazi stepen i  $\|x_i - r_j\|$  Euklidsko rastojanje između  $i$ -te opservacije od  $j$ -tog centroida [5].

Prednosti fazi grupisanja u odnosu na klasične metode grupisanja se ogledaju u tome što pruža detaljnije informacije o podacima, kao i vrednosti funkcija pripadnosti koje su pogodne za komentarisanje, fleksibilan je u korišćenju odstojanja i prilagodljiv je numeričkoj optimizaciji [11]. S druge strane, ima i nedostatke. U

<sup>1</sup> www.tennisabstract.com

slučaju velikog broja opservacija i klastera, teško je sumirati i klasifikovati podatke prilikom mnogobrojnih izlaza. Štaviše, algoritmi fazi klasterovanja, koji se koriste kada postoji nesigurnost, generalno su komplikovani [4].

### 3. OPIS PROBLEMA I PODATAKA

Osnovni motiv ovog rada je primena tehnike fazi klasterovanja kako bi se izvršilo identifikovanje grupa teniskih igrača i analizirale njihove performanse i osobine. Izučavani problem ovog rada je analiza statistika teniskih mečeva kako bi se dalje pronašli obrasci različitih stilova igre za svaku grupu. Klasterovanje metodom fazi *C*-srednjih vrednosti je rađeno sa ciljem da se izdvoje najbolji teniseri, kako bi se utvrdilo koje od karakteristika su najbitnije za uspeh u tenisu. Takođe, posebna pažnja će biti posvećena graničnim slučajevima, odnosno igračima koji sa velikim stepenom pripadaju više klastera što je jedna od prednosti FCM-a.

Na osnovu preuzetih podataka<sup>2</sup> o teniskim mečevima odigranih od januara 2000. do marta 2022. godine, gde svaki red predstavlja jedan meč i uključuje statistike o pobedniku i gubitniku, napravljen je novi skup podataka za potrebe fazi klasterovanja. Izvršeno je filtriranje, prvo teniskih turnira, a potom i samih igrača. U obzir su uzeti samo turniri koji se igraju na tri dobijena seta, igrači koji imaju najmanje 50 mečeva na pomenutim turnirima i koji su ikad bili u prvih 50 pozicija na ATP listi. Usled promene pozicija na ATP listi, u dobijenom skupu podataka analizirana su 133 teniska igrača koji zadovoljavaju prethodne kriterijume.

Odabrani su atributi koji se smatraju osnovnim statistikama u teniskoj analitici:

- procenat osvojenih poena na prvi servis (A1)
- procenat osvojenih poena na drugi servis (A2)
- procenat as servisa (A3)
- procenat duplih grešaka (A4)
- prosečan broj gemova po meču (A5)
- procenat dobijenih tajbrejkova (A6)
- procenat spašenih brejk lopti (A7)
- procenat osvojenih brejkova (A8)

### 4. EKSPERIMENT I REZULTATI

Nakon kreiranja skupa podataka, sledeća faza obuhvatala je detaljniju analizu podataka i pripremu za fazi klasterovanje. U prvom koraku izvršena je standardna min-max normalizacija, kao i analiza nestandardnih opservacija koje su zadržane u skupu zbog poznatog porekla i značajnih informacija. Zarad detaljnijeg razumevanja podataka, izračunata je korelaciona matrica između atributa. Očekivano, procenat osvojenih poena na prvi servis (A1) je veoma visoko korelisan sa atributima: procenat asova (A3) (0.89) i procenat spašenih brejk lopti (A7) (0.7). Imajući u vidu da je procenat prvog servisa veoma bitan faktor u igri tenisera, analizirana su dva modela, sa i bez atributa koji opisuje procenat osvojenih poena na prvi servis (A1). Utvrđeno je da se izbacivanjem promenljive (A1) gube informacije, odnosno da je dobijena bolja struktura klastera sa uključenim svim atributima. Za određivanje optimalnog broja klastera korišćene su lakat metoda (eng. elbow method) i metoda siluete (eng. silhouette method). Primenom lakat metode utvrđeno je da je optimalan broj klastera tri, dok je metoda silueta pokazala dva klastera kao optimalan broj. Nakon analize, tri klastera su se pokazala kao odgovarajuća.

Klasterovanje je izvršeno metodom fazi *C*-srednjih vrednosti, dok je za određivanje distance primenjeno standardno Euklidsko odstojanje. Za potrebe ovog eksperimenta korišćen je programski jezik R, odnosno R studio i biblioteka *fcm*. S obzirom na relativno mali broj ulaznih atributa i opservacija, vremenska i prostorna ograničenja algoritma nisu bila kritična. Korišćene su podrazumevane vrednosti: parameter *m* je 2, maksimalan broj iteracija je 1000, kriterijum zaustavljanja je  $10^{-9}$ , itd.

<sup>2</sup> [https://github.com/JeffSackmann/tennis\\_atp](https://github.com/JeffSackmann/tennis_atp)

Sumirane vrednosti FCM-a, uključujući i promenljivu A1 date su u Tabeli 1. Pripadnost tenisera klasteru određena je na osnovu najveće vrednosti pripadnosti. Jaka pripadnost klasteru je definisana u slučaju kada je pripadnost veća ili jednaka 0.5. Centroidi klastera dati su u Tabeli 2. Slika klastera prikazana je na Slici 1.

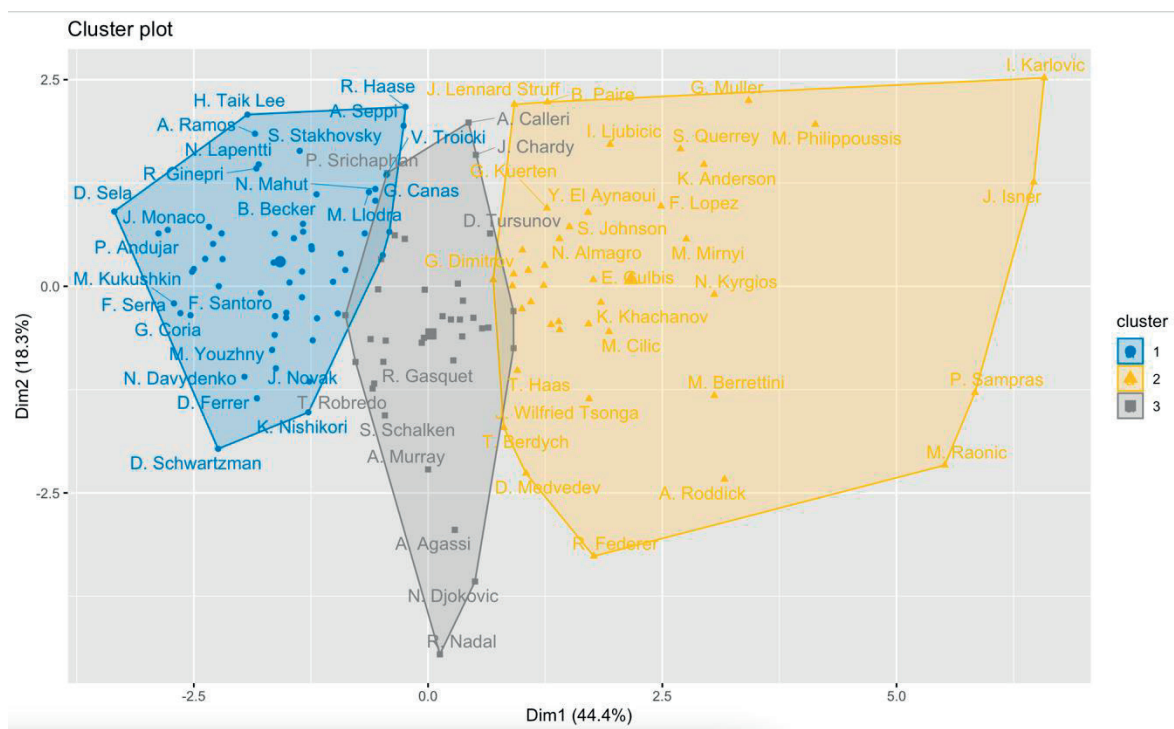
**Tabela 1:** Rezultati klasterovanja metodom FCM-a – metrike klastera

Metrike	Klaster 1	Klaster 2	Klaster 3
Broj tenisera	57	41	35
Najveća pripadnost	0.645	0.776	0.584
Najmanja pripadnost	0.389	0.395	0.359
Prosečna pripadnost	0.516	0.537	0.455

**Tabela 2:** Rezultati klasterovanja metodom FCM-a – centroidi klastera

Promenljive	A1	A2	A3	A4	A5	A6	A7	A8
Klaster 1	0.328	0.319	0.194	0.254	0.330	0.455	0.335	<b>0.673</b>
Klaster 2	<b>0.629</b>	<b>0.476</b>	<b>0.433</b>	<b>0.312</b>	<b>0.474</b>	<b>0.558</b>	<b>0.555</b>	0.548
Klaster 3	0.457	0.399	0.273	0.287	0.368	0.509	0.409	0.666

Na osnovu analize rezultata klasterovanja može se izneti nekoliko zaključaka. Pripadnike prvog klastera odlikuju loše statistike vezane za servis (A1, A2, A3) i nizak procenat spašenih brejk lopti (A7), dok je procenat osvojenih brejk lopti visok (A8), kao što su Švarcman, Ferer, Davidenko, Nišikori. Tenisere iz trećeg klastera, suprotno teniserima iz prvog klastera, odlikuju visoki procenti kada su u pitanju atributi vezani za servis (A1, A2, A3), samim tim i visok procenat osvojenih gemova po meču (A5), visok procenat spašenih brejk lopti (A7) i viši procenat duplih grešaka (A4), dok im je procenat osvojenih brejk lopti (A8) znatno lošiji poput Karlovića, Iznera, Raonića, Samprasa, Federera, itd. Pripadnike drugog klastera odlikuju visoke vrednosti kada su u pitanju procenat osvojenih poena na drugi servis (A2), osvojenih tajbrejkova (A6), sačuvanih brejk lopti (A7) i osvojenih brejk lopti (A8).



**Slika 1:** Prikaz klastera

Nakon detaljnije analize predstavnika klastera, Švarcman (klaster 1 – „Never surrenders“), Đokovića i Nadala (klaster 2 – „Mental rocks“) i Karlovića (klaster 3 – „Top servers“), izvedeno je nekoliko zaključaka. Teniseri iz prvog klastera su uglavnom ljudi nižeg rasta, u odnosu na tenisere iz drugog i trećeg klastera, koji čine mahom visoki teniseri. U skladu sa tim, teniseri iz prvog klastera gube dosta gemova na svoj servis, ali isto tako i prave dosta brejkova, suprotno teniserima iz trećeg klastera, koji svoju igru baziraju na servisu. Pripadnici drugog klastera, u poređenju sa ostala dva klastera, uzimaju najviše poena na svoj drugi servis,



procenat poena na prvi servis je manji od tenisera iz trećeg, ali veći od tenisera iz prvog klastera. Neki od ključnih momenata meča jesu trenuci kada se osvajaju ili brane brejk lopte, kao i igranje samih tajbrejkova. U skladu sa tim, može se zaključiti da teniseri iz drugog klastera imaju dobre ocene kod glavnih atributa za dobijanje meča. Kada pogledamo činjenicu da drugom klasteru pripadaju Đoković, teniser sa najviše nedelja na prvoj poziciji svetske ATP liste i Nadal, teniser sa najviše grend slem titula, potvrđujemo ovu grupu klastera.

Na slici klastera možemo uočiti i neke granične slučajeve. Neki od njih su na primer Grigor Dimitrov, teniser koji se nalazi na granici između drugog i trećeg klastera i Mikael Ljodra koji se nalazi na granici između prvog i drugog klastera. Dimitrov ima visok procenat osvojenih poena na prvi servis (A1) i visok procenat spašenih brejk lopti (A7), koje su odlike trećeg klastera. Dimitrov takođe ima i visok procenat osvojenih poena na drugi servis (A2) što je najviše odlika drugog klastera, dok su mu vrednosti ostalih atributa vezanih za drugi klaster niske. Mikael Ljodra takođe ima visok procenat osvojenih poena na drugi servis (A2) kao odliku drugog klastera, a kao odliku prvog klastera ima visok procenat osvojenih brejk lopti (A8).

U analiziranje je uključen i Rodžer Federer, kao preostali član velike trojke u svetu tenisa. Federer pripada trećem klasteru, ali je jako blizu drugog klastera (Tabela 3). Njegove statistike za servis su dobre, ali isto tako i ocene atributa koje karakterišu pripadnike drugog klastera. Imajući to u vidu, jasno je zašto je smatran za jednog od najvećih tenisera svih vremena. Pripadnost klasterima velike trojke prikazane su u Tabeli 3.

**Tabela 3:** Pripadnost klasterima Đokovića, Nadala i Federera

Pripadnosti klasteru	Klaster 1	Klaster 2	Klaster 3
R. Federer	0.230	0.333	0.436
R. Nadal	0.278	0.382	0.340
N. Đoković	0.309	0.369	0.322

## 5. ZAKLJUČAK

Osnovni cilj ovog rada bio je da odgovori na pitanja: kako grupisati tenisere i koji atributi su najznačajniji za uspešnost tenisera. Na osnovu statistika iz mečeva, primenom algoritma fazi C-srednjih vrednosti izvršeno je klasterovanje 133 teniska igrača. Kao rezultat dobijena su tri klastera koji su u nastavku identifikovani prema stilu igre: „Never surrenders“, „Mental rocks“ i „Top servers“. Time je primarni cilj istraživanja postignut.

Na osnovu pojedinih svetski poznatih kriterijuma, kao što su broj nedelja na prvom mestu i broj grend slem titula, izdvojeni su najbolji i na osnovu njih upoređene bitnosti i karakteristike klastera. U skladu sa informacijama koje smo dobili, može se primetiti da atributi koji karakterišu pripadnike drugog, pa onda i atributi trećeg klastera u kombinaciji sa atributima iz drugog klastera, mogu uzeti za najbitnije u teniskom svetu, jer se u njima nalaze najbolji svih vremena. Zahvaljujući algoritmu fazi klasterovanja, dobijeni su rezultati tenisera koji imaju značajnu pripadnost u više klastera i vrednosti atributa koji ih opisuju su posebno zanimljive za posmatranje.

Mogućnost za unapređenje istraživanja se ogleda u dodavanju atributa koje bi se postiglo uz širenje skupa podataka sa dodatnim statistikama kao što su broj direktnih poena iz forhenda i bekhenda, prosečno trajanje poena, broj grešaka i slično. Takođe, moglo bi se izvršiti poređenje u zavisnosti od podloge na kojoj se igra neki turnir, a pored toga dalji nastavak istraživanja može ići i u smeru primene drugih algoritama klasterovanja.

## LITERATURA

- [1] Askari, S. (2021). Fuzzy C-Means clustering algorithm for data with unequal cluster sizes and contaminated with noise and outliers: Review and development. *Expert Systems with Applications*, 165, 113856.
- [2] Bezdek, J. C. (2017). A primer on cluster analysis: 4 basic methods that (usually) work. First Edition Design Publishing.
- [3] Bonidia, R. P., Rodrigues, L. A., Avila-Santos, A. P., Sanches, D. S., & Brancher, J. D. (2018). Computational intelligence in sports: A systematic literature review. *Advances in Human-Computer Interaction*, 2018.
- [4] De Oliveira, J. V., & Pedrycz, W. (Eds.). (2007). *Advances in fuzzy clustering and its applications*. John Wiley & Sons.
- [5] Erilli, N. A., Öner, Y., Alakuş, K., & Tunç, T. (2010). Classifying Chess Players with Fuzzy Clustering Analysis in Fuzzy Data Using Eco Codes. In *1st International Symposium on Computing in Science & Engineering Proceedings Book* (pp. 1105-1110).

- 
- [6] Fister Jr, I., Ljubič, K., Suganthan, P. N., Perc, M., & Fister, I. (2015). Computational intelligence in sports: challenges and opportunities within a new research domain. *Applied Mathematics and Computation*, 262, 178-186.
- [7] Furlong, J.D.G. (1995). The service in lawn tennis: how important is it?
- [8] Garcia-Gonzalez, L., Diminguez, A.M., Gil, A. & Moreno, M.P. (2014). Effect of decisions making and performance young tennis players: an applied research
- [9] Govindasamy, K., & Velmurugan, T. (2018). Analysis of student academic performance using clustering techniques. *International Journal of Pure and Applied Mathematics*, 119(15), 309-323.
- [10] Murray, N. P., & Hunfalvay, M. (2017). A comparison of visual search strategies of elite and non-elite tennis players through cluster analysis. *Journal of sports sciences*, 35(3), 241-246.
- [11] Næs, T., & Mevik, B. H. (1999). The flexibility of fuzzy clustering illustrated by examples. *Journal of Chemometrics*, 13(3-4), 435-444.
- [12] Ohguni, M., Aoki, M., Sato, H. & Imada, K. (2009). The effect of grip size on the hitting force during a soft tennis forehand stroke.
- [13] Ramamoorthy, V. (2019). Fuzzy C-Mean Clustering Using Data Mining. BookRix: Munich, Germany.
- [14] Reid, M. & Morris, C. (2013). ranking benchmarks of top 100 players in men's professional tennis.
- [15] Sampaio, J., McGarry, T., Calleja-González, J., Jiménez Sáiz, S., Schelling i del Alcázar, X., & Balciunas, M. (2015). Exploring game performance in the National Basketball Association using player tracking data. *PloS one*, 10(7), e0132894.
- [16] Shelly, Z., Burch, R. F., Tian, W., Strawderman, L., Piroli, A., & Bichey, C. (2020). Using K-means clustering to create training groups for elite American football student-athletes based on game demands. *International Journal of Kinesiology and Sports Science*, 8(2), 47-63.
- [17] Syed, A. M. (2019). Performance Attributes and Risk Taking of Players-Evidence from Lawn Tennis. *Journal of Sport and Games 2019*, 1(3), 13-20.
- [18] Vučičević, A., Milošević, P., Poledica, A. (2021). Selecting Case Study members using fuzzy C-means algorithm. *Proc. XLVIII International Symposium on Operational Research, SYMOPIS 2021*. Banja Koviljača, Sept. 20-23, 2021.

## KLASIFIKACIJA NEŽELJENIH SMS PORUKA KORIŠĆENJEM RAČUNARSKE INTELIGENCIJE

### SMS SPAM CLASSIFICATION USING COMPUTATIONAL INTELLIGENCE

DANILO JANJUŠEVIĆ<sup>1</sup>, IVANA DRAGOVIĆ<sup>2</sup>, PAVLE MILOŠEVIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, danilo.janj97@gmail.com

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, ivana.dragovic@fon.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, pavle.milosevic@fon.bg.ac.rs

**Rezime:** *Nepoželjne (spam) poruke deo su komunikacije u digitalnom prostoru od ranih početaka elektronskih sistema za slanje tekstualnog sadržaja. Za veću bezbednost i bolji doživljaj korisnika od značaja je kreiranje sistema za njihovo filtriranje. U ovom radu predložen je jedan takav sistem koji ima za cilj ispravno klasifikovanje SMS poruka koje predstavljaju spam. U tu svrhu korišćen je javno dostupan skup obeleženih SMS poruka napisanih na engleskom jeziku. Posebna pažnja je poklonjena pripremi i normalizaciji elemenata tih poruka. Kroz transformaciju poruka u vektore prikazana su dva pristupa vektorizaciji tekstualnog sadržaja, jedan klasičan TF-IDF model obrade prirodnih jezika i drugi moderniji Word2Vec koji se zasniva na očuvanju semantičke vrednosti reči. Nad vektorima dobijenim pomoću oba pristupa trenirani su algoritmi mašinskog učenja, i to metoda potpornih vektora, neuronske mreže i metoda k najbližih suseda. Postignuti su zavidni rezultati i pokazano je da za klasifikaciju nije od presudnog značaja semantička vrednost reči.*

**Ključne reči:** *SMS, spam, računarska inteligencija, obrada prirodnih jezika*

**Abstract:** *Unsolicited (spam) messages have been a part of the communication in digital space since the early beginnings of electronic systems for text messaging. For greater security and better customer experience the creation of spam filtering systems is of crucial importance. In this paper we propose one such system whose objective is to correctly classify SMS spam messages. For that purpose we used publicly available dataset with labeled SMS messages written in English. Special attention was given to the preparation and normalization of the elements of those messages. Through the transformation of SMS messages into vectors, we presented two approaches for textual content vectorization, one was the classical TF-IDF natural language processing model, and the other one was the more modern Word2Vec model which is based on the preservation of the semantic value of the word. On derived vectors we trained the machine learning algorithms, and those were support vector machines, neural networks and k-nearest neighbours. We achieved outstanding results and it is shown that the semantic value of words is not of crucial importance for the problem of classification.*

**Keywords:** *SMS, spam, computational intelligence, natural language processing*

## 1. UVOD

Internet spam, po definiciji, predstavlja jednu ili više neželjenih poruka, poslatih ili postavljenih putem elektronskih sistema, a koje su deo veće kolekcije poruka čiji je sadržaj u značajnoj meri identičan [3]. Kao takav, spam može, ali i ne mora, biti ilegalnog ili po korisnika opasnog karaktera. Zbog pretežno negativnog uticaja koji neželjene poruke mogu imati javila se potreba za pravljjenjem sistema za filtriranje ovakvih poruka. Takvi sistemi se danas oslanjaju na složene algoritme mašinskog učenja i njihovo razvijanje i usavršavanje predstavlja izazov kako za industriju, tako i za akademsku zajednicu.

SMS spam samo je jedna od potkategorija spama. U ovom radu dat je detaljan prikaz razvijanja sistema za filtriranje neželjenih SMS poruka korišćenjem nekoliko popularnih algoritama mašinskog učenja i kroz više eksperimenata urađenih u programskom jeziku *Python*. Posebna pažnja u radu je data pripremi podataka kao izuzetno važnom koraku pre punjenja algoritama ulaznim podacima. Pripremi podataka i eksperimentima je prethodilo opsežno istraživanje klasičnih i savremenih načina obrade prirodnih jezika kao

grane računarske inteligencije, te je u skladu sa tim u radu izložen i kraći teorijski osvrt na ovu temu. Rad je zaključen uporednom analizom metrika koje su izračunate za svaki od korišćenih algoritama.

Rad je organizovan na sledeći način. U drugom poglavlju opisan je konkretan problem kojim se rad bavi i dat je kratak pregled literature. U trećem poglavlju opisana je metodologija primenjena u ovom radu. Četvrto poglavlje daje kratak opis skupa podataka. Priprema podataka i vektorizacija opisani su u petom poglavlju. Šesto poglavlje odnosilo se na prikaz eksperimenata i rezultate. U poslednjem, sedmom poglavlju, iznet je zaključak ovog rada.

## 2. PROBLEM KLASIFIKACIJE SMS PORUKA

U opštem, najširem smislu klasifikacija SMS poruka potpada pod problem klasifikacije teksta. Automatska klasifikacija teksta podrazumeva dodeljivanje tekstualnih dokumenata predefinisanim skupima klasa korišćenjem tehnika mašinskog učenja [5]. Kao takav ovaj problem je problem nadgledanog učenja. Ono što ga izdvaja od drugih problema klasifikacije jeste to što su podaci sa kojima računar radi čitave rečenične konstrukcije. Zato je ovaj problem usko povezan i sa problemom obrade prirodnih jezika.

Obrada prirodnih jezika (eng. *natural language processing, NLP*) je oblast računarskih nauka, računarske inteligencije i računarske lingvistike koja se bavi interakcijama između računara i jezika ljudi [10]. Ona obuhvata kako algoritme koji prihvataju tekst napisan od strane ljudi, tako i algoritme koji proizvode tekst za svoj izlaz [6]. Istraživanja u ovoj kompleksnoj oblasti počela su još 1950-ih godina i to kao presek računarske inteligencije i lingvistike, ali tek 1980-ih dolazi do prodora moderne statističke obrade prirodnih jezika [8]. Danas se upravo takva obrada prirodnih jezika zasnovana na statistici oslanja na modele mašinskog i dubokog učenja u rešavanju velikog broja problema, a gorepomenuta klasifikacija teksta samo je jedna od oblasti u kojoj je obrada prirodnih jezika našla primenu.

Dobra polazna tačka pri analizi referentne literature na temu klasifikacije teksta je [1]. U ovom radu predstavljeni su koraci neophodni za klasifikaciju teksta, a koji uključuju pretprocesiranje teksta, modelovanje ulaza, smanjenje dimenzija, treniranje modela i evaluaciju. U našem radu vodili smo se upravo koracima opisanim u ovom radu kao polaznom osnovom.

Filtriranje tekstualnih poruka, kao specifičan potproblem klasifikacije teksta, bilo je predmet velikog broja radova u prethodnom periodu. Ti radovi bavili su se klasifikacijom poruka u digitalnom prostoru, načinima na koji se filtriranje može poboljšati i izazovima koji leže u budućnosti na tom polju. Tako je u [2] na skupu SMS poruka koje su prikupljene, a koji smo i mi za potrebe ovog rada koristili, pokazano da je metoda potpornih vektora algoritam koji najbolje klasifikuje spam poruke, te se u tom radu zbog toga i predlaže kao dobra početna tačka za dalja istraživanja. Slično ovom radu i u [9] iznose se rezultati u kojima je korišćeno više algoritama za klasifikaciju elektronske pošte. Po tom istraživanju kao najbolji algoritam se ponovo pokazala metoda potpornih vektora. Ovaj algoritam dao je najvišu tačnost i najmanju stopu lažno pozitivnih rezultata, iako se pokazao kao najsporiji u pogledu treniranja.

## 3. METODOLOGIJA

U narednim pasusima dat je kratak teorijski osvrt na algoritme mašinskog učenja koji su korišćeni u ovom radu u svrhu klasifikovanja SMS poruka. Kao klasifikatori korišćeni su metoda potpornih vektora, neuronska mreža (višeslojni perceptron) i metoda k najbližih suseda.

Metoda potpornih vektora (eng. *Support Vector Machine, SVM*) jedna je od popularnijih tehnika mašinskog učenja i koristi se za različite probleme klasifikacije, regresije i detekcije anomalija. Ova metoda se zasniva na pronalaženju najbolje granice odlučivanja tj. hiperravni koja će odvojiti tačke koje pripadaju različitim klasama, a pod time se podrazumeva ona granica odlučivanja koja je najudaljenija od tačaka najbližih toj hiperravni [4]. Te tačke se nazivaju potporni vektori i oni su tačke koje se koriste prilikom treniranja modela, a njihova udaljenost od hiperravni naziva se margina čija dužina obezbeđuje prostor u kom neće doći do pogrešne klasifikacije novih tačaka. Korišćenjem funkcije jezgra (eng. *kernel*) u višedimenzionalnom prostoru moguće je ovu metodu koristiti i za probleme koji nisu linearno separabilni.

Veštačke neuronske mreže (eng. *Artificial Neural Networks, ANN*) drugi su algoritam korišćen u ovom radu. Ovaj algoritam modelovan je po ugledu na ljudske nervne ćelije gde svaka ćelija prima ulazne signale, obradi ih i šalje ka drugim nervnim ćelijama, tj. neuronima. Svaki neuron implementiran kroz softver poseduje nelinearnu aktivacionu funkciju koja obezbeđuje da granica odlučivanja prilikom klasifikovanja bude nelinearna, te na taj način arhitekture sa više slojeva mogu da aproksimiraju skoro svaku funkciju i da reše mnoge nelinearno separabilne probleme [7]. Parametri se optimizuju kroz algoritam gradijentnog spusta ili neki od algoritama nastalih po uzoru na gradijentni spust gde vektor svih parametara menja vrednosti u

smeru u kom funkcija troška najbrže pada, a ceo proces dobijanja izvoda za gradijent se vrši kroz propagaciju unazad. U prethodnoj deceniji neuronske mreže su doživele proboj u mnogim oblastima, te predstavljaju osnovni gradivni element za duboko učenje. Ipak, za potrebe ovog rada smatrali smo da nije bilo potrebno koristiti duboke neuronske mreže već smo koristili običan višeslojni perceptron (eng. *Multilayer Perceptron, MLP*). Ovaj algoritam, bez slojeva konvolucije, dao je sasvim zadovoljavajuće rezultate u eksperimentima koji su izvršeni.

Poslednji algoritam koji je korišćen bio je kNN, odnosno k najbližih suseda (eng. *k-nearest neighbours*). Ovaj algoritam je zapravo algoritam tzv. lenjog učenja i zasniva se na određivanju klase po tome koja je klasa dominantna među k tačaka najbližih posmatranoj tački, pri čemu se bliskost može definisati na više načina, a najčešće pomoću Euklidskog rastojanja.

#### 4. OPIS SKUPA PODATAKA

Skup podataka korišćen za potrebe ovog rada jeste javno dostupni skup realnih i neobrađenih SMS poruka napisanih na engleskom jeziku, a koje su prikupljene u [2].

Eksplorativna analiza ovog skupa podataka dala je nekoliko interesantnih zapažanja. U osnovi, neobrađeni skup podataka ima svega dve kolone - jednu koja za vrednosti uzima *ham* (za one poruke koje su validne) i *spam* (za sve poruke koje su označene kao nepoželjne od strane tima koji je poruke prikupio), i drugu kolonu koja za vrednost uzima promenljivi broj karaktera koji čine SMS poruku. Upravo ta promenljivost broja karaktera za atribut SMS poruke zahteva opsežnu pripremu i vektorizaciju koja je obavljena.

Analiza je dalje pokazala da skup sadrži 5572 poruka, od čega je 747 poruka obeleženo kao spam. Među spam porukama postoje 653 jedinstvene poruke, dok je među validnim porukama 4516 jedinstvenih. Ovakav odnos ukazuje na relativno visoku nebalansiranost skupa podataka, što je i očekivano. Ipak, i pored toga, za potrebe ovog rada se smatra da spam poruke ne predstavljaju anomalije jer ih ima dovoljno za učenje modela, te se problem filtriranja takvih poruka u ovom radu posmatra kao problem nadgledanog učenja i klasifikacije, a ne kao problem detekcije anomalija.

Za potrebe dalje analize kreiran je novi atribut koji se odnosi na dužinu SMS poruke. Eksplorativna analiza nad ovim atributom pokazala je određene statistike, te se tako moglo videti da je srednja vrednost dužine poruka na celom skupu podataka 80.12, da je standardna devijacija 59.69, najmanji broj karaktera koje neka poruka ima je 2, a najviše čak 910 karaktera. Takođe, utvrđeno je da SMS poruke koje spadaju u spam klasu imaju tendenciju da sadrže veći broj karaktera nego obične SMS poruke, što bi za određene manje sofisticirane sisteme bio dovonjan indikator za filtriranje. Ipak, radi kreiranja inteligentnih sistema bilo je potrebno od SMS poruka napraviti vektore za ulaze u takve sisteme.

#### 5. PRIPREMA PODATAKA - VEKTORIZACIJA I NORMALIZACIJA

Kako je već primećeno, tekstualni sadržaj dokumenta poput SMS poruka nije moguće lako i direktno pretvoriti u vektor u  $n$ -dimenzionalnom prostoru zbog promenljivosti sadržaja tih dokumenata - broj karaktera stoga ne može biti broj dimenzija prostora u kom se vektor nalazi. Zbog toga se mora pristupiti vektorizaciji teksta na drugačiji način.

U ovom radu za prvi niz eksperimenata sledili su se principi vektorizacije SMS poruka pomoću modela vreće reči (eng. *bag of words*) kao prostog, ali moćnog modela obrade prirodnih jezika. U ovom modelu za attribute se uzimaju reči iz dokumenata, ali njihova pozicija u rečenicama, semantika i gramatika nisu od značaja već je bitan broj koliko se određena reč pojavljuje u dokumentu. Tako se svaka reč koja se pojavljuje mapira u učestalost te reči u dokumentu. Kako je tema ovoga rada klasifikacija spam SMS poruka, smatramo da ovaj pristup može biti dobar - klasifikatoru je bitno da nauči šta je dominantna tema onih poruka koje nisu poželjne, te da ih na osnovu toga odvoji od drugih poruka. Dominantnu temu klasifikator u ovom slučaju prepoznaje preko najučestalijih reči koje se nalaze u SMS porukama označenim kao spam (reči kao što su novac, nagrada, ponuda i sl. mogu se češće nalaziti u nepoželjnim porukama nego u poželjnim). Kako bi vektor imao nepromenljiv broj atributa za sve SMS poruke, pre određivanja učestalosti reči potrebno je sačiniti rečnik svih reči koje se pojavljuju u SMS porukama, a zatim svaku reč iz tog rečnika (ili njegove skraćene verzije) za svaku SMS poruku mapirati u broj pojavljivanja te reči u datoj SMS poruci, te tako mapirane vrednosti koristiti kao vektor ulaza za modele mašinskog učenja.

Da bi rečnik bio kvalitetan potrebno je pre njegovog pravljenja odraditi opsežnu pripremu podataka. Za potrebe ovog rada ta priprema je podrazumevala nekoliko koraka.

U prvom koraku su sve SMS poruke transformisane da budu u formatu malih slova. Na taj način se prilikom mapiranja neće praviti razlika između reči koje imaju u sebi mala i velika slova već će se te reči posmatrati kao identične, što u govornom jeziku i jesu.

U drugom koraku koji se često radi pri ovakvim problemima iz svih SMS poruka su izbačene tzv. zaustavne reči (eng. *stop words*). To su najčešće korišćene reči u jeziku i za engleski jezik su to lične zamenice, pomoćni glagoli i njihove negacije, predlozi, priloz, gramatički članovi i sl. Taj korak je neophodan zato što takve reči nisu relevantne za odlučivanje, a takođe bi mogle da odnesu uvek prevagu u odnosu na ostale reči.

Potom su u SMS porukama svi veb linkovi i mejl adrese zamenjeni predefinisanim konstrukcijama. Inicijalna pretpostavka je da veliki broj spam poruka u stvarnom svetu sadrži veb linkove i mejlove sa ciljem agresivne promotivne kampanje, širenja računarskih virusa, krađe identiteta, promovisanja sadržaja neprikladnog za određene uzraste itd. Zato je ovaj korak od izuzetne važnosti za pripremu podataka kako bi se mogli takvi linkovi identifikovati i kako bi se na osnovu njihove učestalosti donela odluka o klasifikaciji. Slične zamene standardizovanim konstrukcijama su urađene i za brojeve i valute.

Poslednji korak u pripremi podataka, a pre modelovanja ulaza, bio je stemovanje reči, odnosno svođenje reči na korensku osnovu (eng. *stemming*). Ovo je urađeno kako klasifikator ne bi pravio razliku između reči koje u svojoj suštini imaju istu semantičku vrednost, ali se razlikuju gramatički. Na taj način je omogućeno računaru da prepozna različite forme jedne iste reči ako je ona dominantna u nekoj SMS poruci, te da prema njoj i drugim rečima iz SMS-a donese odluku o klasifikaciji. Stemovanje je izvršeno pomoću algoritma Porter za stemovanje, a realizovano korišćenjem klase *PorterStemmer* iz *nlTK* biblioteke programskog jezika *Python*.

Nakon što su konačno pripremljeni i normalizovani podaci, pristupilo se formiranju vektora ulaza. Poruke su prvo podeljene na skup poruka predviđenih za treniranje modela koje čine 80% originalnog skupa SMS poruka i na skup poruka predviđenih za testiranje istreniranih modela, a koje čine 20% prvobitnog skupa. Zatim se izvršilo kreiranje modela vreće reči gde se prvo uči rečnik svih reči iz SMS poruka iz skupa za treniranje, a zatim se svakoj SMS poruci dodeljuje tačno jedan vektor u  $n$ -dimenzionalnom prostoru gde je  $n$  ukupan broj reči u rečniku pripremljenih reči.

Ipak, ovakav pristup vektorizaciji je moguće poboljšati pomoću statističke metode učestalosti termina-inverzne učestalosti dokumenta (eng. *term frequency-inverse document frequency*, *TF-IDF*). Ova metoda ne posmatra samo učestalost reči u samom dokumentu, već nastoji da težinama kazni one termine koji se nalaze prečesto u čitavom korpusu dokumenata, a koji nisu relevantni, i obrnuto, stoga smo u našem radu model vreće reči nadogradili ovom statističkom metodom.

Ovim su svi koraci vezani za pripremu podataka završeni i moglo se pristupiti kreiranju i treniranju modela mašinskog učenja. Prvi niz eksperimenata rađen je upravo sa ovako vektorizovanim porukama, dok je drugi niz eksperimenata rađen korišćenjem *Word2Vec* neuronske mreže, o čemu će biti reči više u narednim pasusima.

## 6. EKSPERIMENTI I REZULTATI

Najpre su algoritmi mašinskog učenja primenjeni nad *TF-IDF* vektorima. Mere performansi korišćene za evaluaciju i poređenje rezultata eksperimenata su tačnost (eng. *accuracy*), koja se odnosi na procenat ispravno predviđenih opservacija, i *F1* mera koja predstavlja harmonijsku sredinu preciznosti (eng. *precision*) kao udela ispravno predviđenih pozitivnih opservacija u ukupnom broju opservacija koje je algoritam klasifikovao kao pozitivne, i odziva (eng. *recall*) koji predstavlja udeo ispravno predviđenih pozitivnih opservacija u ukupnom broju pozitivnih opservacija. Prvi klasifikator koji je korišćen bila je metoda potpornih vektora. Za ovaj algoritam su varirani hiperparametar regularizacije  $C$  i hiperparametar zakrivljenosti granice odlučivanja  $\gamma$ . Bilo je potrebno odrediti koja kombinacija hiperparametara za metodu potpornih vektora je optimalna, pa je urađena unakrsna validacija. Od 12 različitih kombinacija najbolje se pokazala kombinacija u kojoj je hiperparametar regularizacije bio 100, a  $\gamma$  0.01 i za tu kombinaciju hiperparametara na skupu za testiranje pokazana je tačnost od 99.1% i *F1*-skor od 99.49%.

Slično, za višeslojni perceptron rađena je unakrsna validacija za hiperparametre aktivacione funkcije, algoritme optimizacije i broj neurona u skrivenim slojevima. Najbolje se pokazala arhitektura sa jednim slojem od 10 neurona, adaptivnom estimacijom momenta kao optimizatorom i rektifikovanom linearnom jedinicom kao aktivacionom funkcijom. Za ovu kombinaciju hiperparametara postignuta je tačnost od 99.28% na skupu za testiranje i *F1*-skor od 99.59%.

Poslednji algoritam za koji su merene metrike nad ovako vektorizovanim podacima bio je algoritam  $k$  najbližih suseda. Pri primeni ovog algoritma variran je broj suseda od 3 do 29 sa korakom 2. Iako algoritam

lenjog učenja, i ovaj algoritam je pokazao zavidne rezultate - najbolje metrike dao je za 15 najbližih suseda, te je na skupu za testiranje za taj broj postignuta tačnost od 98.02% i F1-skor od 98.88%.

Do sada smo pri eksperimentisanju koristili TF-IDF model vektorizacije dokumenata koji se, za potrebe ovog rada, pokazao i više nego dobar. Međutim, ovaj model zanemaruje semantičku vrednost reči i ne može se koristiti za neke složenije probleme obrade prirodnih jezika. Iako je prvi pristup postigao dobre rezultate, odlučili smo da bez obzira na to probamo iste eksperimente i nad porukama vektorizovanim pomoću Word2Vec tehnike.

Word2Vec je tehnika ugrađivanja reči u vektorski prostor (eng. *word embeddings*) uz pomoć treniranja istoimene neuronske mreže. Ova metoda zasniva se na ideji da reči koje imaju slično značenje često dele i sličan kontekst i okruženje u kojem se nalaze. Kako Word2Vec neuronska mreža za treniranje zahteva ogroman broj reči, a naš skup podataka nema dovoljno bogat rečnik, za potrebe ovog rada korišćena je pretrenirana Word2Vec mreža koja je istrenirana nad 100 milijardi reči iz Google News skupa podataka i koja za svaku reč iz tog skupa (a tako sigurno i za veliki broj reči u SMS porukama) daje vektor u 300-dimenzionalnom prostoru. Svaka SMS poruka pretvorena je u listu pojedinačnih reči koje su potom pretvorene u Word2Vec vektore, a zatim su ti vektori sabrani zajedno u vektor koji predstavlja jednu SMS poruku.

Za ovako vektorizovane SMS poruke rezultati eksperimenata su bili sledeći nakon optimizovanja hiperparametara. Za metodu potpornih vektora pronađeni su optimalni hiperparametar regularizacije koji iznosi 10 i hiperparametar gama koji iznosi 0.01 za Gausovu kernel funkciju, i za njih su dobijeni rezultati tačnosti 98.92% i F1-skor od 99.39% na skupu za testiranje. Za višeslojni perceptron ponovo je bio dovoljan jedan sloj od 10 neurona sa adam algoritmom optimizacije, ali se najbolje pokazala ovaj put aktivaciona funkcija tangens hiperbolički, te je ova kombinacija na test skupu pokazala tačnost od 98.65% i F1-skor od 99.23%. Kada je reč o kNN algoritmu njegov optimalni hiperparametar bio je 3 najbliža suseda i za njega je na skupu za testiranje postignuta tačnost od 97.4% i F1-skor od 98.53%. Primećeno je da su svi algoritmi znatno brže istrenirani nego u TF-IDF modelu, pri čemu je vreme treniranja bilo i do 10 puta manje nego prvobitno, što je i logično jer je vektor ulaza smanjen sa prvobitnih 5711 dimenzija na svega 300 atributa.

U Tabeli 1 može se videti uporedni prikaz svih dobijenih rezultata.

**Tabela 1:** Tabelarni prikaz performansi za sve korišćene algoritme i za sve primenjene metode vektorizacije

	metoda potpornih vektora		neuronska mreža		kNN	
	tačnost	F1-skor	tačnost	F1-skor	tačnost	F1-skor
<b>TF-IDF</b>	99.1%	99.49%	99.28%	99.59%	98.02%	98.88%
<b>Word2Vec</b>	98.92%	99.39%	98.65%	99.23%	97.4%	98.53%

Kao što se iz tabele može primetiti najbolje rezultate dao je višeslojni perceptron i to za SMS poruke koje su vektorizovane pomoću TF-IDF principa. Ono što je bitno primetiti jeste da su i ostali rezultati samo neznatno slabiji, kao i to da, iako je korišćen pretrenirani Word2Vec model to nije uticalo u velikoj meri na performanse, štaviše dobijeni su nešto slabiji rezultati za takav pristup. Dobijeni rezultati nadmašuju rezultate objavljene u [2] gde je primenom metode potpornih vektora ostvarena tačnost od 97.64% nad istim skupom poruka, a bez prethodnog stemovanja, brisanja zaustavnih reči i identifikovanja brojeva, valuta, mejlova i veb adresa.

Svega nekoliko poruka bilo je pogrešno klasifikovano primenom neuronske mreže, a kao primer pogrešno klasifikovane poruke navodimo poruku koja se može videti na Slici 1. Naša pretpostavka je da je neuronska mreža pogrešno klasifikovala ovu poruku kao spam zbog prisustva cifara, reči koja označava trošak ili potencijalno zbog ponavljanja jedne iste reči nekoliko puta.

'1.20 that call cost. Which i guess isnt bad. Miss ya, need ya, want ya, love ya'

**Slika 1:** Prikaz pogrešno klasifikovane poruke

## 7. ZAKLJUČAK

Cilj ovog rada bilo je istraživanje načina za filtriranje SMS poruka putem računarske inteligencije. Ovo istraživanje obuhvatilo je kako načine za pripremu, normalizaciju i vektorizaciju SMS poruka kao tekstualnih ulaza, tako i podešavanje hiperparametara samih klasifikatora SMS poruka. U radu smo došli do zaključka da je za ovakav problem dovoljno koristiti klasične modele računarske obrade prirodnih jezika, poput vreće reči i TF-IDF modela, te da za problem klasifikacije SMS poruka nije od presudne važnosti semantička vrednost samih reči koja se zadržava Word2Vec pristupom. Takođe, pokazano je da su klasični algoritmi

mašinskog učenja u osnovi dobri klasifikatori za ovakav problem. Kako su se modeli crne kutije, poput neuronskih mreža, pokazali kao sasvim prihvatljivi, zaključujemo da za ovaj problem nisu potrebni algoritmi koji imaju mogućnost detaljnog tumačenja. Tokom izvođenja eksperimenata primećeno je da se unakrsna validacija izvršava znatno brže za SMS poruke vektorizovane Word2Vec pristupom zbog smanjenja dimenzionalnosti u odnosu na TF-IDF, stoga se ovaj pristup predlaže ukoliko su pri rešavanju problema vreme i računarski resursi ključni faktori. Smatramo da postoji potencijal za dalja poboljšanja Word2Vec pristupa, te da bi sa nekim većim skupom SMS poruka nad kojim bi se eventualno istrenirao Word2Vec ovo rešenje dalo još bolje rezultate. Takođe, potencijal postoji i u načinu na koji se vrši agregacija vektora reči u drugom pristupu, gde bi neki sofisticiraniji način od prostog sabiranja vektora mogao da pruži bolje performanse za testirane modele.

Kako su SMS poruke u ovom radu bile napisane na engleskom jeziku, smatramo da postoji potencijal u budućnosti i za slično istraživanje vezano za poruke napisane na srpskom jeziku. Trenutna potencijalna ograničenja za istraživanja u tom pravcu nalaze se u tome što je mašinska obrada srpskog jezika kao prirodnog jezika u začetku, te to što su pojedine biblioteke korišćene u ovom radu prilagođene isključivo engleskom jeziku. Za treniranje modela poput Word2Vec neuronske mreže potrebne su i velike količine podataka, pa izazovi leže i u prikupljanju tekstualnog sadržaja na srpskom jeziku, a zatim i u treniranju modela nad tim sadržajem radi dobijanja vektorskih reprezentacija reči.

## LITERATURA

- [1] Aas, K., & Eikvil, L. (1999). Text categorisation: A survey (p. 3). Technical report, Norwegian computing center.
- [2] Almeida, T. A., Hidalgo, J. M. G., & Yamakami, A. (2011, September). Contributions to the study of SMS spam filtering: new collection and results. In *Proceedings of the 11th ACM symposium on Document engineering* (pp. 259-262).
- [3] Blanzieri, E., & Bryl, A. (2008). A survey of learning-based techniques of email spam filtering. *Artificial Intelligence Review*, 29(1), 63-92.
- [4] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.
- [5] Dalal, M. K., & Zaveri, M. A. (2011). Automatic text classification: a technical review. *International Journal of Computer Applications*, 28(2), 37-40.
- [6] Goldberg, Y., & Hirst, G. (2017). *Neural network methods in natural language processing. Synthesis Lectures on Human Language Technologies (2017)*. Morgan & Claypool Publishers. DOI: 10.2200/S00762ED1V01Y201703HLT037
- [7] Hagan, M., Demuth, H., Beale, M., & DeJesus O. (2014). *Neural Network Design (2nd ed.)*. Martin Hagan.
- [8] Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18(5), 544-551
- [9] Trivedi, M., Sharma, S., Soni, N., & Nair, S. (2015). Comparison of text classification algorithms. *International Journal of Engineering Research & Technology (IJERT)*, 4(02).
- [10] Reshamwala, A., Mishra, D., & Pawar, P. (2013). Review on natural language processing. *IRACST Engineering Science and Technology: An International Journal*, 3(1), 113-116



## PREPOZNAVANJE TALENTOVANIH UČENIKA OSNOVNE ŠKOLE PRIMENOM NEURONSKIH MREŽA I FAZI LOGIKE

### TALENT RECOGNITION IN ELEMENTARY SCHOOL STUDENTS USING NEURAL NETWORKS AND FUZZY LOGIC

NATALIJA DAMJANOVIĆ<sup>1</sup>, IVANA DRAGOVIĆ<sup>2</sup>, PAVLE MILOŠEVIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, natalija.damjanovic.98@gmail.com

<sup>2</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, ivana.dragovic@fon.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, pavle.milosevic@fon.bg.ac.rs

**Rezime:** Rano prepoznavanje talentovanih učenika od velike je važnosti za njihovo dalje usmeravanje i obrazovanje. Pitanje šta se tačno podrazumeva pod pojmom „talentovanost“ je otvoreno. U ovom radu ispituje se mogućnost primene veštačkih neuronskih mreža, fazi logike i neuro-fazi sistema u oblasti obrazovanja. Kako bi se identifikovali talentovani učenici osnovne škole, sproveden je eksperiment u kojem su primenjene metode računarske inteligencije i varirani parametri ovih metoda. Izvršena je uporedna analiza kako bi se došlo do metode koja daje najbolje rezultate za postavljeni problem.

**Ključne reči:** neuronske mreže, fazi logika, neuro-fazi sistemi, talentovanost

**Abstract:** Early recognition of talented students is of great importance for their further education and guidance. The meaning of the term “talent” is open to debate. In this paper the potential of application of artificial neural networks, fuzzy logic and neuro-fuzzy systems in the field of education is discussed. In order to identify talented primary school students, the experiment was conducted in which methods of computational intelligence were applied and their parameters varied. A comparative analysis was conducted in order to determine which method gives the best results for this particular problem.

**Keywords:** neural networks, fuzzy logic, neuro-fuzzy systems, talent

#### 1. UVOD

Prepoznavanje talenta u mladom uzrastu izuzetno je korisno zarad prilagođavanja pristupa nastavnika pojedinačnom učeniku, odnosno daljeg usmeravanja đaka i negovanja njihovih talenata. Naravno, pošto je „talentovanost“ širok pojam, otvoreno je pitanje šta on tačno podrazumeva, koje različite vrste talenta postoje, na koji način je talenat merljiv, itd.

Trend sve veće upotrebe metoda računarske inteligencije u ovu svrhu je evidentan, kao i da oni daju vrlo dobre rezultate. Često se kao izlazna promenljiva koja se želi predvideti koriste objektivni pokazatelji, poput rezultata učenika na određenim testovima [1,8,9]. Sa druge strane, postoje radovi koji se bave predviđanjem nadarenosti učenika procenjene od strane nastavnika [6]. Autori su se uglavnom susretali sa sličnim poteškoćama tokom svojih istraživanja, npr. sa malim brojem instanci u skupu podataka i preteranom prilagođenošću modela podacima (eng. *overfitting*).

Ovaj rad bavi se ispitivanjem mogućnosti primene metoda računarske inteligencije za identifikaciju talentovanih učenika trećeg i četvrtog razreda osnovne škole. Skup podataka koji je korišćen u ovom radu sadrži podatke o 151 učeniku, gde je svaki učenik opisan sa 48 različitih atributa. Na osnovu tih atributa formirane su 4 ulazne promenljive, i to: procena učenika od strane nastavnika, procena učenika od strane roditelja, procena učenika od strane vršnjaka i samoprocena učenika. Izlazna promenljiva formirana je na osnovu rezultata koji je učenik postigao na testu progresivnih matrica u boji. Kao klasifikatori u ovom radu korišćeni su veštačka neuronska mreža (eng. *artificial neural network*, NN), fazi sistem zaključivanja (eng. *fuzzy inference system*, FIS) i adaptivni neuro-fazi sistem zaključivanja (eng. *adaptive neuro-fuzzy inference system*, ANFIS). Po svim posmatranim metrikama kao najbolje rešenje se pokazao ANFIS model, čijom primenom su postignuti izvanredni rezultati (tačnost 92.05%, preciznost 95.58%, odziv 93.91%, F1 94.74%). Ipak, treba uzeti u obzir da su rezultati dobijeni uz prethodno poznavanje skupa podataka.

Ovaj rad organizovan na sledeći način. U drugom poglavlju dat je kratak pregled literature na navedenu temu. Treće poglavlje bavi se teorijskim osnovama. U četvrtom poglavlju bliže je opisan skup podataka koji je korišćen u eksperimentu. Peto poglavlje detaljnije objašnjava način izvođenja eksperimenta, dok su u

šestom poglavlju uporedno prikazani dobijeni rezultati. U poslednjem, sedmom odeljku izneti su zaključci ovog rada, kao i ideje za naredna istraživanja.

## 2. PREGLED RADOVA

Niz različitih algoritama računarske inteligencije i mašinskog učenja je korišćen u svrhu procene talentovanosti dece. U radu [9], testirana su dva algoritma: višeslojni perceptron i neuronske mreže sa radialno zasnovanim funkcijama. Autori su koristili podatke o 221 učeniku prosečnog uzrasta 10.1 godina. Učenike opisuju 23 ulazne promenljive (npr. procene nastavnika i drugih đaka, procenjena spremnosti đaka pred polazak u prvi razred, postignutim ocenama na polugodištu i na kraju školske godine, kao i na obrazovanju roditelja). Izlazna promenljiva bio je rezultat postignut na SPM (standardne progresivne matrice), gde se učenik smatrao nadarenim ukoliko je postigao bar 95%, odnosno 90% po navedenim normama [9]. Autori ističu da su postignuti dobri rezultati prilikom testiranja, dok kao probleme sa kojima su se suočili tokom istog navode mali broj nadarenih učenika, kao i *overfitting*.

U radu [6], obrađen je problem detekcije najvažnijih kriterijuma za pronalaženje talentovanih učenika. Takođe, autori su želeli da pronađu najbolji algoritam mašinskog učenja za prepoznavanje talentovanih učenika. Podaci korišćeni u ovom istraživanju odnosili su se na 400 đaka prosečnog uzrasta 8 godina. Kao izlazna varijabla posmatrana je nadarenost procenjena od strane nastavnika. Sledeći kriterijumi procenjeni su kao najznačajniji: postavljanje pitanja na visokom nivou, unikatne ideje u poređenju sa vršnjacima, razumevanje tema lekcija brže od vršnjaka, itd. Autori navode da je dobijeni model zasnovan na naivnom Bajesovom klasifikatoru imao tačnost od 92.4%. Ističu i da je neophodna obuka nastavnika, jer su često za pojam nadarenosti vezivali samo pozitivne attribute, iako talentovani đaci poseduju i zajedničke negativne osobine, npr. ometanje nastave, anksioznost, kao i to da ih vršnjaci smatraju za preozbiljne ili dosadne [6].

U radu [3], autori su posmatrali uzorak od 17 učenika osnovne škole. Kao ulazne podatke koristili su EEG rezultate prikupljene tokom rešavanja zadataka, uzrast učenika, težinu zadataka, vreme potrošeno na rešavanje, postignute bodove, kao i da li su ti bodovi iznad proseka grupe. Izlazna promenljiva bila je talentovanost u skladu sa njihovom pretpostavkom. Koristili su četiri različita algoritma mašinskog učenja. Tokom svog istraživanja naišli su na tipične probleme koji nastaju zbog malog skupa podataka, zbog čega su morali da balansiraju skup podataka. Najbolji rezultat dobijen je korišćenjem stabla odlučivanja.

Uprkos tome što se prilikom primene metoda mašinskog učenja dobijaju odlični rezultati, na osnovu gorenavedenih istraživanja može se zaključiti da je nezahvalno govoriti o njihovoj stvarnoj upotrebnoj vrednosti, uzevši u obzir činjenicu da je za početak vrlo teško odabrati izlaznu promenljivu. Drugim rečima, sam pojam nadarenosti nije dovoljno jasan, odnosno postoji više definicija šta ona tačno obuhvata. Stoga su često odabrane izlazne promenljive poprilično subjektivne, uzevši u obzir da se često zasnivaju na mišljenju odabranih osoba. Zbog toga bi, iako bi ovaj problem i dalje bio prisutan u značajnoj meri, bilo lakše govoriti o konkretnim talentima koji bi mogli biti merljivi, kao što su talenti za određenu disciplinu. U nastavku će biti reči o radovima koji su se fokusirali na preciznije traženi talenat.

U radu [1], autori su se bavili predviđanjem talenta za programiranje kod učenika prosečne starosti 11.29 godina. Kao izlaznu varijablu odabrali su rezultat koji su učenici postigli na kursu *Classic Maze Course* sa sajta *Coursera*, dok su ulazne promenljive bili 13 odgovora koje su učenici dali prilikom popunjavanja ankete. Korišćene su neuronske mreže i variran je broj neurona u skrivenom sloju. Na njihovom primeru Bajesova regularizacija koja je dala najbolje rezultate, ali autori napominju da je eksperiment vršen na malom broju uzoraka.

U radu [8], autori su radili su na predviđanju dara za matematiku kod učenika četvrtog razreda osnovne škole uz pomoć neuronske mreže. Oni su odabrali 60 ulaznih promenljivih, kojima se opisuje 5 komponenti matematičke nadarenosti, dok je kao izlazna promenljiva upotrebljena psihološka evaluacija darovitosti utemeljena u Ravenovim progresivnim matricama. Najbolji rezultat je sa 100% tačnosti prepoznao učenike sa darom, dok je za nenadarene bio u pravu u 71.43% slučajeva [8]. Autori su konstatovali da je to značajno bolji rezultat u odnosu na predviđanja vršena od strane nastavnog osoblja.

## 3. TEORIJSKE OSNOVE

### 3.1. Veštačke neuronske mreže

Veštačke neuronske mreže inspirisane su prirodnim neuronskim mrežama, kojima po kompleksnosti nisu blizu, ali postoje ključne sličnosti. Obe vrste mreža se sastoje od jednostavnih delova koji su međusobno povezani, i baš te konekcije određuju funkciju mreže [5]. Takođe, obe vrste mreža imaju paralelnu strukturu – svi neuroni rade istovremeno.

Veštačku neuronsku mrežu čine međusobno povezani neuroni, koji su organizovani po slojevima. Kada se vrednosti na ulazu neurona pomnože težinskim koeficijentima, dobijaju se ulazni podaci. Nad njima se potom primenjuje aktivaciona funkcija (eng. *activation function*), koja se bira u skladu sa problemom koji pokušava da se reši. Neuronska mreža ima definisano pravilo učenja na osnovu kojeg podešava parameter mreže. One uče preko primera i poseduju sposobnost generalizovanja. Veštačka neuronska mreža može da radi sa nepotpunim i nepreciznim podacima, sposobna je da generalizuje i obavlja kompleksna nelinearna ulazno/izlazna preslikavanja bez definisanja eksplicitnog matematičkog modela [7].

### 3.2. Fazi logika i fazi sistemi zaključivanja

Često je kod velikog broja problema teško povući jasnu granicu između crnog i belog, odnosno između tačnog ili netačnog kao kod klasične Bulove logike. Fazi logika omogućava rad sa bilo kojom vrednošću iz zatvorenog jediničnog intervala  $[0,1]$ , te je na taj način postignuto da se objekti mogu detaljnije opisati, lakše rangirati i tačnije upoređivati, dok se istinitost izraza u fazi svetu može gradirati [7]. Funkcija pripadnosti opisuje stepen pripadnosti elementa fazi skupu, pri čemu jedna fazi promenljiva može biti opisana pomoću jedne ili više funkcija pripadnosti, dok je process zaključivanja opisan uzročno posledičnim (ako-onda) pravilima. Stoga, može se reći da ovakav način rasuđivanja imitira ljudsko rezonovanje.

Fazi sistemi zaključivanja su sistemi koji pomoću teorije fazi skupova i fazi logičkih operacija formalizuju ekspertsko znanje radi modelovanja veza između ulaza i izlaza [7]. Za modelovanje FIS-a nije neophodno imati precizna saznanja o podacima, već na odgovarajuć način definisati ako-onda pravila, funkcije pripadnosti svake ulazne promenljive, kao i izlaznu promenljivu.

Adaptivni neuro-fazi sistem zaključivanja je sistem koji kombinuje principe fazi sistema zaključivanja i neuronskih mreža. On koristi trening parametre neuronske mreže da bi na adekvatan način modelovao parametre fazi sistema zaključivanja.

## 4. SKUP PODATAKA

Skup podataka koji je korišćen u eksperimentu sadrži podatke o 151 učeniku trećeg i četvrtog razreda osnovne škole. Svaki učenik je opisan sa 48 različitih atributa. Posmatrano je devet različitih vrsta inteligencije, odnosno darovitosti, i to: verbalno-lingvistička, logičko-matematička, vizuelno-prostorna, telesno-kinestetička, muzičko-ritmička, interpersonalna, intrapersonalna, prirodnjačka i filozofsko-duhovna.

Može se reći da su atributi u ovom skupu podataka intuitivno podeljeni u pet grupa, i to:

1. podaci o nastavniku i njegove procene sposobnosti posmatranog učenika (ukupno 12 atributa)
2. podaci o roditelju i njegove procene sposobnosti posmatranog učenika (ukupno 11 atributa)
3. podaci o vršnjaku i njegove procene sposobnosti posmatranog učenika (ukupno 11 atributa)
4. podaci o učeniku i njegova samoprocena sposobnosti (ukupno 11 atributa)
5. rezultati koje je posmatrani učenik postigao na testovima (ukupno 3 atributa)

Za izlaznu promenljivu mogao bi se odabrati neki od tri atributa koja se tiču rezultata koje je učenik ostvario na testovima, i to na: testu progresivnih matrica u boji, testu shvatanja i testu verbalnih sposobnosti.

U eksperimentu će biti korišćeni agregisani podaci kao ulazne promenljive, gde svaka od njih zapravo predstavlja agregisanu procenu jedne osobe, i to:

- procena učenika od strane nastavnika,
- procena učenika od strane roditelja,
- procena učenika od strane vršnjaka,
- samoprocena učenika.

Podaci su agregisani pomoću zbira procena svih devet vrsta inteligencije od strane jedne osobe. Pošto dalji tok eksperimenta nije zavisao od načina agregacije podataka, kao agregacioni metodi bi se mogli primeniti i složenije agregacije, poput Šokeovog integrala [4], logičke agregacije [10] ili LSP metoda [2], kako bi se rezultati unapredili.

## 5. POSTAVKA EKSPERIMENTA

Za potrebe odabira izlazne promenljive formiran je novi atribut, čija će vrednost biti 0 ili 1, a formirana na osnovu rezultata koji je učenik postigao na testu progresivnih matrica u boji. U kontekstu ovog eksperimenta, to je određeno na sledeći način: ukoliko je u pitanju učenik trećeg razreda, smatra se talentovanim ako je ostvario 31 poen ili više na pomenutom testu, odnosno makar 32 poena ako se radi o učeniku četvrtog razreda. U ovom eksperimentu kao klasifikatori su korišćeni prvo specijalizovana neuronska mreža za klasifikaciju, a zatim FIS i ANFIS.

Bitno je obratiti pažnju na raspodelu izlazne varijable po klasama, prilikom čega je uočeno da je značajno veći broj učenika koji su se pokazali kao talentovani od onih koji to nisu. Očekivano je da ovakva situacija može dovesti do priklanjanja neuronske mreže dominantnoj klasi.

U cilju klasifikacije neuronskom mrežom neophodno je pripremiti trening i test skupove. Test skup uvek obuhvata manji broj instanci, obično oko 20% originalnog skupa, ali rigidan odabir podataka za test skup ne bi bilo dobro rešenje jer bi onda i dalje postojala mogućnost da se u njemu nađu podaci koji bi bili važni za trening, odnosno proces učenja neuronske mreže. Ovo se rešava korišćenjem unakrsne validacije sa  $k$  particija. U ovom eksperimentu korišćena je unakrsna validacija sa 10 particija. Odabran način unakrsne validacije nosi rizik malog broja instanci za treniranje mreže.

## 6. REZULTATI

U nastavku teksta prikazani i analizirani su dobijeni rezultati. Svi eksperimenti su sprovedeni u MATLAB okruženju. Za evaluaciju rezultata ovog eksperimenta posmatrane su sledeće metrike: tačnost, preciznost, odziv i F1 metrika.

### 6.1. Rezultati dobijeni primenom veštačke neuronske mreže

Za predviđanje darovitosti učenika korišćeno je 15 arhitektura mreža, gde je variran broj skrivenih slojeva, broj neurona u sloju i aktivaciona funkcija.

Poređenjem inicijalnih rezultata mreža primećeno je da je tačnost poprilično visoka u sva tri posmatrana slučaja (od 75.66% do 81.32%), ali da je preciznost veoma niska (od 19.78% do 32.37%). Uzevši u obzir da preciznost predstavlja procenat stvarno pozitivnih opservacija u odnosu na sve one za koje je predviđeno da su pozitivne, kao i činjenicu da je u ovom skupu podataka značajno veći broj učenika koji su se pokazali kao talentovani od onih koji to nisu, možemo zaključiti da se neuronska mreža priklonila većinskoj klasi. Drugim rečima, desilo se upravo ono što smo očekivali - imajući u vidu da je 76.16% posmatranih učenika smatrano talentovanim, čak i da je neuronska mreža za svakog od 151 učenika predvidela da je talentovan, bila bi u pravu u 76.16% slučajeva. Stoga je bilo neophodno sprovesti neku od tehnika balansiranja.

Za ovaj konkretan problem odabrana je SMOTE tehnika, koja na osnovu već postojećih instanci manjinske klase generiše nove instance. Nakon njene primene postignuta je skoro savršena balansiranost, takva da 51.57% instanci pripada pozitivnoj klasi.

Nakon primene SMOTE tehnike za balansiranje podataka, eksperiment je ponovljen. Poređenjem rezultata jasno se istaklo da se mreža sa 3 skrivena sloja sa po redom 10, 20 i 30 neurona sa korišćenjem funkcije tangens hiperbolički pokazala kao najbolja po svakoj metrici sem odziva. Dobro su se pokazale i mreža sa jednim skrivenim slojem sa 30 neurona gde je takođe korišćena aktivaciona funkcija tangens hiperbolički, kao i mreža sa tri skrivena sloja, sa redom 20, 30, 40 neurona i aktivacionom funkcijom tangens hiperbolički. U Tabeli 1. su redom navedeni najbolji rezultati.

**Tabela 1:** Najbolji rezultati dobijeni primenom neuronske mreže

	tačnost	preciznost	odziv	F1
[10*20*30]	64.98%	71.33%	62.20%	66.46%
(30)	61.48%	67.81%	59.63%	63.46%
[20*30*40]	64.04%	58.54%	64.84%	61.53%

Valja naglasiti da je tačnost očekivano značajno opala u odnosu na rezultate dobijene pre balansiranja klasa izlazne promenljive, odnosno pre primene SMOTE tehnike. Sa druge strane, ako posmatramo preciznost i F1 metriku, novodobijeni rezultati su bespogovorno bolji. Stoga, možemo tvrditi da ovako definisane mreže modeluju posmatrani problem na realniji način.

### 6.2. Rezultati dobijeni primenom FIS modela

U nastavku eksperimenta formiran je FIS model Mamdani tipa radi određivanja da li je učenik talentovan ili ne. Za ulazne promenljive, gde svaka od njih zapravo predstavlja agregisanu procenu jedne osobe, korišćene su Gausove funkcije pripadnosti. Posmatrane su po 3 funkcije pripadnosti, odnosno procena može biti okarakterisana kao niska, srednja ili visoka. Za izlaznu promenljivu korišćena je kombinacija dve Gausove funkcije.

Zaključivanje se vrši na osnovu 31 ekspertski definisanog pravila. Naime, učenik se smatra talentovanim ako:

- sve četiri procene potpadaju pod visoke
- su bilo koje tri procene visoke, a preostala jedna srednja
- su bilo koje dve procene visoke, a druge dve srednje

Učenik se ne smatra talentovanim ako:

- sve četiri procene potpadaju pod niske
- sve četiri procene potpadaju pod srednje
- su bilo koje tri procene niske, a preostala jedna srednja
- su bilo koje dve procene niske, a druge dve srednje
- su bilo koje tri procene srednje, a preostala jedna niska

Ipak, testiranjem sačinjenog modela zaključeno je da je model isuviše strogo podešen. Za potrebe nastavka eksperimenta pretpostavljeno je da je učenik talentovan ukoliko je izlazna vrednost koju mu je model dodelio veća od prosečne. Na osnovu tako dobijenih rezultata (tačnost 49.67%, preciznost 80%, odziv 45.22% i F1 57.78%) se dalo zaključiti da je odabran previsok kriterijum. Naime, pošto je preciznost visoka, možemo zaključiti da su u većini slučajeva kada su opservacije predviđene kao pozitivne one to zaista i bile. Ipak, kada pogledamo odziv koji je daleko niži od preciznosti, jasno je da je veliki broj stvarno pozitivnih opservacijama pogrešno predviđen. Ovo se dogodilo zbog činjenice da je početnom pretpostavkom u kom slučaju je učenik talentovan njih 115 od 151 u posmatranom skupu smatrano talentovanim.

**Tabela 2:** Najbolji rezultat dobijen primenom FIS modela

	tačnost	preciznost	odziv	F1
FIS	66.89%	78.76%	77.39%	78.07%

Dakle, pošto smatramo da je 76% posmatranih učenika talentovano, više bi imalo smisla u kontekstu izlaza FIS modela reći da je predviđeno da je učenik netaalentovan ako potpada ispod 0.25, drugim rečima ako pripada prvom kvartilu. U sledećem prolazu je testirana ova pretpostavka i primećeno je da su vrednosti značajno bolje od prethodnog prolaza, ali pored toga značajno bolje i od vrednosti metrika dobijenih uz pomoć neuronske mreže. Ipak, valja napomenuti da je ovo postignuto „veštački“, tj. da je kriterijum podešen poznajući skup podataka. Finalni rezultati su sumirani u Tabeli 2.

### 6.3. Rezultati dobijeni primenom ANFIS modela

Za potrebe rada sa ANFIS modelom skup podataka podeljen je na delove za test i trening, i to na način da je zadržan odnos klasa, odnosno u delu za trening nalazi se 80% opservacija pozitivne klase i 80% opservacija negativne klase. Nakon što su učitani podaci, generisan je FIS tako da je za svaku ulaznu promenljivu potrebno definisati funkcije pripadnosti. Pri treniranju se bira optimizacioni metod, gde su ponuđene opcije *hybrid* i *backpropagation*. Takođe se odabira i dozvoljena greška, kao i broj epoha treniranja. Pravila se automatski generišu. Kao što je ranije pomenuto, pošto smatramo da je 76% posmatranih učenika talentovano, imalo bi smisla u kontekstu izlaza modela reći da je predviđeno da je učenik netaalentovan ako potpada ispod 0.25, drugim rečima ako pripada prvom kvartilu. Kao najbolji ANFIS model pokazao se onaj gde je za svaku ulaznu promenljivu korišćeno po 3 funkcije pripadnosti, i to trougaonog tipa, dok je tip funkcije pripadnosti izlazne promenljive podešen kao linearan. Takođe, pri treniranju FIS odabran je hibridni optimizacioni metod, 3 epohe, a pokazalo se da nije bitno da li je dozvoljena greška podešena na 0.03 ili 0.05. U nastavku je tabela iz koje se jasno vidi da je na ovaj način postignut najbolji rezultat.

**Tabela 3:** Najbolji rezultat dobijen primenom ANFIS modela

	tačnost	preciznost	odziv	F1
ANFIS	92.05%	95.58%	93.91%	94.74%

Ispostavilo se da su automatski generisana pravila dala bolje rezultate od ručno podešenih. Ipak, neophodno je ponovo naznačiti da su oni dobijeni uz prethodno poznavanje skupa podataka. Drugim rečima, ovi rezultati su postignuti nad skupom podataka gde je značajno veći broj talentovanih od netaalentovanih učenika, odnosno gde postoji nebalansiranost izlazne klase. Da je izlazna klasa bila balansirana, predviđeni izlazni podaci bi bili drugačije posmatrani, tj. umesto prvog kvartila posmatrala bi se srednja vrednost.

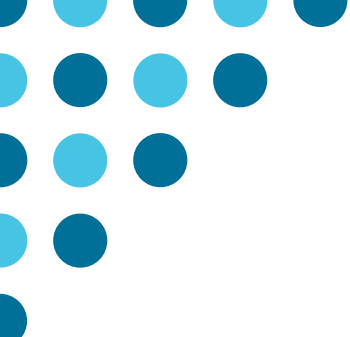
## 7. ZAKLJUČAK

Cilj ovog rada bio je ispitivanje mogućnosti primene veštačkih neuronskih mreža, fazi logike i neuro-fazi sistema u oblasti obrazovanja, a u svrhu otkrivanja talentovanih učenika osnovne škole. U radu je vršena uporedna analiza ovih tehnika kako bi se došlo do najboljeg rezultata. Pokazano je da je na postojećem skupu podataka ANFIS dao najbolje rezultate. Imajući u vidu da je ovaj sistem izuzetno jednostavan za tumačenje, možemo reći da je najpodesniji za modelovanje ovog problema. Ekspertski definisan FIS i neuronska mreža su dali približne rezultate, mada je za neuronsku mrežu koršćen modifikovan skup podataka. Ipak, imajući u vidu da je neuronska mreža model crne kutije bez lake interpretacije, smatramo da se ona najlošije pokazala za modelovanje ovog problema. Osnovni problemi koji su doveli do ovoga su pre svega mali uzorak, a zatim i potencijalno loše odabrane arhitekture mreže i algoritam učenja.

U radu je posebna pažnja posvećena i analizi samog pojma „talentovanost“, odnosno šta taj pojam tačno podrazumeva i na koji način bi se moglo definisati da li je učenik talentovan ili ne. Ovo pitanje otvara vrata daljim eksperimentima, u kojima bi se moglo početi od drugačije pretpostavke šta talentovanost predstavlja. U ovom radu, učenik je smatran talentovanim ako je ostvario određeni broj poena na testu progresivnih matrica u boji. Buduća istraživanja mogu uzeti u obzir i učenikove rezultate postignute na testovima verbalnih sposobnosti i testovima shvatanja. Pored toga, budući da je u ovom radu učenik smatran talentovanim ako je ostvario 31, odnosno 32 poena na posmatranom testu, može se polemisati o tome da li je taj kriterijum strog. S tim u vezi, mogli bi se posmatrati i granični slučajevi („možda“ talentovani), tako da bi se ovaj problem predstavio kroz klasifikaciju na tri klase. Takođe, postoji mogućnost da bi drugačiji način agregacije ulaznih podataka doveo do drugačijih rezultata. Svakako, može se pretpostaviti da bi rezultati bili još bolji ukoliko bi skup podataka sadržao više instanci.

## LITERATURA

- [1] Çetinkaya, A., & Baykan, Ö. K. (2020). Prediction of middle school students' programming talent using artificial neural networks. *Engineering Science and Technology, an International Journal*, 23(6), 1301-1307.
- [2] Dujmović, J. J., & Nagashima, H. (2006). LSP method and its use for evaluation of Java IDEs. *International Journal of Approximate Reasoning*, 41(1), 3-22.
- [3] Ghali, R., Tato, A., & Nkambou, R. (2019). Using EEG Features and Machine Learning to Predict Gifted Children. In *The Thirty-Second International Flairs Conference*.
- [4] Grabisch, M., & Roubens, M. (2000). Application of the Choquet integral in multicriteria decision making. *Fuzzy Measures and Integrals-Theory and Applications*, 348-374.
- [5] Hagan, M., Demuth, H., Beale, M., & DeJesus O. (2014). *Neural Network Design* (2nd ed.). Martin Hagan.
- [6] Kartal, E., Özyaprak, M., Özen, Z., Şimşek, İ., Köse Biber, S., Biber, M., & Can, T. (2020). Asking the Right Questions to Nominate A Student as Gifted and Talented: A Machine Learning Approach. *Bilişim Teknolojileri Dergisi*, 13(4), 385-400.
- [7] Milošević, P., Rakićević, A., Dragović, I., Poledica, A., Petrović, B., Vukićević, A., & Zukanović, M. (2021). *Računarska inteligencija: praktikum u MATLAB-u*. Beograd: Fakultet organizacionih nauka.
- [8] Pavleković, M., Zekić-Sušac, M., & Đurđević, I. (2011). Model neuronskih mreža za predviđanje matematičke darovitosti u djece. *Croatian Journal of Education: Hrvatski časopis za odgoj i obrazovanje*, 13(1), 10-41.
- [9] Pavlin-Bernardić, N., Ravić, S., & Matić, I. P. (2016). The application of artificial neural networks in predicting children's giftedness. *Suvremena psihologija*, 19(1), 49-58.
- [10] Radojevic, D. (2008). Logical aggregation based on interpolative. *Mathware & Soft Computing*, 15(1), 125-141.



# MENADŽMENT

---

# MANAGEMENT



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business





## UTICAJ SOCIO-KULTURALNIH KARAKTERISTIKA POTROŠAČA NA KUPOVINU GLOBALNIH BRENDOVA

### INFLUENCE OF SOCIO-CULTURAL CHARACTERISTICS OF CONSUMERS ON THE PURCHASE OF GLOBAL BRANDS

STEFAN ZDRAVKOVIĆ<sup>1</sup>, ALEKSANDAR JOVANOVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Kragujevcu – Ekonomski fakultet, szdravkovic034@yahoo.com

<sup>2</sup> Univerzitet u Kragujevcu – Fakultet inženjerskih nauka, a.jovanovic@kg.ac.rs

**Rezime:** Međunarodne ekonomske integracije i globalizacija svetskog tržišta dovele su do pojave velikog broja multinacionalnih kompanija. Konkurencija na tržištu je veoma izražena, tako da formulisanje i implementacija optimalne marketing strategije (globalna, lokalna ili glokalna) predstavlja veliki izazov za marketing menadžere multinacionalnih kompanija. Odluka potrošača da prihvate globalne brendove se nalazi pod uticajem brojnih faktora. Prilikom procesa internacionalizacije poslovanja, potrebno je sprovesti istraživanje tržišta, kako bi se dobile informacije o karakteristikama lokalnih potrošača. Kultura oblikuje sistem vrednosti pojedinca, stil života, kao i ponašanje u procesu kupovine. Cilj istraživanja je da se utvrdi uticaj kulturalne otvorenosti (svetska promišljenost, kulturalna inteligencija) i potrošačkog etnocentrizma (animozitet, tradicija) na kupovinu globalnih brendova od strane potrošača i na osnovu dobijenih rezultata da preporuka marketarima o primeni optimalne marketing strategije (globalna, lokalna, glokalna). Sprovedeno je empirijsko istraživanje metodom ankete na teritoriji Republike Srbije. Dobijeni podaci su obrađeni putem statističkog softvera SPSS. Rezultati su pokazali da kulturalna otvorenost ostvaruje pozitivan, a potrošački etnocentrizam negativan uticaj na kupovinu globalnih brendova od strane potrošača, tako da istraživanje omogućava menadžerske implikacije.

**Ključne reči:** Kulturalna otvorenost, Potrošački etnocentrizam, Globalni brendovi

**Abstract:** International economic integration and globalization of the world market have led to the emergence of a large number of multinational companies. Competition in the market is very pronounced, so the formulation and implementation of optimal marketing strategy (global, local or gloocal) is a great challenge for marketing managers of multinational companies. Consumer decisions to accept global brands are influenced by a number of factors. During the process of internationalization of business, it is necessary to conduct market research, in order to obtain information on the characteristics of local consumers. Culture shapes an individual's value system, lifestyle and behavior in the buying process. The aim of the research is to determine the influence of cultural openness (worldly providence, cultural intelligence) and consumer ethnocentrism (animosity, tradition) on the purchase of global brands by consumers and based on the results to recommend marketers on the implementation of optimal marketing strategy (global, local, gloocal). Empirical research was conducted using the survey method on the territory of the Republic of Serbia. The obtained data were processed through the statistical software SPSS. The results showed that cultural openness has a positive and consumers ethnocentrism has a negative impact on the purchase of global brands by consumers, so that the research provides managerial implications.

**Keywords:** Cultural openness, Consumer ethnocentrism, Global brands

## 1. UVOD

Proces globalizacije je doveo do razvoja ekonomske saradnje između država, ukidanja trgovinskih barijera i liberalizacije spoljnotrgovinskog prometa roba i usluga (Cleveland et al., 2022) Došlo je do pojave velikog broja međunarodno aktivnih kompanija, koje plasiraju svoje brendove na tržišta širom sveta. Kompanije koje su prepoznatljive na globalnom nivou mogu da formulišu *globalnu marketing strategiju*, koja podrazumeva da kompanija na svako tržište na kome nastupa plasira potpuno identične brendove bez ikakvih prilagođavanja, a navedena strategija omogućava niske troškove proizvodnje i realizaciju ekonomije obima (Lysonski & Durvasula, 2013). Takođe, određeni broj potrošača ima razvijenu globalnu potrošačku kulturu, kulturalnu

otvorenost, svetsku promišljenost, visok stepen razvijenosti kulturalne inteligencije, kosmopolitistička shvatanja i orijentisanost, što implicira kupovinu brendova koji su prepoznatljivi na globalnom nivou (Cleveland & Bartsch, 2019). Sa druge strane, kompanija može prilikom nastupa na stranom tržištu da primeni *lokalnu marketing strategiju*, koja podrazumeva da su brendovi kompanije prilagođeni karakteristikama lokalnog tržišta, što podrazumeva da se uvažava kulturalni okvir i etnocentrične tendencije potrošača, tako što se proizvodi reklamiraju putem domaćih medija, sadrže određene simbole koji su karakteristični za tu državu i slično (Prince et al., 2016). Određeni broj ljudi ima izražen potrošački etnocentrizam, patriotizam, nacionalni identitet, poštuju tradiciju svoje države i imaju animozitet prema nekim drugim kulturama, što implicira kupovinu proizvoda koji su prilagođeni njihovom kulturalnom i etnocentričnom okviru (Zdravković, 2021). Kultura oblikuje sistem vrednosti pojedinca, verovanja, pravila ponašanja, stil života, a u značajnoj meri utiče i na odluke koje potrošači donose u procesu kupovine (Huang & He, 2021). Autori Strizhakova & Coulter (2019), kao i studija Gammoh et al. (2020) predlažu da prilikom internacionalizacije poslovanja marketing menadžeri međunarodno aktivnih kompanija implementiraju *glokalnu marketing strategiju* koja može da predstavlja optimalno rešenje, jer omogućava standardizaciju poslovanja, ali i diferenciranje i prilagođavanje u određenoj meri kulturalnim karakteristikama i zahtevima potrošača na stranom tržištu (Tabela 1).

Tabela 1: Izbor marketing strategije prilikom internacionalizacije poslovanja

	Poslovanje kompanije	Socio-kulturalne karakteristike potrošača
<b>Globalna marketing strategija</b>	-niski troškovi proizvodnje; -uštede usled realizacije ekonomije obima; -jednostavno uvođenje proizvoda na strano tržište.	-razvijena globalna potrošačka kultura; -otvorenost prema drugim kulturama, visok stepen kulturalne inteligencije; -kosmopolitistička shvatanja, svetska promišljenost.
<b>Lokalna marketing strategija</b>	-diferenciranje u odnosu na konkurenciju; -usklađivanje poslovanja i dobijanje subvencija od strane države; -prilagođavanje kulturalnom okviru lokalnih potrošača.	-izražen visok stepen potrošačkog etnocentrizma; -patriotizam, nacionalni identitet, konzervativizam; -poštovanje tradicije, animozitet prema drugim državama.
<b>Glokalna marketing strategija</b>	-primena globalnog i lokalnog pristupa u poslovanju; -mogućnost primene standardizacije i diferenciranja u poslovanju; -privlačenje globalnih i lokalnih potrošača.	-kosmopolitistička orijentisanost i nacionalna svest; -kulturalna otvorenost i poštovanje sopstvene kulture; -razvijen stepen kulturalne inteligencije i potrošačkog etnocentrizma.

Predmet istraživanja je analiza uticaja socio-kulturalnih karakteristika potrošača na kupovinu globalnih brendova kao i sagledavanje različitih tipova marketing strategija koje kompanija može da primeni prilikom internacionalizacije poslovanja. Cilj istraživanja je da se utvrdi da li kulturalna otvorenost (svetska promišljenost, kulturalna inteligencija) i potrošački etnocentrizam (animozitet, tradicija) ostvaruju uticaj na kupovinu globalnih brendova od strane potrošača. Sprovedeno je empirijsko istraživanje, a na osnovu rezultata, donose se zaključci o uticaju navedenih socio-kulturalnih karakteristika potrošača na kupovinu globalnih brendova i daje se predlog o primeni određenog tipa marketing strategija na tržištu Republike Srbije.

## 2. PREGLED LITERATURE

Proces globalizacije i porast broja multinacionalnih kompanija su implicirali razvoj globalne potrošačke kulture (Pratono & Arli, 2020). Određena grupa potrošača smatra da su globalni brendovi garancija kvaliteta i pouzdanosti, kao i da su brendirani proizvodi prepoznatljivu po odličnim tehničkim specifikacijama i performansama. Osobe koje kupuju globalne brendove imaju razvijenu *kulturalnu otvorenost*, koja podrazumeva želju pojedinca da upozna istoriju, tradiciju, običaje, pravila ponašanja, verovanja i sistem vrednosti drugih kultura (Shi & Shan, 2019). Neki od njenih najznačajnijih pokretača su *svetska promišljenost* i *kulturalna inteligencija* (Earley & Ang, 2003). *Svetska promišljenost* podrazumeva da pojedinac prati globalne trendove, pojavu tehnoloških otkrića, često putuje van granica svoje zemlje, ima kosmopolitistička shvatanja i orijentisanost (Frias-Jamilena et al., 2018). *Kulturalna inteligencija* podrazumeva da osoba ima razvijen skup metakognitivnih, kognitivnih, motivacionih i bihevioralnih sposobnosti, koje joj omogućavaju

da se adaptira prilikom situacije koje su multikulturalne prirode (Tuan, 2016). Osobe koje imaju razvijen visok stepen kulturalne inteligencije poznaju društvene sisteme drugih država (pravni, ekonomski), motivisane su da ostvaraju interakciju sa ljudima koji potiču iz drugih kultura, da uče strane jezike i da na taj način dodatno unapređuju svoju kulturalnu inteligenciju (Lorenz et al., 2018). Takođe, takve osobe se putem svojih verbalnih i neverbalnih kompetencija (akcent, izraz lica, govor tela) na jednostavan način prilagođavaju situacijama koje karakteriše multikulturalnost, kao što su na primer komunikacija i saradnja sa poslovnim partnerima koji dolaze iz drugih kultura, ili odlazak studenata u inostranstvo na svetska takmičenja (Rahman et al., 2021).

Jimenez & San-Martin (2016) u svojoj studiji navode da svetska promišljenost ostvaruje pozitivan uticaj na odluku potrošača da kupe globalne brendove. Nie & Wang (2021) navode da je svetska promisljenost jedan od glavnih pokretača kulturalne otvorenosti i da ostvaruje pozitivan uticaj na kupovinu stranih i globalno prepoznatljivih brendova. Pratonu & Arli (2020) su utvrdili da kulturalna inteligencija ostvaruje pozitivan uticaj na odluku potrošača da kupe strane brendove. Zdravković & Peković (2021) su utvrdili da kulturalna inteligencija ostvaruje pozitivan uticaj na turiste da posete inostrane turističke destinacije, kao i da pozitivno utiče na razvoj globalne potrošačke kulture i kupovinu brendova koji su prepoznatljivi na globalnom nivou. Na osnovu navedenih studija, mogu se formulisati sledeće istraživačke hipoteze:

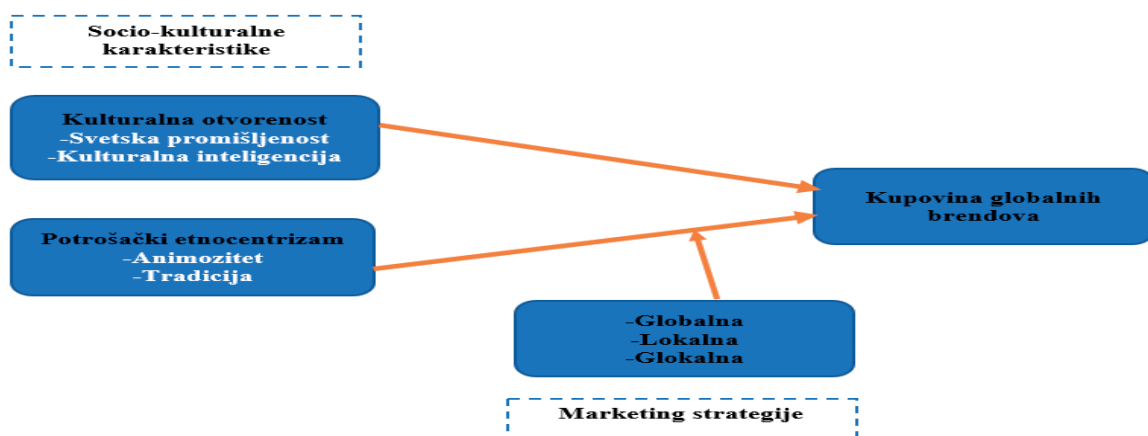
*H1a: Svetska promišljenost ostvaruje pozitivan statistički značajan uticaj na kupovinu globalnih brendova od strane potrošača.*

*H1b: Kulturalna inteligencija ostvaruje pozitivan statistički značajan uticaj na kupovinu globalnih brendova od strane potrošača.*

Određeni broj ljudi ima izražen visok stepen etnocentrizma, koji predstavlja sociološki fenomen i može se definisati kao shvatanje ljudi da je njihova sopstvena kultura po elementima poput istorije, tradicije, kulturalnog nasleđa dominantnija u poređenju sa kulturom bilo koje druge zemlje (Shimp & Sharma, 1987). Iz bazičnog koncepta, razvio se koncept *potrošačkog etnocentrizma* koji predstavlja ekonomski fenomen i može se definisati kao sklonost potrošača da kupuju proizvode domaćeg porekla (Pentz et al., 2017). Neki od najznačajnijih pokretača potrošačkog etnocentrizma su *animozitet* i *tradicija* (Shankarmahesh, 2006). Određeni broj ljudi prema nekim državama razvija ratni ili ekonomski animozitet. Ratni animozitet nastaje kao posledica vojnih sukoba sa drugom zemljom, dok ekonomski animozitet pojedinci razvijaju prema zemljama koje njihovoj matičnoj zemlji uvode određene ekonomske sankcije, trgovinske barijere i onemogućavaju izvoz robe na njeno tržište (Marinković, 2017). Pojam tradicije podrazumeva poštovanje istorije, običaja, pravila ponašanja koje su karakteristični za jedno kulturalno područje (Šapić, 2017). Takođe, tradicija podrazumeva poštovanje društvenog poretka, kao i kulturalnog nasleđa jedne države. Souiden et al. (2018) su utvrdili da animozitet prema drugim državama ostvaruje negativan uticaj na odluku potrošača da kupe strane i globalne brendove. Fernandez-Ferrin et al. (2020) su utvrdili da animozitet prema drugim državama ostvaruje pozitivan uticaj na odluku potrošača da kupe domaće proizvode i usluge, a negativno utiče na kupovinu stranih brendova. Prethodna istraživanja su utvrdila da tradicija ostvaruje pozitivan uticaj na kupovinu proizvoda domaćeg porekla, kao i negativan uticaj na kupovinu stranih brendova (Šapić, 2017; Gammoh et al, 2020). Na osnovu navedenog, mogu se formulisati sledeće istraživačke hipoteze, kao i istraživački model (Slika 1).

*H2a: Animozitet prema drugim državama ostvaruje negativan statistički značajan uticaj na kupovinu globalnih brendova od strane potrošača.*

*H2b: Tradicija ostvaruje negativan statistički značajan uticaj na kupovinu globalnih brendova od strane potrošača.*



### 3. METODOLOGIJA ISTRAŽIVANJA

Empirijsko istraživanje metodom ankete je sprovedeno u aprilu 2022. godine na teritoriji Republike Srbije (Beograd, Kragujevac). Ukupan broj ispitanika u uzorku je 218 i oni su davali odgovore na konstatacije putem Likertove skale ocenama od 1 do 7 (1- apsolutno se ne slažem; 7- apsolutno se slažem). U uzorku ima 98 ispitanika ženskog i 120 ispitanika muškog pola. Najveći broj ispitanika je starosne uzrasti od 21 do 30 godina (43.6%). Kada je obrazovanje u pitanju, najveći broj ispitanika u uzorku ima završen fakultet, njih 110 (50.5%), a kada je u pitanju status, najveći broj ispitanika u uzorku čine zaposleni i studenti.

Konstatacije putem kojih su merene istraživačke varijable svetska promišljenost, kulturalna inteligencija, animozitet i tradicija su preuzete iz relevantne marketinške literature, a pregled je dat u Tabeli 2.

Tabela 2: Istraživačke varijable i konstatacije putem kojih su merene

Istraživačke varijable		Konstatacije	Izvor
Kulturalna otvorenost	Svetska promišljenost	1. Pratim inovativna tehnološka otkrića koja se pojavljuju na tržištu.	Frank & Watchravesringkan (2016)
		2. Često putujem u inostranstvo.	
		3. Pratim globalne trendove.	
	Kulturalna inteligencija	1. Često razgovaram sa ljudima iz drugih kultura.	Earley & Ang (2003)
		2. Učim strane jezike.	
		3. Upoznajem kulturalno nasleđe drugih zemalja.	
Potrošački etnocentrizam	Animozitet	1. Sumnjam u kvalitet stranih brendova.	Marinković (2017)
		2. Nemam poverenje u strane poslovne partnere.	
		3. Kupovina stranih brendova slabi nacionalnu ekonomiju.	
	Tradicija	1. Poštujem istoriju svoje zemlje.	Šapić (2017)
		2. Porodične vrednosti su veoma važne.	
		3. Iz moje zemlje potiču brojne poznate ličnosti.	
Kupovina globalnih brendova		1. Globalni brendovi su signal kvaliteta proizvoda.	Strizhakova & Coulter (2019)
		2. Globalni brendovi omogućavaju status u društvu.	
		3. Preporučujem prijateljima globalne brendove.	

Statistički softver SPSS je korišćen za analizu podataka. Od statističkih analiza primenjena je analiza pouzdanosti, kao i višestruka regresiona analiza u funkciji potvrđivanja ili odbacivanja istraživačkih hipoteza.

### 4. REZULTATI ISTRAŽIVANJA

Prva statistička analiza koja je primenjena u okviru ovog istraživanja je analiza pouzdanosti, kako bi se ispitalo da li su konstatacije putem kojih se mere istraživačke varijable međusobno interno konzistentne (Tabela 3).

Tabela 3: Rezultati analize pouzdanosti

Istraživačke varijable	Crnobach's Alpha
Svetska promišljenost	0.78
Kulturalna inteligencija	0.82
Animozitet	0.84
Tradicija	0.88
Kupovina globalnih brendova	0.87

Izvor: kalkulacija autora u statističkom programu SPSS

Analiza pouzdanosti se sagledava na osnovu vrednosti *Crnobach's Alpha* koeficijenta. Da bi konstatacije bile međusobno interno konzistentne i istraživački model relevantan i pouzdan potrebno je da vrednost koeficijenta bude veća od praga 0.70 (Nunally, 1978). Navedeni uslov je zadovoljen kod svih istraživačkih varijabli. U istraživanju je primenjena višestruka regresiona analiza, kako bi se ispitalo uticaj *kulturalne otvorenosti* (*svetska promišljenost*, *kulturalna inteligencija*) i *potrošačkog etnocentrizma* (*animozitet*, *tradicija*) na kupovinu globalnih brendova od strane potrošača. Rezultati su prikazani putem Tabele 4.

Tabela 4: Rezultati višestruke regresione analize (zavisna varijabla: *kupovina globalnih brendova*)

Nezavisne varijable		Beta koeficijent	T test	Sig vrednost	VIF koeficijent
Kulturalna otvorenost	Svetska promišljenost	0.223**	4.254	0.000	2.541
	Kulturalna inteligencija	0.281**	4.579	0.000	3.471
Potrošački etnocentrizam	Animozitet	-0.187**	-3.354	0.002	2.887
	Tradicija	-0.193**	-3.567	0.005	1.896

Izvor: kalkulacija autora u statističkom programu SPSS Napomena: \*\*-nivo signifikantnosti 0.01;  $R^2 = 0.621$

Koeficijent determinacije  $R^2 = 0.621$  pokazuje da je postavljenim regresionim modelom objašnjeno 62.1% procenata varijabiliteta zavisne varijable *kupovina globalnih brendova*. U istraživanju se ne javlja problem multikolinearnosti jer je vrednost VIF koeficijenta manja od praga 5 (Field, 2000). *Kulturalna otvorenost* preko svojih pokretača *svetska promišljenost* ( $\beta = 0.223$ ; Sig= 0.000) i *kulturalna inteligencija* ( $\beta = 0.281$ ; Sig= 0.000) ostvaruje pozitivan statistički značajan uticaj na kupovinu globalnih brendova od strane potrošača, tako da su **potvrđene istraživačke hipoteze H1a i H1b**. *Potrošački etnocentrizam* preko svojih pokretača *animozitet* ( $\beta = -0.187$ ; Sig= 0.002) i *tradicija* ( $\beta = -0.193$ ; Sig= 0.005) ostvaruje negativan statistički značajan uticaj na kupovinu globalnih brendova od strane potrošača, tako da su **potvrđene istraživačke hipoteze H2a i H2b**.

## 5. ZAKLJUČAK

Istraživanje je sprovedeno na teritoriji Republike Srbije sa ciljem da se ispita uticaj socio-kulturalnih karakteristika potrošača na kupovinu globalnih brendova. Rezultati su pokazali da **kulturalna otvorenost** (svetska promišljenost, kulturalna inteligencija) ostvaruje pozitivan, a **potrošački etnocentrizam** (animozitet, tradicija) negativan uticaj na kupovinu globalnih brendova od strane potrošača. Cilj istraživanja je bio i da se na osnovu rezultata da preporuka marketarima o tome koja od mogućih marketing strategija (globalna, lokalna, globalna) je optimalna za primenu na tržištu Republike Srbije. **Globalna marketing strategija** podrazumeva da multinacionalna kompanija na svako tržište plasira identične brendove, dok **lokalna marketing strategija** podrazumeva prilagodjavanje proizvoda karakteristikama lokalnih potrošača, na svakom tržištu na kome kompanija nastupa. Rezultati ovog istraživanja pokazuju da je na **tržištu Republike Srbije** optimalno primeniti **globalnu marketing strategiju**, jer kulturalna otvorenost ostvaruje pozitivan uticaj, a potrošački etnocentrizam negativan uticaj na kupovinu globalnih brendova od strane potrošača. **Globalna marketing strategija** kombinuje globalni i lokalni pristup, odnosno uvažava potrebe globalnih potrošača koji imaju kosmopolitistička shvatanja, svetsku promišljenost, kulturalnu inteligenciju, ali i lokalnih potrošača koji imaju izražen potrošački etnocentrizam, animozitet prema drugim državama i poštuju tradiciju svoje zemlje. **Glavni doprinos i originalnost istraživanja** jeste predstavljenje rezultata marketing menadžerima o uticaju socio-kulturalnih karakteristika potrošača na kupovinu globalnih brendova, kao i preporuka o određenom tipu marketing strategije koji je optimalno primeniti na tržištu Republike Srbije. **Ograničenje istraživanja** se ogleda u relativno maloj veličini uzorka, kao i činjenici da je istraživanje sprovedeno na teritoriji jedne zemlje. **Pravci budućih istraživanja** bi mogli da se zasnivaju na analizi uticaja navedenih socio-kulturalnih karakteristika na kupovinu određenih kategorija proizvoda (hrana, garderoba, automobili), jer potrošač može da kupuje domaće prehrambene proizvode, a sa druge strane luksuznu garderobu i automobile globalnih proizvođača, kako bi mu oni omogućili status u društvu. Takođe, trebalo bi uključiti još neku determinantu, poput kosmopolitizma, ksenocentrizma, a imidž zemlje porekla može uticati na kupovne odluke. Naredno istraživanje bi moglo da se sprovede na teritoriji više zemalja, jer kros-kulturalna istraživanja su sveobuhvatnija i pružaju mogućnost da se izvrši komparativna analiza stavova pripadnika različitih nacija u pogledu kupovine globalnih brendova.

## LITERATURA

- [1] Cleveland, M. & Bartsch, F. (2019). Global consumer culture: epistemology and ontology. *International Marketing Review*, 36(4), 556-580. doi: 10.1108/IMR-10-2018-0287
- [2] Cleveland, M., Papadopoulos, N. & Laroche, M. (2022). Global consumer culture and national identity as drivers of materialism: an international study of convergence and divergence. *International Marketing Review*, 39(2), 207-241. doi: 10.1108/IMR-02-2021-0097
- [3] Earley, P. C. & Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures*. Palo Alto, CA: Stanford University Press
- [4] Field, A. (2000). *Discovering statistics using SPSS for Windows*. Thousand Oaks: Sage Publication.
- [5] Frank, P. & Watchravesringkan, K.(T). (2016). Exploring antecedents and consequences of young consumers' perceived global brand equity. *Journal of Product and Brand Management*, 25(2), 160-170.

- [6] Frias-Jamilena, D., Sabiote-Ortiz, C., Martin-Santana, J. & Beerli-Palacio, A. (2018). The effect of Cultural Intelligence on consumer-based destination brand equity. *Annals of Tourism Research*, 72(September): 22-36. doi: 10.1016/j.annals.2018.05.009
- [7] Fernández-Ferrín, P. Bande, B., Martín-Consuegra, D., Díaz, E. & Kastenholz, E. (2020). Sub-national consumer ethnocentrism and the importance of the origin of food products: an exploratory analysis. *British Food Journal*, 122(3), 995-1010. doi: 10.1108/BFJ-09-2019-0746
- [8] Gammoh, B.S., Koh, A.C. & Okoroafo, S.C. (2020). Positioning strategies of high-tech products: cross-cultural moderating effects of ethnocentrism and cultural openness. *Journal of Product and Brand management*, 29(3), 369-385. doi: 10.1108/JPBM-10-2018-2048
- [9] Huang, H. & He, J. (2021). When face meets globalization: How face drives consumers' attitudes toward global consumer culture positioning. *International Marketing Review*, 38(1), 184-203.
- [10] Jiménez, N. & San-Martin, S. (2016). The central role of the reputation of country-of-origin firms in developing markets. *Journal of Business and Industrial Marketing*, 31(3), 349-364.
- [11] Lysonski, S. & Durvasula, S. (2013). Nigeria in transition: acculturation to global consumer culture. *Journal of Consumer Marketing*, 30(6), 493-508.
- [12] Lorenz, M., Ramsey, J. & Glenn Richey, R. (2018). Expatriates' international opportunity recognition and innovativeness: The role of metacognitive and cognitive cultural intelligence. *Journal of World Business*, 53(2), 222-236. doi: 10.1016/j.jwb.2017.11.004
- [13] Marinković, V. (2017). Efekti animoziteta prema Evropskoj Uniji i patriotizma na potrošački etnocentrizam građana Srbije. *Ekonomski Horizonti*, 19(1), 3-15. doi: 10.5937/ekonhor1701003M
- [14] Nunnally, J. C. (1978). Introduction to Psychological Measurement. New York: McGraw-Hill
- [15] Nie, C. & Wang, T. (2021). How global brands incorporate local cultural elements to improve brand evaluations: A perspective on cultural mixing. *International Marketing Review*, 38(1), 163-183.
- [16] Prince, M., Davies, M.A.P., Cleveland, M. & Palihawadana, D. (2016). Here, there and everywhere: a study of consumer centrism. *International Marketing Review*, 33(5), 715-754.
- [17] Pentz, C., Terblanche, N. & Boshoff, C. (2017). Antecedents and consequences of consumer ethnocentrism: evidence from South Africa. *International Journal of Emerging Markets*, 12(2), 199-218.
- [18] Pratono, A.H. & Arli, D. (2020). Linking global consumer culture and ethnocentric consumerism to global citizenship: exploring the mediating effect of cultural intelligence. *International Journal of Sociology and Social Policy*, 40(7/8), 659-675. doi: 10.1108/IJSSP-10-2019-0212
- [19] Rahman, M.S., Abdel Fattah, F.A.M., Hussain, B. & Hossain, M.A. (2021). An integrative model of consumer-based heritage destination brand equity. *Tourism Review*, 76(2), 358-373.
- [20] Shimp, T.A. & Sharma, S. (1987). Consumer ethnocentrism: Construction and validation of CETSCALE. *Journal of Marketing Research*, 24(3), 280-289.
- [21] Shankarmahesh, M.M. (2006). Consumer ethnocentrism: An integrative review of its antecedents and consequences. *International Marketing Review*, 23(2), 146-172. doi: 10.1108/02651330610660065
- [22] Souiden, N., Ladhari, R. & Chang, L. (2018). Chinese perception and willingness to buy Taiwanese brands: The role of ethnocentrism and animosity. *Asia Pacific Journal of Marketing and Logistics*, 30(4), 816-836. doi: 10.1108/APJML-09-2017-0203
- [23] Shi, X. & Shan, X. (2019). Cross-cultural impact on financial companies' online brand personality. *Marketing Intelligence and Planning*, 37(5), 482-496. doi: 10.1108/MIP-06-2018-0233
- [24] Strizhakova, Y. & Coulter, R. (2019). Consumer cultural identity: local and global cultural identities and measurement implications. *International Marketing Review*, 36(5), 610-627.
- [25] Tuan, L.T. (2016). From cultural intelligence to supply chain performance. *The International Journal of Logistics Management*, 27(1), 95-121. doi: 10.1108/IJLM-01-2014-0009
- [26] Zdravković, S. (2021). Imidž zemlje porekla i potrošački ksenocentrizam u kontekstu formiranja stavova potrošača i lojalnosti prema stranim brendovima- moderatorski efekat kulturalne inteligencije. *Marketing- Časopis za marketing teoriju i praksu*, 52(1), 12-22. doi: 10.5937/markt2101012Z
- [27] Zdravković, S. & Peković, J. (2021). Cultural intelligence and heritage impact on choosing foreign tourist destination. *Hotel and Tourism Management*, 9(1), 27-42. doi:10.5937/menhottur2101027Z
- [28] Šapić, S. (2017). Efekti kosmopolitizma i tradicije na procene i namere korisnika usluga restorana brze hrane. *Ekonomski horizonti*, 19(2), 81-93. doi: 10.5937/ekonhor1702081S



# NAUKA O PODACIMA

---

# DATA SCIENCE



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business







## ALGORITHM FOR VISUAL ASSESSMENT OF THE CLUSTERING TENDENCY WITH SPECIAL DISTANCE MEASURES

GUZEL SHKABERINA<sup>1</sup>, LEV KAZAKOVTSSEV<sup>2</sup>, NATALIA REZOVA<sup>3</sup>, IVAN ROZHNOV<sup>4</sup>

<sup>1</sup> Reshetnev Siberian State University of Science and Technology – Institute of Informatics and Telecommunications  
z\_guzel@mail.ru

<sup>2</sup> Reshetnev Siberian State University of Science and Technology – Institute of Informatics and Telecommunications  
levk@bk.ru

<sup>3</sup> Reshetnev Siberian State University of Science and Technology – Institute of Informatics and Telecommunications  
natalyakl@yandex.ru

<sup>4</sup> Reshetnev Siberian State University of Science and Technology – Institute of Informatics and Telecommunications  
ris2005@mail.ru

---

**Abstract:** *In multidimensional data, the role of visualization in gaining knowledge about the structure of the data is critical. Assessing the clustering trend, i.e. determining if a data set has any inherent clusters and how many of them, is an important pre-clustering problem. The work is devoted to the study of methods and algorithms allowing to display information about the data internal structure before clustering. In our study, we propose a modified algorithm for visual assessment of the tendency to clustering based on the VAT and iVAT algorithms with different types of distance measures (Euclidean, Squared Euclidean, Mahalanobis, and Manhattan distances). For experiments, we used classical data sets from Machine Learning Repository, Basic Benchmark Repositories, and we considered the sample consisting of four different homogeneous batches of electronic radio components. Computational experiments demonstrate the highest efficiency of using the Euclidean and Squared Euclidean distances in VAT and iVAT.*

**Keywords:** *pre-clustering problem, distance measure, VAT, iVAT*



## IMPROVING QUALITY OF CONVOLUTIONAL NEURAL NETWORK MODELS FOR OFFLINE SIGNATURE VERIFICATION

SINIŠA STANIVUK<sup>1</sup>, DEJAN SIMIĆ<sup>1</sup>

<sup>1</sup> University of Belgrade – Faculty of Organizational Sciences, sinisa.stanivuk@gmail.com  
University of Belgrade – Faculty of Organizational Sciences, dejan.simic@fon.bg.ac.rs

**Abstract:** *With biometric technology becoming more prevalent in today's society, tasks like fingerprint or face recognition are being solved by using machine learning. One task particularly challenging and important is signature verification. The models need to learn specific minutiae that differentiate between genuine and forged signatures. This proves difficult if the forgeries are skillful. We introduce TripSigNet, a model based on the convolutional siamese networks, similar to SigNet model, differing in the loss employed. By introducing this change, we achieved the same accuracy results, but much faster and with greater quality. Training time has been improved by more than 10 times on a given dataset.*

**Keywords:** *Biometrics, Machine Learning, Neural Networks, Signature Verification*

### 1. INTRODUCTION

Biometric technology is becoming more prevalent in our modern society. Fingerprint scanning, face recognition, iris or retina scanning are just some of the more common tasks in biometrics. These approaches represent physiological characteristics, meaning each individual has different and unique face, iris, retina and fingerprint. Another group of tasks in biometrics relies on behavioral characteristics such as gait and signatures. Signatures are one type of characteristic that is easily attainable and vastly used for individual verification in real life.

Since 2012 [1], image processing models in machine learning have advanced immeasurably. They are used for different tasks such as image recognition [15], object detection [5], image generation [7] etc. Naturally, these models can also be used in biometric domain for tasks such as face recognition [16], but also signature verification [13]. We will investigate the performance of these models and present minor changes that made big improvements in performance and quality.

The remainder of the paper is structured as follows. The section 2 covers related works in domains such as convolutional neural networks, machine learning in biometrics and signature verification. Used methodology is presented in section 3, and section 4 describes experimental setup and achieved results. The conclusions are drawn in the final section.

### 2. RELATED WORK

#### 2.1. Convolutional Neural Networks

Convolutional neural networks (CNNs) are based on the convolution layers which are defined as convolutions of an input image and a convolution kernel. Convolution kernels are weight matrices which extract lower-level features at the starting layers, and higher-level features at the ending layers by combining lower-level features of previous layers [1]. Periodically, subsampling layers such as maximum or average pooling are inserted between convolution layers. Subsampling layers gradually reduce the size of the input image which in turn reduces the consumption of hardware resources and number of model parameters [1]. After extracting important features from an input image, CNN model transforms them into fully connected layers. These layers usually end with a softmax function which is used for classification.

LeCun et al. [2] were the first authors to introduce a deep CNN model used for image classification called LeNet-5. Unfortunately, due to insufficient data and hardware resources, it didn't attract enough attention [1]. In 2012, after winning the ILSVRC competition [3] by a large margin, the AlexNet [4] model popularized CNN models. Since then, the field of computer vision advanced significantly and branched into tasks like object detection [5], image segmentation [6], image generation [7], pose estimation [8] to name a few.

## 2.2. Biometrics and Machine Learning

Before applying machine learning to biometric data, there are certain factors that biometric characteristics being analyzed need satisfy. Those factors include universality, uniqueness, performance, permanence, measurability, acceptance, circumference [10]. Of these, performance and circumference are perhaps most important when employing machine learning on this type of data.

Most popular application of supervised learning involves face recognition [10]. Schroff et al. [15] developed a FaceNet model which used a novel loss function to create more expressive embeddings of face images used in downstream recognition and clustering tasks. Because of the preference for facial recognition to work in night-time scenarios, Sarfraz & Stiefelhagen [16] tried to create a model for mapping between visible and thermal image domains. Their deep neural network learned non-linear mapping from visible to thermal spectrum, which bridged a gap in previous performances by 40%. Another interesting application is iris recognition. Nguyen et al. [12] used CNN models pre-trained on ImageNet dataset to generate image features, which were further fed to a classification model. The model is a multi-class SVM trained in a one-against-all manner. We can conclude that more recently deep learning approaches dominate in the biometric domain.

## 2.3. Signature Verification

Previously stated applications of machine learning were based on physiological biometric data, but machine learning models can also be applied to behavioral data. Signature verification is a frequent application. We can divide signature verification into online and offline types [13]. Online signature verification relies on electronic writing devices such as pads and stylus pens. By using these, we can gather advanced information about a signature such as writing speed and even pressure. By combining all this information, we can create reliable systems for signature verification. Unfortunately, this approach can be costly, which is why offline signature verification exists. Offline signature verification relies exclusively on the 2D image of the signature, thus being more cost-friendly.

Dutta et al. [11] developed a model for writer independent offline signature verification based on highly correlated features of the signature. They extract important features from a signature, generate a codebook, code the features using nearest neighbor search, and then capture the distribution of image statistics of these features. After doing so for each image, they develop a kernel function to compute the similarity between any two signatures. The kernel produces a kernel matrix which an SVM model uses to make a final prediction.

Dey et al. [13] used an approach based on deep learning models. They developed a siamese CNN model for writer independent offline signature verification. SigNet accepts two images and compares them using the same model weights. This model achieved state-of-the-art performances on several benchmark datasets for signature verification.

## 3. METHODOLOGY

### 3.1. Loss Functions

Chopra et al. [14] introduced the contrastive loss, which is a type of loss function used for similarity learning. The authors advocate the usage of similarity learning when the number of classes is either very large or not precisely known during training. In this setting, the model tries to learn how to map similarities from input space into a lower-dimensional target space.

The definition of the contrastive loss is given below:

$$L(s_1, s_2, y) = (1 - y)D_w^2 + y \max(0, m - D_w^2), \quad (1)$$

where  $s_1$  and  $s_2$  represent image vector representations in the target space,  $y$  represents the binary label for genuine vs. forged signature.  $D_w$  represents Euclidean distance between two vector representations  $s_1$  and  $s_2$  produced by the weights  $w$  of the network. Finally,  $m$  is a parameter of the loss function called margin. The model should learn to separate genuine and forged signatures in the target space by at least margin  $m$ .

Schroff et al.[15] introduced the Triplet loss function in their FaceNet model, for the task of facial recognition. The definition of this version of triplet loss is given below:

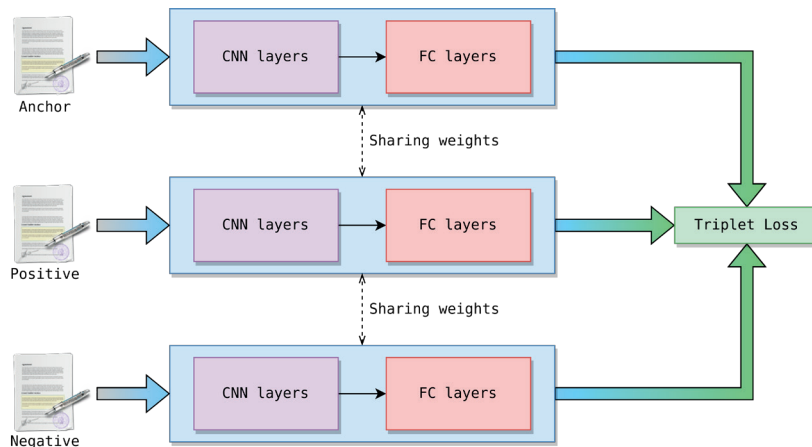
$$L(s_o, s_p, s_n) = \max(0, D_w(s_o, s_p) - D_w(s_o, s_n) + m), \quad (2)$$

where  $s_o$  is the representation of the original, genuine signature, called anchor,  $s_p$  is the representation of the positive class (another genuine signature), and  $s_n$  is the representation of the negative class (forged signature).  $D_w$  represents the Euclidean distance between two vector representations produced by the weights  $w$  of the network. Finally, same as in the contrastive loss,  $m$  is a parameter of the loss function called margin.

We argue that by employing the triplet loss instead of the contrastive loss, model training time can be reduced and the quality of the representations increased.

### 3.2. Models

SigNet [13] is a siamese CNN model used for signature verification. Siamese CNN models are type of CNNs which have twin heads consisting of convolutional layers which accept two input images. These two input images are analyzed in the same manner since twin heads share the same weights. The twin heads produce the final vector representations for both images. These vector representations are then sent as inputs to the loss function which computes how similar input images are.



**Figure 1** TripSigNet model achitecture

Our model, TripSigNet, has three input images corresponding to each of the inputs of triplet loss function. The basis is the same as for the SigNet [13] model, only difference being the three heads instead of two. The model's architecture can be seen in the **Figure 1**, while the architecture of one head can be seen in the **Table 1**.

Layers	Size	Parameters
Convolution	$96 \times 11 \times 11$	stride = 1
Local Response Norm	/	$k = 2, n = 5, \alpha = 10^{-4}, \beta = 0.75$
Pooling	$96 \times 3 \times 3$	stride = 2
Convolution	$256 \times 5 \times 5$	stride = 1, pad = 2
Local Response Norm	/	$k = 2, n = 5, \alpha = 10^{-4}, \beta = 0.75$
Pooling + Dropout	$256 \times 3 \times 3$	stride = 2, p = 0.3
Convolution	$384 \times 3 \times 3$	stride = 1, pad = 1
Convolution	$256 \times 3 \times 3$	stride = 1, pad = 1
Pooling + Dropout	$256 \times 3 \times 3$	stride = 2, p = 0.3
Fully Connected + Dropout	1024	p = 0.5
Fully Connected	128	/

**Table 1:** Architecture of one head as defined in the SigNet model

## 4. EXPERIMENTS AND RESULTS

### 4.1. Data

For evaluating the task of offline signature verification we used CEDAR<sup>1</sup> dataset - one of the most popular benchmark datasets for this task. The dataset is created in an experimental setup where each of the 55 signers would give out their genuine signatures 24 times every 20 minutes. Afterward, they would try to forge the signatures of 3 other signers, 8 times each. As a result, every signer would have 24 genuine and 24 forged signatures. Because of our hardware constraints, we selected this dataset since it's the smallest compared to other mainly used benchmark datasets [13].

We resize all images to  $105 \times 105$  using bilinear interpolation and invert pixel color so that the background is black and the signature white. We keep the first 50 signers' data as the training data, and the remaining 5 signers' data as the test data.

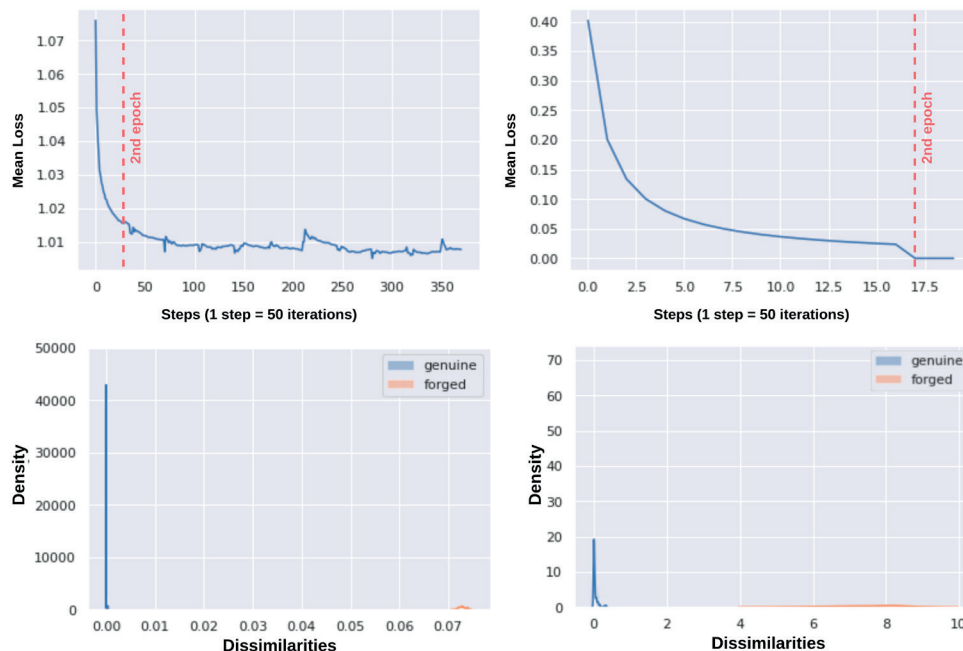
When forming triplets, Schroff et al. [15] argue that generating all possible triplets is a bad practice since some triplets easily satisfy the constraint of triplet loss<sup>2</sup>. We believe this is not true for our task since the main goal of forged signatures is to be as similar as possible to the genuine signatures.

### 4.2. Training

We compare and evaluate two models, SigNet [13] and our TripSigNet. We train both models for 20 epochs with Adam optimizer, learning rate of  $5 \times 10^{-5}$ , and in batches of 8 samples. We ran all experiments in the Google Colab setting which provides a free NVIDIA Tesla T4 GPU, 12GB of RAM, and 78GB of disk space.

### 4.3. Results

For evaluating the models, we used popular metrics such as accuracy, false acceptance rate (FAR) and false rejection rate (FRR). Beside these, we also evaluated training time and signature representation quality. We didn't use any explicit measure of representation quality, we analyzed it subjectively via plots of **Figure 2**.



**Figure 2** Plots depicting the differences between SigNet (left) and TripSigNet (right) models. Upper plots show loss function over time during training, and bottom plots show distributions of dissimilarities between original and genuine, and original and forged signatures.

<sup>1</sup> Can be downloaded from <https://cedar.buffalo.edu/NIJ/data/>

<sup>2</sup>  $D_w(s_o, s_p) + m < D_w(s_o, s_n)$

Even though SigNet model has the same accuracy, FAR and FRR metrics, as shown in **Table 2**, it's obvious that training time is much larger. We trained both SigNet and TripSigNet models for 10 epochs. SigNet training time lasted 146 minutes, and TripSigNet training time lasted around 10 minutes. The reason for this is because TripSigNet training actually stopped at the start of second epoch because the model was fully optimized. This can be seen in the upper right plot of **Figure 2** as loss reaches its optimum in 0.

Model	Accuracy	FAR	FRR	Training Time (minutes)
SigNet	100.00	0.00	0.00	146
TripSigNet	100.00	0.00	0.00	10

**Table 2:** Performance comparison of the SigNet and TripSigNet models

Comparing bottom two plots of **Figure 2**, we conclude that the quality of signature representations is much higher in our TripSigNet model. The quality is higher because of the larger gap between genuine and forged signature dissimilarities, which could prove beneficial in the future use of these models on real world data.

## 5. CONCLUSION

In this paper we presented an updated version of SigNet model called TripSigNet. We showed that by employing triplet loss instead of contrastive loss, the model can achieve the same performance much faster and have greater quality. We believe this is because the model can now simultaneously learn representations of genuine and forged signatures. In SigNet model this was separated as the model could be trained only using pairs of images. This improvement of quality should be beneficial in future uses of the model in the real world scenario. The future work may test TripSigNet model's capabilities against other benchmark datasets.

## REFERENCES

- [1] Wang, W., Yang, Y., Wang, X., Wang, W., & Li, J. (2019). Development of convolutional neural network and its application in image classification: a survey. *Optical Engineering*, 58(4), 040901.
- [2] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- [3] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. & Berg, A.C. (2015). ImageNet large scale visual recognition challenge. *International journal of computer vision*, 115(3), 211-252.
- [4] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [5] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- [6] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
- [7] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of StyleGAN. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 8110-8119).
- [8] Jin, S., Xu, L., Xu, J., Wang, C., Liu, W., Qian, C., Ouyang, W. & Luo, P. (2020, August). Whole-body human pose estimation in the wild. In *European Conference on Computer Vision* (pp. 196-214). Springer, Cham.
- [9] Ortiz, N., Hernández, R. D., Jimenez, R., Mauledeoux, M., & Avilés, O. (2018). Survey of biometric pattern recognition via machine learning techniques. *Contemporary Engineering Sciences*, 11(34), 1677-1694.
- [10] Jain, A., Bolle, R., & Pankanti, S. (Eds.). (1999). *Biometrics: personal identification in networked society* (Vol. 479). Springer Science & Business Media.
- [11] Dutta, A., Pal, U., & Lladós, J. (2016, December). Compact correlated features for writer independent signature verification. In *2016 23rd international conference on pattern recognition (ICPR)* (pp. 3422-3427). IEEE.
- [12] Nguyen, K., Fookes, C., Ross, A., & Sridharan, S. (2017). Iris recognition with off-the-shelf CNN features: A deep learning perspective. *IEEE Access*, 6, 18848-18855.

- [13] Dey, S., Dutta, A., Toledo, J. I., Ghosh, S. K., Lladós, J., & Pal, U. (2017). SigNet: Convolutional siamese network for writer independent offline signature verification. arXiv preprint arXiv:1707.02131.
- [14] Chopra, S., Hadsell, R., & LeCun, Y. (2005, June). Learning a similarity metric discriminatively, with application to face verification. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 1, pp. 539-546). IEEE.
- [15] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).
- [16] Sarfraz, M. S., & Stiefelwagen, R. (2017). Deep perceptual mapping for cross-modal face recognition. *International Journal of Computer Vision*, 122(3), 426-438.



## NEKI ASPEKTI DEMOKRATIZACIJE NAUKE O PODACIMA

### SOME ASPECTS OF DEMOCRATIZATION OF DATA SCIENCE

JASNA SOLDIĆ-ALEKSIĆ<sup>1</sup>, RADE STANKIĆ<sup>2</sup>, BILJANA CHRONEOS KRASAVAC<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, jasna.soldic@ekof.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, rade.stankic@ekof.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Ekonomski fakultet, biljana.krasavac@ekof.bg.ac.rs

**Rezime:** Cilj ovog rada je da prikaže neke aspekte kompleksnog procesa demokratizacije nauke o podacima. Ovo pitanje smo sagledali u kontekstu opštih trendova demokratizacije primene različitih informaciono-komunikacionih tehnologija. Kako je mašinsko učenje jedan od osnovnih alata nauke o podacima, posebno smo se bavili pitanjem razvoja različitih softverskih alata i platformi za automatizaciju mašinskog učenja (AutoML alati). Ali, pored AutoML alata, ukazali smo da process demokratizacije nauke o podacima uključuje aktivnosti koje se teško mogu automatizovati. Takođe, u radu smo sagledali stepen prihvatanja ideje demokratizacije nauke o podacima i načine njene realizacije u poslovnim organizacijama.

**Ključne reči:** demokratizacija nauke o podacima, automatizovano mašinsko učenje, automatizovana nauka o podacima, CRISP-DM metodologija

**Abstract:** The aim of this paper is to present some aspects of the complex process of democratization of data science. We have considered this issue in the context of general trends of democratization of the application of various information and communication technologies. As machine learning is one of the basic tools of data science, we have especially dealt with the development of various software tools and platforms for machine learning automation (AutoML tools). But in addition to the AutoML tools, we pointed out that the process of data science democratization involves activities that are difficult to automate. Also, in this paper, we looked at the degree of acceptance of the idea of democratization of data science and ways of its implementation in business organizations.

**Keywords:** data science democratization, automated machine learning, automated data science, CRISP-DM methodology

“Ignorance more frequently begets confidence than does knowledge”  
Charles Darwin

## 1. UVOD

Jedna od definicija nauke o podacima (*Data Science* – DS) polazi od toga da je reč o polju primene naprednih analitičkih tehnika i naučnih principa sa ciljem da se iz podataka izdvoje vredne informacije potrebne za donošenje poslovnih odluka i strateško planiranje. Na osnovu toga jasno je da primena nauke o podacima može da ima izuzetno važnu ulogu u poslovnom okruženju: informacije i znanje do koga se dolazi primenom tehnika i alata nauke o podacima mogu da pomognu organizaciji da poveća operativnu efikasnost, identifikuje nove poslovne mogućnosti, unapredi marketinške aktivnosti i planove prodaje, unapredi upravljanje finansijskim rizikom, na vreme otkrije nelegalne transakcije, obezbedi efikasnije lance snabdevanja, ojača zaštitu svojih informacionih sistema, a sve u cilju sticanja kompetitivnih prednosti u odnosu na poslovnu konkurenciju. Uobičajene aplikacije iz domena nauke o podacima odnose se na: prediktivno modeliranje, klasifikacije, kategorizacije i segmentacije, prepoznavanje obrazaca (*pattern recognition*), otkrivanje neobičnih opservacija (*anomaly detection*), analizu sentimenata (*opinion mining*), kao i razvoj tehnologija kao što su algoritmi preporuka (*recommendation engines*), personalizovani sistemi i alati veštačke inteligencije kao što su *chatbots*, AI asistenti, samoupravljujuća vozila i mašine, i dr.

S obzirom na svoju multisicplinarnost, kompleksnost tehnika i algoritama, rad sa malim i velikim obimom podataka (*small and big data*), kao i različitu prirodu podataka (struktuirani, polustruktuirani, nestruktuirani, *in-rest* vs *in-motion*) oblast nauke o podacima je puna izazova. Tipični izazovi u oblasti nauke o podacima odnose se na sledeće:

- Preuzimanje podataka iz različitih izvora i njihovo arhiviranje,
- Izbor adekvatnog analitičkog alata,
- Upravljanje kompleksnošću procesa analitičkog modeliranja,
- Suočavanje sa problemom pristrasnosti u podacima,
- Upravljanje procesom implementacije izabranog modela,
- Merenje poslovne vrednosti i/ili poslovnog uticaja primene modela,
- Održavanje modela.

Pored navedenih izazova pojavljuje se i pitanje raspoloživosti ljudskih resursa koji su osposobljeni da se bave ovom zahtevnom oblašću. Na prvom mestu to su naučnici podataka (*data scientists*), za koje ne postoji jedinstvena definicija poslova i kvalifikacija koje moraju da poseduju. Postavlja se pitanje da li se svim aktivnostima pod okriljem nauke o podacima može baviti jedan čovek – *data scientist*, ili je u pitanju timski rad većeg broja ljudi sa različitim kvalifikacijama (*data engineer, data analyst, machine learning engineer, data visualization developer, data translator, data architect*). Takođe, pitanje je da li je dovoljan jedan tim ljudi – „*centers of excellence*“ za sveobuhvatnu primenu koncepta nauke o podacima u poslovnom sistemu i/ili postoje i druga rešenja? Koliko može da pomogne strategija demokratizacije nauke o podacima u organizaciji? Razmišljajući o navedenim pitanjima, u radu smo pokušali da se približimo mogućim odgovorima.

## 2. OPŠTI TRENDVI DEMOKRATIZACIJE

Posmatrajući duži vremenski period – nekoliko prethodnih decenija možemo zapaziti mnoge primere demokratizacije primene različitih informaciono-komunikacionih tehnologija (personalni računari, internet, laptop računari, mobilni telefoni i slično). Ovaj trend demokratizacije je prisutan i danas, a povezan je sa sledećim fenomenima: fenomen *big data*, *open data* inicijativa, *open-source* softver i programski jezici, „*open*“ inicijative u oblasti posebnih industrijskih grana, na primer, *open banking* i sl. U tom smislu možemo govoriti o demokratizaciji podataka, prediktivne analitike, celokupne poslovne analitike, pa i demokratizaciji primene nauke o podacima. Polovinom prošle dekade vladalo je mišljenje da je „demokratizacija podataka nezaustavljiva sila u digitalnom poslovanju“, odnosno da „demokratizacija podataka i analitike je sledeća velika obećavajuća oblast za poslovni uspeh i ekonomiju zasnovanu na podacima“ [10]. Ovo mišljenje se dalje proširilo na pitanje demokratizacije nauke o podacima. Proces demokratizacije nauke o podacima znači da se ova oblast učini dostupnom ne samo ekspertima iz oblasti, već svima kojima se omogući pristup obimnoj količini podataka i analitičkim alatima koji imaju jednostavan korisnički interfejs.

Koliko je realan ovaj pristup? Postoje i liberalna i konzervativna razmišljanja povodom ovog pitanja. Zagovornici demokratizacije nauke o podacima smatraju da je u organizacijama neodrživ koncept da samo relativno mali broj zaposlenih ima pristup podacima i bavi se naukom o podacima. Po njima, naučnici podataka ne mogu da ispune sve obimnije zahteve koji se pred njih postavljaju u poslovnim sistemima, a takođe, nailaze na teškoće kako da adekvatno predstavljaju svoje zaključke zaposlenima koji ne poseduju odgovarajuće razumevanje oblasti rada sa podacima. Stoga je, po njihovom mišljenju, potrebno demokratizovati korišćenje podataka i svih alata za analitiku, pri čemu se podrazumeva obuka zaposlenih i uvođenje pozicije *citizen data scientist*. Naravno, u obuci zaposlenih može se poći od jednostavnijih DS algoritama za predikciju i klasifikaciju, a komplikovanije algoritme ostaviti u domenu naučnika podataka. Oponenti ovog mišljenja iznose stavove da je uvođenje *citizen data scientist*-a jednostavno iluzija, pozivajući se na tvrdnju da „malo znanja može biti veoma opasno“, odnosno na poznati fenomen iz psihologije – *Dunning-Kruger* efekat, koji opisuje kognitivnu pristrasnost, gde ljudi sa malim ograničenim znanjem iz određene oblasti imaju tendenciju da precene svoje znanje i kompetencije u tom domenu, u poređenju sa objektivnim kriterijumima. Pobornici ovog mišljenja navode primer da *citizen data scientist* neće prepoznati da „korelacija ne znači kauzalnost“, što u eri velikog obima podataka ima poseban značaj, jer se veoma lako može dogoditi da se uoče sumnjive korelacije (*spurious correlations*), i da se na osnovu toga generišu pogrešni zaključci. Stoga, ova grupa autora i istraživača se zalaže da se ne pristupa demokratizaciji nauke o podacima.

Ako pogledamo realnost, uočljiva je značajna tendencija u softverskoj industriji u pravcu razvoja različitih softverskih alata i platformi koje doprinose demokratizaciji nauke o podacima. Većina tih alata pripada grupi alata za automatizovano mašinsko učenje. U nastavku ćemo se baviti pitanjem automatizovanog mašinskog učenja i automatizove nauke o podacima.

## 3. AUTOMATIZOVANO MAŠINSKO UČENJE VS AUTOMATIZOVANA NAUKA O PODACIMA

Između dva koncepta - mašinskog učenja (*Machine Learning* – ML) i nauke o podacima (*Data Science* – DS) postoji značajna razlika. Nauka o podacima se odnosi na aplikacije naučnog metoda na veoma široki

koncept izdvajanja znanja i bitnih sadržaja (*insight*) iz podataka. Mašinsko učenje se odnosi na „pitanje kako kreirati računarski program koji se automatski unapređuje sa iskustvom“ [5].

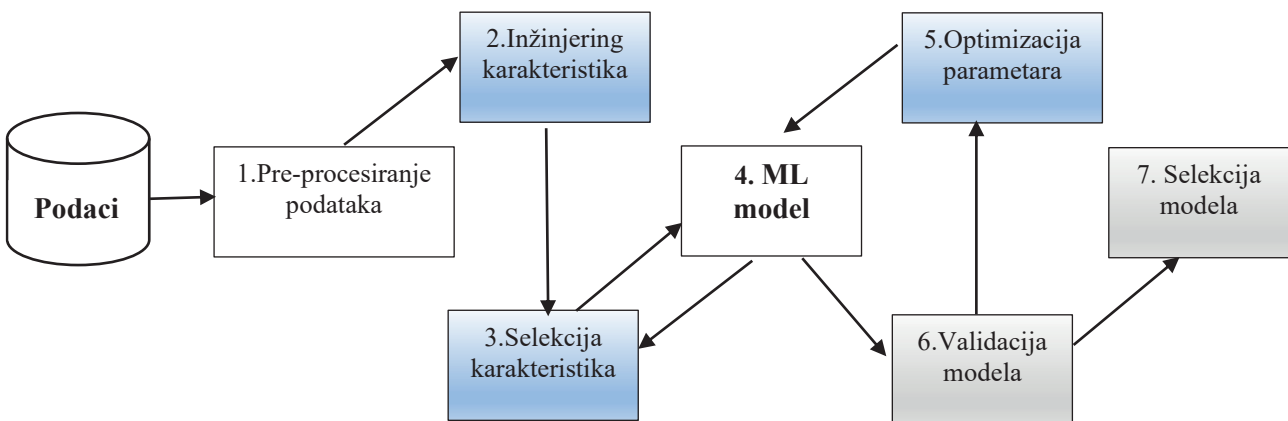
Po svojoj prirodi mašinsko učenje je interdisciplinarno, jer uključuje tehnike iz oblasti statistike, računarskih nauka, veštačke inteligencije i dr. Mašinsko učenje je samo jedna vrsta alata koje koriste naučnici podataka, ali nije preterano reći da predstavlja centralni deo nauke o podacima. Naime, ako imamo u vidu da je suština DS da se izdvoje bitne informacije (*insight*) iz podataka, onda se može reći da je ML mašina ili motor koji omogućava da se taj proces izdvajanja automatizuje.

Poslednjih godina uočava se u literaturi trend povećane zainteresovanosti istraživača, ali i kompanija koje rade na razvoju softvera, ka automatizaciji ova dva koncepta. Šta znači automatizovano mašinsko učenje, odnosno automatizovana nauka o podacima? Postoje mišljenja da se radi o istom postupku, ili u krajnjoj liniji o značajnom preklapanju dva pojma. Ali kao što smo videli iz definicija, postoje značajne razlike u konceptima, tako da postoje razlike i u njihovoj automatizaciji.

Automatizovano mašinsko učenje (AutoML) ima za cilj da automatizuje sve ili samo neke korake procesa mašinskog učenja, što uključuje sledeće:

- Pre-procesiranje podataka (*data pre-processing*) – uključuje unapređenje kvaliteta podataka, prečišćavanje podataka, integraciju podataka, transformaciju i redukciju podataka,
- Kreiranje bitnih karakteristika (*feature engineering* i *feature extraction*) – iz podataka koje će obezbediti preciznije rezultate i smanjiti obim podataka koji se obrađuju,
- Selekciju karakteristika (*feature selection*) – AutoML može da automatizuje selekciju samo korisnih atributa,
- Selekciju modela (*model selection*) – izbor modela,
- Optimizaciju parametara/hiperparametara (*hyperparameter optimization*) – podešavanje hiperparametara izabranih algoritama,
- Validaciju modela.

Prema autoru Mayo, M. proces automatizacije mašinskog učenja podrazumeva automatizaciju jezgra ovog procesa, a to znači pre svega automatizaciju sledećih postupaka, tj. faza ili koraka: optimizaciju postupka izdvajanja i inženjeringa karakteristika (*feature engineering* i *feature selection*) i optimizaciju hiper/parametara (*hyperparameter optimization*) [4]. Pored ovih koraka, moguća je i poželjna i automatizacija validacije modela i selekcije modela. Na slici 1 prikazan je životni ciklus projekta mašinskog učenja i istaknute su faze koje podležu automatizaciji (redni broj 2, 3 i 5, a dodatno je moguća automatizacija faza 6 i 7).



**Slika 1.** Životni ciklus projekta mašinskog učenja

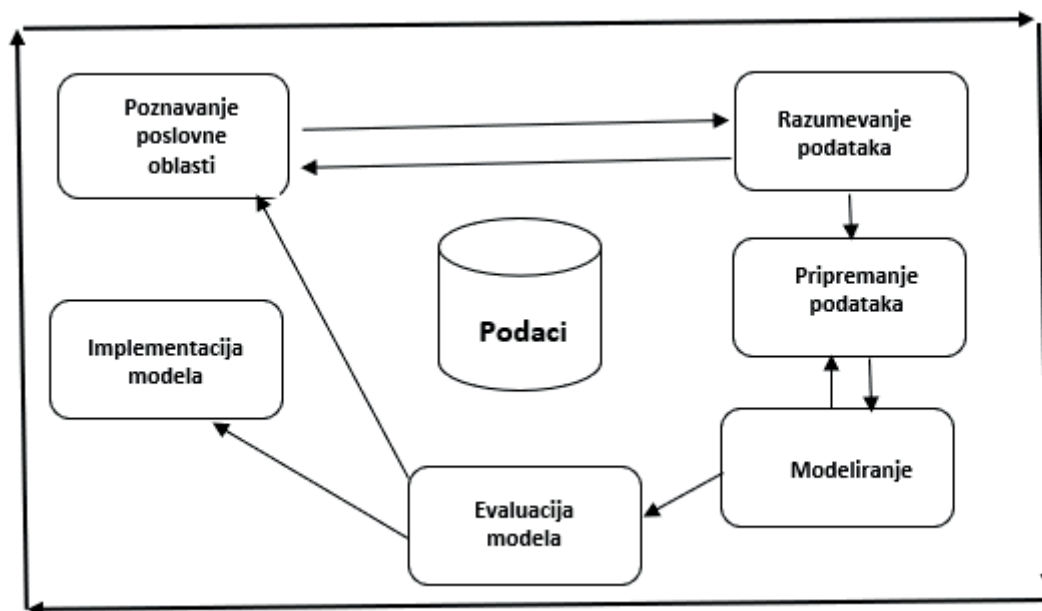
Izvor: Obradeno prema tekstu Mayo, M. [4]

Očigledno je da kreiranje AutoML alata potiče od sledeće ideje: ako se mora kreirati veliki broj ML modela, koristeći različite algoritme i različite konfiguracije hiper/parametara, onda postupak gradnje modela može biti automatizovan, kao i poređenje performansi i tačnosti modela [3]. Randy Olson (kreator AutoML alata TPOT), istakao je da takvi AutoML alati nisu zamena za naučnika podataka, već su viđeni kao njihovi asistenti (TPOT je nazvao „*your data science assistant*“)[6]. Dalje, on ističe da efektivni ML dizajn zahteva od čoveka sledeće:

- Konstantno podešavanje hiperparametara izabranih modela,
- Neprekidno isprobavanje mnogih različitih modela,

- Neprekidno istraživanje brojnih karakteristika podataka (*feature representations*).

Proces automatizacije nauke o podacima je mnogo širi u odnosu na automatizaciju mašinskog učenja, i može se posmatrati kao kišobran koncept, koji pokriva sledeće: korišćenje alata za automatizovano mašinsko učenje (AutoML), kao i automatizaciju drugih zadataka, kao što su: prikupljanje podataka (*data collection*), pripremanje podataka za analizu (*data preparation*), eksplorativnu analizu podataka i vizualizaciju podataka, interpretaciju rezultata. Dakle, AutoDS ima za cilj da pokrije širi proces, koji se može predstaviti poznatim metodološkim okvirom iz područja napredne pretrage podataka (*data mining-a*), CRISP-DM (*The Cross-Industry Standard Process for Data Mining - CRISP-DM*) – (slika 2). Naime ova metodologija se često koristi kao vodič za realizaciju procesa nauke o podacima u praksi.



Slika2. Ilustracija CRISP-DM metodologije

AutoML ima potencijal da značajno poveća produktivnost naučnika koji se bavi podacima. Može se reći dok se najveći deo ML procesa može automatizovati, to ne važi za DS proces. Osnovni razlog je taj da mnogi poslovni problemi pripadaju grupi nestruktuiranih i/ili slabo struktuiranih problema, koje je teško i definisati. Dodatno, DS proces uključuje znanje o poslovnoj oblasti, definisanje problema, ali i predstavljanje rezultata – komunikaciju sa korisnicima koji ne moraju da poznaju bilo koji segment rada sa podacima (veština *data storytelling*). Takođe, veoma je bitna činjenica da se u poslovnoj praksi, značaj naučnika podataka meri na bazi poslovne vrednosti koju stvaraju, pre nego na osnovu tačnosti, preciznosti i drugih metrika koje važe za ocenu performansi algoritama. Tako, na primer, projekat u okviru koga se rešava problem klasifikacije može da ostvari klasifikacionu preciznost od 99%, ali ako nije adekvatno implementiran u rešava je poslovnog problema, možda ne donosi poslovnu dobit kompaniji.

Kakva je trenutna situacija na tržištu AutoML alata? Tržište AutoML alata generisalo je prihod od 270 miliona dolara 2019 godine, a očekuje se da do 2030 godine ostvari prihod od 14,5 milijardi dolara, čime se predviđa godišnja stopa rasta od 43,7% u periodu 2020-2030. Imajući to u vidu, smatra se da ova industrija nije dostigla svoj maksimum, već se očekuje njen dalji ubrzani razvoj [2].

#### 4. ML ALATI I PLATFORME ZA DEMOKRATIZACIJU NAUKE O PODACIMA

U prethodnom delu ukazali smo na ključne osobenosti i razlike između automatizovanog mašinskog učenja i automatizovanog procesa nauke o podacima, te je jasno da AutoML alati mogu da doprinesu široj primeni algoritama mašinskog učenja, što znači i samoj demokratizaciji nauke o podacima. Danas postoji široka lepeza raznovrsnih AutoML alata i platformi. Jedna od najstarijih AutoML platformi jeste H2O.ai (pojavi se 2012 godine), posle koje se 2015 pojavljuje TPOT, a kada su par godina kasnije tehnološki giganti Amazon Web Service - AWS, Google Cloud i Microsoft Azure predstavili svoje AutoML alate, popularnost i interes za ovu softversku grupu alata i usluga naglo su porasli. U nastavku dajemo kraći prikaz nekih poznatijih ML alata i platformi, koji su obeležili period ML automatizacije.

**H2O.ai** je jedna od vodećih AI platformi u oblaku, za koju se može reći da ima misiju upravo da demokratizuje veštačku inteligenciju za svakoga („*AI for everyone*“). Postoje dve verzije: *open-source* paket i, od 2017 godine,

komercijalni proizvod koji pruža AutoML usluge - *Driverless AI*. Od svog nastanka H2O je stekao veliku popularnost i komercijalnu primenu, posebno u oblasti finansijskih usluga i maloprodaje.

Već smo naveli paket **TPOT**, koji je razvijen na Univerzitetu u Pensilvaniji, Institutu za Biomedicinsku informatiku. Reč je o *open-source* Python DS automatizovanom alatu, koji vrši optimizaciju *ML pipelines* koristeći genetsko programiranje. Jedan primer primene ovog alata (ML klasifikator) je selekcija atributa za dati skup podataka gde se koristi više algoritama za izbor atributa, a zatim na kombinovani niz svih atributa primenjuje se *Random Forest* klasifikator.

**Amazon** se uključio u utakmicu kreiranja AutoML alata sa proizvodom *AutoGluon*. Reč je o *open-source* AutoML biblioteci, tj. kolekciji alata koja pruža korisnicima transparentan način da vizuelno prate sve faze ML procesa, kao i mogućnost da uključe svoje algoritme u taj proces. Veoma popularan ML alat je *Amazon SageMaker Autopilot*, koji automatski kreira ML modele za dati niz podataka, pri čemu obezbeđuje korisniku potpuno praćenje i vizualizaciju procesa.

**Google Cloud AutoML** se pojavio 2018 godine i veoma brzo stekao popularnost zahvaljujući veoma jednostavnom i za korisnika prihvatljivom korisničkom interfejsu i visokim performansama. Google je ponudio niz proizvoda za rad sa struktuiranim, polustruktuiranim i nestruktuiranim podacima (tekst, slike, video i sl.), kao što su: *Vertex AI*, *AutoML Tabular*, *AutoML Image*, *AutoML Video*, *AutoML Text*, *AutoML Translation*.

**Microsoft Azure AutoML**, koji se takođe pojavio 2018 godine, pruža transparentan proces selekcije modela svojim korisnicima, koji ne moraju da poznaju programiranje. Važno je istaći da je u ovu platformu uključen MLOps (*Machine Learning Operations*) koncept, tj. DevOps za mašinsko učenje, koji omogućava IT timu da saraduje, automatizuje i ubrzava ML životni ciklus.

**DataRobot** je napredna AI platforma koja pruža mogućnost automatizacije DS procesa preko raznovrsnih AutoML funkcionalnosti. Platforma je opremljena sa savremenim *open-source* algoritmima, počev od klasične regresije, kompleksnih višeklasnih klasifikatora pa sve do *deep learning* algoritama. Na raspolaganju su verzije u oblaku i *on-premise*.

**Salesforce Einstein platforma** se pojavila 2016 godine i veliku popularnost je stekla zahvaljujući mogućnostima vezanim za AutoML funkcionalnosti, koje su poznate pod imenom *TransmogriAI*. Reč je o AutoML biblioteci namenjenoj za struktuirane podatke, a napisana je na programskoj jeziku Scala. Einstein platforma, zahvaljujući AutoML uslugama obavlja veliki broj predikcija – modelira prodaju i marketinške aktivnosti (80 milijardi predikcija na dan), čime je postala u svetskim okvirima jedan od najboljih primera primene AutoML alata u praksi.

## 5. DEMOKRATIZACIJA NAUKE O PODACIMA U ORGANIZACIJI

Na koji način će biti prihvaćena nauka/naučnici o podacima u poslovnim organizacijama? U odgovoru na ovo pitanje možemo se podsetiti činjenice da opšti stepen prihvatljivosti mnogih tehnološkim novina u poslovnim sistemima se može objasniti teorijom širenja inovacija (*Diffusion of Innovation Theory*). Prema ovoj teoriji stepen prihvatanja inovacija može se uslovno predstaviti približno normalnom raspodelom, gde se u korišćenju novih tehnologija najpre pojavljuju malobrojni inovatori praćeni sa nešto većim broj prvih korisnika (*early adopters*), a zatim slede dva talasa prihvatanja inovacija od strane rane većine (*early majority*), odnosno kasnije većine korisnika (*late majority*), da bi na kraju inovacije bile prihvaćene od strane najkonzervativnijih pojedinaca koji su generalno skeptični prema inovacijama [8]. Takođe, na prihvatanje koncepta nauke o podacima u organizacijama veoma je bitna konkretna organizaciona struktura i kultura kompanija. Međutim, nezavisno od toga, generalno se može reći da *data-intensive* kompanije imaju veliki potencijal za izdvajanje bitnih sadržaja (*insight*) iz podataka uz pomoć nauke o podacima. Za mnoge takve kompanije vezuje se ideja sprovođenja *data-science* transformacije na nivou cele organizacije. Da bi se ta ideja što uspešnije realizovala, *Redman* i *Davenport* predlažu da se učini da nauka o podacima dobije strateški značaj u poslovnom sistemu i da se radi na njenoj demokratizaciji [7]. Njihovo mišljenje je da mnoge organizacije koriste *data science* koncept time što formiraju „*centers of excellence*“, što znači da na jednom mestu koncentrišu raspoložive ljudske resurse – naučnike podataka, kako bi generisali što bolje rezultate. Međutim, u praksi ovakav pristup nije pružio najbolje rezultate. Stoga ovi autori ukazuju na sledeća četiri pravca aktivnosti u kompanijama:

- Fokusirati se na probleme koji mogu doneti strateške koristi,
- Raditi na demokratizaciji celokupnog koncepta nauke o podacima u organizaciji,
- Napraviti prioritete u rešavanju problema nauke o podacima i dodeljivanju zadataka adekvatnom osoblju,
- Raditi na razvoju i širenju vizije o primeni nauke o podacima u celoj kompaniji.

U skladu sa prethodnim, i mnogi drugi autori govore o potrebi demokratizacije primene nauke o podacima u poslovnim organizacijama. Njihov glavni argument u prilog ove teze jeste postojanje velikog obima podataka, sa jedne strane, i relativno malih grupa ljudi koji mogu da se bave tim podacima u organizacijama

(*data scientists*), sa druge strane. Ovaj nesklad, odnosno jaz može postati neodrživ. Stoga se nameće rešenje: raširiti koncept nauke o podacima na celu organizaciju. Naime, kompanije koje žele da aktivno učestvuju u tržišnoj utakmici u vreme koje se označava kao „*age of data*“, bilo bi dobro da razmišljaju o sledećem: da rade na disperziji korišćenja alata za analitiku podataka, da rade na obuci i osposobljavanju zaposlenih za rad sa podacima i da razvijaju i šire kulturu odgovornosti u radu sa podacima [1]. Kao dobar primer u tom smislu može poslužiti kompanija Airbnb, koja se opredelila za strateško uključanje koncepta nauke o podacima i demoktizaciju njene primene u svom poslovanju. Kompanija je postavila cilj da svakog radnika osposobi za donošenje odluka na bazi podataka (*data-driven decisions*). Da bi ostvarila postavljeni cilj kompanija je kreirala svoj Univerzitet podataka (*Data University*), kao i Repozitorijum znanja (*Knowledge Repository*), preko koga se promovišu nova znanja i iskustva zaposlenih se čine dostupna svima. Iskustva i drugih kompanija su pokazala da nauka o podacima ne mora biti namenjena samo naučnicima podataka, već „pametne“ kompanije (*smart companies*) se trude da njihni zaposleni razumeju i koriste jezik podataka (da postanu *data-literate*), kako bi otvorili prostor za inovacije i povećali efikasnost poslovanja.

## 6. ZAKLJUČAK

U ovom radu smo pokušali da sagledamo samo neke aspekte kompleksnog procesa demokratizacije nauke o podacima. Istakavši trendove demokratizacije, posebno smo ukazali na pojavljivanje softverskih alata i platformi koje mogu da doprinesu demokratizaciji nauke o podacima. - alati za automatizovano mašinsko učenje. Cilj primene AutoML alata jeste da oslobode naučnike podataka repetitivnih i vremenski zahtevnih zadataka, kao što je, na primer, optimizacija hiper/parametara modela, tako da se oni mogu posvetiti zadacima koji se teže mogu automatizovati. Posebno smo ukazali da su alati mašinskog učenja postali integralni deo *data science* protoka i obrade podataka (*data science pipeline*). Takođe, pokušali smo da sagledamo stepen i načine realizacije koncepta demokratizacije nauke o podacima u poslovnim organizacijama. Mada ima *pro et contra* mišljenja povodom ove ideja, činjenica je da u mnogim organizacija (posebno *data-intesive* organizacijama) ova ideja se stavlja u fokus poslovne strategije.

## LITERATURA

- [1] Cornelissen, J., (2018). The Democratization of Data Science, *Harvard Business Review*, Analytics and Data Science, July, 2018, <https://hbr.org/2018/07/the-democratization-of-data-science>.
- [2] Dilmegani, C., (2022). AutoML: In depth Guide to Automated Machine Learning [2022], *AI Multiple*, Automated Machine Learning posts, June 14, 2022. <https://research.aimultiple.com/auto-ml/>.
- [3] Mayo, M., (2017). The Current State of Automated Machine Learning, *KDnuggets* on January 18, 2017 in Automated, Automated Data Science, Automated Machine Learning, Retrieved from: <https://www.kdnuggets.com/2017/01/current-state-automated-machine-learning.html>.
- [4] Mayo, M., (2018). Automated Machine Learning vs Automated Data Science, *KDnuggets* on July 2, 2018, <https://www.kdnuggets.com/2018/07/automated-machine-learning-vs-automated-data-science.html>.
- [5] Mitchell, T., (1997). *Machine Learning*, 1<sup>st</sup> ed., McGraw Hill Education.
- [6] Olson, R.S., Bartley, N., Urbanowicz, R.J., & Moore, J. (2016). Evaluation of a Tree-based Pipeline Optimization Tool for Automating Data Science. *Proceedings of the Genetic and Evolutionary Computation Conference*, July 2016, Pages 485–492, <https://doi.org/10.1145/2908812.2908918>.
- [7] Redman, T. C, Davenport, T. H. (2021). 4 Ways to Democratize Data Science in Your Organization, *Harvard Business Review*, Analytics and data Science, March, 2021., <https://hbr.org/2021/03/4-ways-to-democratize-data-science-in-your-organization>.
- [8] Reynolds, G.W. (2016), *Information Technology for Managers*, 2ed. Cengage Learning, Boston, USA
- [9] Soldic, J., Chronos Krasavac, B. (2018). Demokratizacija podataka i poslovne analitike, *Zbornik radova, XLV Symposium on Operational Research, SYM-OP-IS 2018*, Zlatibor, 16-18, 09. 2018, 228-234.
- [10] Srujana, H.M., Sanjay S. Sharma & Amitava, D. (2016). Democratization of analytics: New frontier of data economy, *Analytics Magazine*, INFORMS, March 7, 2016, <https://doi.org/10.1287/LYTX.2016.02.02>

## PRIMENA STABALA ODLUČIVANJA U ANALIZI UZROKA I VRSTA SAOBRAĆAJNIH NEZGODA NA TERITORIJI REPUBLIKE SRBIJE

### APPLICATION OF DECISION TREES IN THE ANALYSIS OF CAUSES AND TYPES OF TRAFFIC ACCIDENTS ON THE TERRITORY OF THE REPUBLIC OF SERBIA

ALEKSA MAKSIMOVIĆ<sup>1</sup>, VLADICA STOJANOVIĆ<sup>2</sup>

<sup>1</sup> Sektor za analitiku, telekomunikacione i informacione tehnologije MUP-a RS, maksimovic1919@gmail.com

<sup>2</sup> Kriminalističko-policijski univerzitet – Departman informatike i računarstva, Beograd, vladica.stojanovic@kpu.edu.rs

**Rezime:** U radu je ispitan uticaj i značaj određenih faktora na oblik i težinu saobraćajnih nezgoda koje su se desile na teritoriji Republike Srbije u periodu od 2015. do 2021. godine. Nad prikupljenim podacima su implementirani algoritmi nadgledanog mašinskog učenja, klasterizacije i klasifikacije, u cilju učenja iz podataka i uočavanja obrazaca unutar njih. U cilju kvalitetnije obrade i analize, originalni podaci su transformisani u pogodniji oblik, a kao glavni rezultat istraživanja, predstavljeno je stablo odlučivanja o uticaju uočenih atributa na vrstu, težinu i oblik saobraćajnih nezgoda.

**Ključne reči:** Stabla odlučivanja, bezbednost saobraćaja, mašinsko učenje, klasterizacija, klasifikacija.

**Abstract:** The paper examines the impact and significance of certain factors on the form and severity of traffic accidents that occurred on the territory of the Republic of Serbia in the period from 2015 to 2021. Algorithms of supervised machine learning, clustering and classification have been implemented over the collected data, to learn from the data and observe the patterns within them. To better process and analyse, the original data were transformed into a more suitable form, and as the main result of the research, a decision tree on the impact of the observed attributes on the type, severity and shape of traffic accidents was presented.

**Keywords:** Decision trees, traffic safety, machine learning, clustering, classification.

## 1. UVOD

Moderan način življenja i užurbani stil života dovodi do toga da danas automobil postaje potreba, a nikako luksuz. Veliki broj automobila neminovno uslovljava i porast broja saobraćajnih nezgoda, iako se uvođenjem novih zakona i strogih metoda kontrole saobraćaja neprekidno nastoji da se poveća opšta bezbednost u saobraćaju. No, da li je zaista tako? Da li smo vremenom, u eri informacionih tehnologija i pametnih automobila zaista sve bezbedniji kao učesnici u saobraćaju? Da li je moguće dodatno doprineti bezbednosti saobraćaja? Da li je moguće uočiti određene zakonitosti koje dovode do saobraćajnih nezgoda sa različitim vrstama ishoda? Da li su neki faktori bitniji od drugih i koliko? Ovaj rad će pokušati da odgovori na neka od ovih pitanja. Preciznije rečeno, predmet istraživanja jeste učenje iz podataka o saobraćajnim nezgodama evidentiranim na teritoriju Republike Srbije u periodu od 01.01.2015. do 31.12.2021. godine.

U cilju određivanje važnosti svake od odlika koje karakterišu saobraćajnu nezgodu i njihovu klasifikaciju, korišćene su neke od metoda mašinskog učenja, odnosno *metode pretrage i otkrivanja znanja u podacima*. Poseban akcenat dat je nekim od najčešće korišćenih tehnika otkrivanja znanja, koje imaju itekako veliku i značajnu primenu u praksi. Reč je, pre svega, o tehnikama *heurističkog tipa*, odnosno postupcima *segmentacije (klasterizacije), klasifikacije i predviđanja* koje daju ogroman doprinos savremenoj analitici podataka [6]. Ove tehnike omogućavaju formiranje posebnih *logički povezanih grupa podataka (klastera)*, kao i tzv. *stabala odlučivanja* (engl. *Decision Trees*). Stabla odlučivanja predstavljaju jedan od najmoćnijih i najpopularnijih alata učenja i predikcije na osnovu pretrage velikih skupova podataka. Ona istovremeno predstavljaju jednostavan grafički alat za podršku odlučivanje i često se koriste u operacionim istraživanjima, posebno u analizi odluka i identifikaciji optimalne strategije koja će najverovatnije dostići cilj, a u zavisnosti od slučajeva koji utiču na nju [1], [4]. Upravo konstrukcija stabla odlučivanja koje ukazuje na mogućnost predviđanja oblika i težine saobraćajnih nezgoda (tzv. klasa), a na koje može uticati čitav niz faktora koji su takođe analizirani i formirani u obliku tzv. nominalnih varijabli (atributa), jeste osnovni cilj ovog istraživanja.

## 2. OSNOVNI MODEL PODATAKA

Kao što je već rečeno u uvodnom delu, u radu su korišćeni podaci koji sadrže informacije o ukupno 217.191 saobraćajnih nezgoda na teritoriji Republike Srbije, u periodu od početka 2015. pa sve do kraja 2021. godine. Reč je o tzv. „sirovim“ podacima predstavljenim tabelarno u tzv. *Open Data Source (.ODS)* formatu, a koji su dostupni u elektronskom obliku na Web portalu otvorenih podataka Republike Srbije [7]. Osnovne atribute, tj. kolone unutar tabela predstavljaju, redom, jedinstveni ID broj nezgode, policijsku upravu, opštinu, datum i vreme saobraćajne nezgode, geo-lokacija mesta saobraćajne nezgode, vrstu saobraćajne nezgode (sa materijalnom štetom, povređenim ili poginulim licima), naziv, odnosno oblik saobraćajne nezgode, broj vozila - učesnika u saobraćajnoj nezgodi, status lica koja su učestvovala (povređeni, poginuli), kao i detaljni opis saobraćajne nezgode. Na slici 1 dat je prikaz ovog skupa podataka, konkretno, za 2017. godinu.

1	2	3	4	5	6	7	8	9
1164470	BEGRAD	BARAJEVO	08.01.2017.05:00	03.3885	44.55065	Sa mat.stetom	SN SA JEDINIM VOZILOM	Nezgodе sa učešćem jednog vozila i životinja
1164587	BEGRAD	BARAJEVO	09.01.2017.12:45	03.304345	44.570626	Sa mat.stetom	SN SA PARKIRANIM VOZILIMA	Ostali sudari sa parkiranim vozilom
1165056	BEGRAD	BARAJEVO	09.01.2017.11:20	03.379753	44.630332	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – SKRETANJE ILI PRELAZAK	Najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje, polukružno okretanje ispred drugog vozila
1166560	BEGRAD	BARAJEVO	09.01.2017.11:30	03.30813	44.57349	Sa povredjenim	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila – čeon sudar
1164594	BEGRAD	BARAJEVO	09.01.2017.09:40	03.3696	44.61692	Sa mat.stetom	SN SA JEDINIM VOZILOM	Nezgodа sa jednim vozilom – silazak udesno sa kolovoza na pravcu
1164452	BEGRAD	BARAJEVO	10.01.2017.07:30	03.41846	44.57097	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila – čeon sudar
1166161	BEGRAD	BARAJEVO	16.01.2017.15:30	03.2991	44.55503	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Ostale nezgode sa učešćem najmanje dva vozila bez skretanja (bez podataka o smeru)
1166218	BEGRAD	BARAJEVO	18.01.2017.13:10	03.4063	44.54152	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez skretanja
1166501	BEGRAD	BARAJEVO	19.01.2017.13:40	03.39906	44.59506	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez skretanja
1165176	BEGRAD	BARAJEVO	20.01.2017.16:55	03.41131	44.61404	Sa poginulim	SN SA PEŠACIMA	Prelazak pešaka izdesno, van raskrsnice, bez skretanja vozila
1165059	BEGRAD	BARAJEVO	20.01.2017.07:40	03.2986	44.55812	Sa povredjenim	SN SA NAJMANJE DVA VOZILA – SKRETANJE ILI PRELAZAK	Najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje ulevo ispred drugog vozila
1165434	BEGRAD	BARAJEVO	22.01.2017.20:30	03.427796	44.57906	Sa mat.stetom	SN SA JEDINIM VOZILOM	Nezgodе sa učešćem jednog vozila i životinja
1166291	BEGRAD	BARAJEVO	23.01.2017.13:55	03.413526	44.5798	Sa mat.stetom	SN SA PARKIRANIM VOZILIMA	Ostali sudari sa parkiranim vozilom
1166083	BEGRAD	BARAJEVO	25.01.2017.16:30	03.30358	44.59267	Sa mat.stetom	SN SA JEDINIM VOZILOM	Nezgodа sa jednim vozilom – silazak udesno sa kolovoza na pravcu
1166103	BEGRAD	BARAJEVO	25.01.2017.02:00	03.30159	44.49991	Sa mat.stetom	SN SA PARKIRANIM VOZILIMA	Sudar sa parkiranim vozilom sa desne strane kolovoza
1166447	BEGRAD	BARAJEVO	31.01.2017.15:15	03.37942	44.54337	Sa poginulim	SN SA PARKIRANIM VOZILIMA	Sudar sa parkiranim vozilom sa desne strane kolovoza
1166978	BEGRAD	BARAJEVO	03.02.2017.22:10	03.36284	44.61513	Sa povredjenim	SN SA NAJMANJE DVA VOZILA – SKRETANJE ILI PRELAZAK	Najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje ulevo ispred drugog vozila
1168066	BEGRAD	BARAJEVO	10.02.2017.07:15	03.41814	44.57517	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez skretanja
1167998	BEGRAD	BARAJEVO	13.02.2017.14:00	03.415245	44.502434	Sa mat.stetom	SN SA JEDINIM VOZILOM	Nezgodа sa jednim vozilom na kolovozu
1168002	BEGRAD	BARAJEVO	13.02.2017.16:00	03.41328	44.60614	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila koja se kreću u istom smeru – preticanje
1169126	BEGRAD	BARAJEVO	16.02.2017.16:20	03.381955	44.633166	Sa povredjenim	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila koja se kreću u istom smeru – susretanje
1168148	BEGRAD	BARAJEVO	17.02.2017.05:00	03.361684	44.613804	Sa poginulim	SN SA PEŠACIMA	Prelazak pešaka sleva, van raskrsnice, bez skretanja vozila
1168257	BEGRAD	BARAJEVO	17.02.2017.21:25	03.43572	44.60663	Sa povredjenim	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila koja se kreću u istom smeru – susretanje
1169156	BEGRAD	BARAJEVO	19.02.2017.02:30	03.39603	44.63261	Sa mat.stetom	SN SA JEDINIM VOZILOM	Nezgodа sa jednim vozilom – silazak ulevo sa kolovoza na pravcu
1169248	BEGRAD	BARAJEVO	20.02.2017.08:10	03.36853278	44.62181578	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – SKRETANJE ILI PRELAZAK	Ostale nezgode sa najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje
1169577	BEGRAD	BARAJEVO	20.02.2017.11:30	03.41513	44.58301	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila – suprotni smerovi bez skretanja – kretanje unazad
1168777	BEGRAD	BARAJEVO	20.02.2017.14:50	03.3820542	44.63232311	Sa povredjenim	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila koja se kreću u istom smeru – susretanje
1169538	BEGRAD	BARAJEVO	21.02.2017.15:45	03.378437	44.628869	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – SKRETANJE ILI PRELAZAK	Najmanje dva vozila koja se kreću istim putem u suprotnim smerovima uz skretanje ulevo ispred drugog vozila
1169561	BEGRAD	BARAJEVO	22.02.2017.11:50	03.300615	44.54518483	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila koja se kreću u istom smeru – susretanje
1169556	BEGRAD	BARAJEVO	24.02.2017.14:00	03.46785	44.60033	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – SKRETANJE ILI PRELAZAK	Najmanje dva vozila uz skretanje – nepoznato
1170132	BEGRAD	BARAJEVO	03.03.2017.09:00	03.38662	44.56281	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – SKRETANJE ILI PRELAZAK	Najmanje dva vozila koja se kreću različitim putevima uz prolazak kroz raskrsnicu, ili od kojih jedno prelazi preko kolov
1170283	BEGRAD	BARAJEVO	03.03.2017.18:00	03.371732	44.62455	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila koja se kreću u istom smeru – susretanje
1172446	BEGRAD	BARAJEVO	04.03.2017.05:35	03.29963	44.59921	Sa mat.stetom	SN SA JEDINIM VOZILOM	Nezgodе sa učešćem jednog vozila i životinja
1170362	BEGRAD	BARAJEVO	06.03.2017.19:30	03.36989	44.62317	Sa povredjenim	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila koja se kreću u istom smeru – susretanje
1172797	BEGRAD	BARAJEVO	07.03.2017.11:30	03.41128	44.5526	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – BEZ SKRETANJA	Najmanje dva vozila koja se kreću u istom smeru – susretanje
1170709	BEGRAD	BARAJEVO	08.03.2017.11:20	03.329279	44.581576	Sa mat.stetom	SN SA NAJMANJE DVA VOZILA – SKRETANJE ILI PRELAZAK	Najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje ulevo ispred drugog vozila

Slika 1: Prikaz osnovnog skupa podataka

Kao što se sa lako može uočiti, ovako oblikovani izvorni podaci nisu normalizovani, niti uređeni, te su stoga neadekvatni za ozbiljniju analitiku i zaključivanje. Stoga je najpre bilo potrebno transformisati sve podatke u oblik koji će biti adekvatniji za dalje istraživanje. Pre svega, vremenski podaci o saobraćajnim nezgodama su atomski podeljeni na datum i vreme saobraćajne nezgode, tj. ne govore mnogo i nisu efikasni u smislu analitičke obrade. Iz tih razloga, pristupilo se njihovoj transformaciji, odnosno „izvlačenju“ informacija iz njih. Kako bi se pratila dinamika dešavanja saobraćajnih nezgoda, najpre je kolona koja predstavlja datum nezgode, primenom Excel-funkcije funkcije `TEXT()`, klasterizovana po godišnjim intervalima na sledeći način:

$$=TEXT([address], "yyyy").$$

Isti postupak korišćen je zatim za prikaz podatka o mesecu kada se dogodila neka saobraćajna nezgoda, s tim što se kao drugi argument, koristi zapis „mmmm“. Najzad, na isti način određen je i dan saobraćajne nezgode, gde je za drugi argument željenog prikaza korišćen zapis „dddd“.

Dalje, kako bi se odredio kvartal kada se dogodila saobraćajna nezgoda korištena je višestruko ugnježdjena IF-funkcija, čiji su argumenti nazivi meseca u odgovarajućem kvartalu (na primer, januar, februar i mart predstavljaju mesece I kvartala, itd). Kako bi brojevi kvartala bili ispisani rimskim brojevima korišćena je funkcija Excel-funkcija `ROMAN()` na sledeći način:

$$=IF(OR([address]="januar",[address]="februar",[address]="mart"),ROMAN(1),IF(OR([address]="april",[address]="maj",[address]="jun"),ROMAN(2),IF(OR([address]="jul",[address]="avgust",[address]="septembar"),ROMAN(3),ROMAN(4))))).$$

Nakon toga, za svaki evidentirani slog u bazi podataka utvrđeno je da li je dan na koji se dogodila saobraćajna nezgoda bio radni dan ili ne. Ovo je takođe urađeno kombinacijom Excel-funkcija `IF` i `AND`, gde su argumenti bili nazivi dana:

$$=IF(OR([address]="nedelja",[address]="subota"),"NE","DA").$$

Dakle, ukoliko je evidentiran neradni dan, tj. subota ili nedelja, vrednost koju funkcija vraća je „NE“ i obratno.



Kako bi se podaci o saobraćajnim nezgodama što kvalitetnije analizirali sa aspekta uslova saobraćaja, tj. frekvencije saobraćaja na dnevnom nivou, period od jednog dana (24h) je podeljen na tri jednaka intervala od po 8 sati. Na taj način, dobijeni su vremenski intervali koji su označeni simbolima „00-08h“, „08-16h“ i „16-00h“, a za koje se smatra da mogu bitnije uticati na oblik, težinu i vrstu saobraćajne nezgode. U samom postupku korišćen je niz ugnježenih IF-funkcija u kombinaciji sa AND-funkcijom.

Još jedan od novih atributa koji je formiran ovde odnosi se na teritoriju policijske uprave gde je nastala saobraćajna nezgoda. Kako teritorija Grada Beograda predstavlja ubedljivo najveću policijska upravu, izvršena je takođe i klasifikacija svih saobraćajnih nezgoda na one koje su se desile na teritoriji Beograda ili van njega. Ovo je urađeno pomoću IF-funkcije, gde je kao argument prosleđen podatak o tome da li je reč o policijskoj upravi Beograda ili ne:

$$=IF([address]="BEOGRAD","DA","NE")$$

Najzad, jedan od najznačajnijih događaja u prethodnom vremenskom periodu je svakako pandemija izazvana virusom SARS-COV2, tako da je i ova činjenica uzeta u razmatranje kao potencijalni atribut koji može uticati na oblik i težinu saobraćajnih nezgoda. U tu svrhu, sve saobraćajne nezgode iz originalnog skupa podataka klasifikovane su na one koje su se desile pre i posle pandemije ovog virusa. Postupak segmentacije podataka se opet zasniva na primeni ugnježenih IF-funkcija, gde su kao argumenti postavljene godine u kojoj se saobraćajna nezgoda dogodila. Ukoliko je nezgoda bila pre 2020. godine funkcija vraća vrednost „PRE“, a u suprotnom vraćena vrednost je „POSLE“. Preciznije, sintaksa ove funkcije glasi:

$$=IF(OR([address]="2015",[address]="2016",[address]="2017",[address]="2018", [address]="2019"),"PRE","POSLE").$$

Na kraju ovog postupka transformacije originalnih podataka, kao *osnovni izlazni kriterijum*, odnosno tzv. *klasa odlučivanja* uzeta je vrsta, odnosno težina saobraćajne nezgode u pogledu materijalne štete, odnosno eventualno povređenih ili stradalih lica. S obzirom na to da su sve vrste saobraćajnih nezgoda unutar originalnih podataka posmatrane na tri različita nivoa, ove vrednosti klasifikovane su abecednim redom u zavisnosti od ishoda i težine same nezgode. Tako je za najblaži oblik saobraćajnih nezgoda, u kojima je došlo „samo“ do materijalne štete, korišćen simbol „A“, za nezgode sa povređenim licima simbol „B“, dok je najteži oblik nezgoda sa tragičnim ishodom označen simbolom „C“. I ovde su u postupku klasifikacije korišćene višestruke ugnježdene IF-funkcije, čiji je argument podatak tekstualnog tipa koji govori o vrsti same saobraćajne nezgode. Kao izlazni rezultat dobija se jedna od pomenutih „mapiranih“ vrednosti „A“, „B“ ili „C“, na sledeći način:

$$=IF([address]="Sa mat.stetom","A", IF([address]="Sa povredjenim","B","C")).$$

Na ovaj način je transformacija izvornog skupa podataka završena, a finalni oblik nove baze podataka prikazan je na sledećoj slici 2.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	policijskaUprava	BEOGRAD?	opstina	datum	godina	COVID	mesec	kvartal	dan	radniDan	vreme	dobaD	Y	X	vrsta
2	BEOGRAD	TRUE	MILADENOVAC	1.1.2015	2015	PRE	janjvar	I	četvrtak	TRUE	13:15 08-16h	20,67108	44,55664	Sa mat.stetom	A
3	SOMBOR	FALSE	SOMBOR	1.1.2015	2015	PRE	janjvar	I	četvrtak	TRUE	4:45 00-08h	19,19975	43,664816	Sa mat.stetom	A
4	SOMBOR	FALSE	SOMBOR	1.1.2015	2015	PRE	janjvar	I	četvrtak	TRUE	3:00 00-08h	19,11643	43,770619	Sa mat.stetom	A
5	POZAREVAC	FALSE	POZAREVAC	1.1.2015	2015	PRE	janjvar	I	četvrtak	TRUE	16:25 16-00h	21,18786	44,6179727	Sa mat.stetom	A
6	SMEDEREVO	FALSE	SMEDEREVSKA PALANKA	1.1.2015	2015	PRE	janjvar	I	četvrtak	TRUE	7:40 00-08h	20,9689	44,368428	Sa mat.stetom	A
7	NOVI PAZAR	FALSE	SIENICA	1.1.2015	2015	PRE	janjvar	I	četvrtak	TRUE	13:15 08-16h	20,08804	43,254961	Sa povredjenim	B
8	ČAČAK	FALSE	GORNJI MILANOVAC	2.1.2015	2015	PRE	janjvar	I	petak	TRUE	17:30 16-00h	20,2147	44,025718	Sa povredjenim	B
9	SMEDEREVO	FALSE	VELIKA PLANA	2.1.2015	2015	PRE	janjvar	I	petak	TRUE	11:35 08-16h	21,08325	44,30493465	Sa povredjenim	B
10	POZAREVAC	FALSE	POZAREVAC	2.1.2015	2015	PRE	janjvar	I	petak	TRUE	3:40 00-08h	21,18305	44,62257038	Sa mat.stetom	A
11	SOMBOR	FALSE	SOMBOR	3.1.2015	2015	PRE	janjvar	I	subota	FALSE	14:10 08-16h	19,13002	43,786521	Sa povredjenim	B
12	NOVI SAD	FALSE	SIENICA	4.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	12:30 08-16h	19,8285	43,517667	Sa povredjenim	B
13	UŽICE	FALSE	BAJINA BAŠTA	4.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	15:30 08-16h	19,43418	43,88902	Sa mat.stetom	A
14	BEOGRAD	TRUE	NOVI BEOGRAD	4.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	13:45 08-16h	20,38542	44,81848	Sa povredjenim	B
15	NOVI SAD	FALSE	NOVI SAD	4.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	19:15 16-00h	19,81794	45,253366	Sa povredjenim	B
16	SOMBOR	FALSE	KULA	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	16:00 16-00h	19,3786	43,70135731	Sa povredjenim	B
17	SMEDEREVO	FALSE	SMEDEREVSKA PALANKA	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	16:25 16-00h	20,96157	44,366218	Sa mat.stetom	A
18	NOVI PAZAR	FALSE	SIENICA	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	10:00 08-16h	19,98986	43,274049	Sa mat.stetom	A
19	BEOGRAD	TRUE	STARI GRAD	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	13:30 08-16h	20,46572	44,81664	Sa mat.stetom	A
20	BEOGRAD	TRUE	SAVSKI VENAC	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	9:00 08-16h	20,45471	44,80104556	Sa mat.stetom	A
21	ČAČAK	FALSE	GORNJI MILANOVAC	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	10:00 08-16h	20,46435	44,021552	Sa povredjenim	B
22	SOMBOR	FALSE	SOMBOR	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	12:15 08-16h	19,10474	43,76339	Sa povredjenim	B
23	NOVI SAD	FALSE	NOVI SAD	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	19:03 16-00h	19,83899	43,276384	Sa povredjenim	B
24	UŽICE	FALSE	ARIJEU	5.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	9:30 08-16h	20,64709	43,710984	Sa povredjenim	B
25	SOMBOR	FALSE	KULA	6.1.2015	2015	PRE	janjvar	I	utorak	TRUE	14:55 08-16h	19,54036	43,607824	Sa povredjenim	C
26	ČAČAK	FALSE	GORNJI MILANOVAC	6.1.2015	2015	PRE	janjvar	I	utorak	TRUE	17:35 16-00h	20,46408	44,027817	Sa povredjenim	B
27	ZRENJANIN	FALSE	NOVA CRNIA	7.1.2015	2015	PRE	janjvar	I	sreda	TRUE	13:00 08-16h	20,55193	43,6654	Sa mat.stetom	A
28	KRUŠEVAC	FALSE	KRUŠEVAC	7.1.2015	2015	PRE	janjvar	I	sreda	TRUE	11:15 08-16h	21,33929	43,571993	Sa povredjenim	B
29	NOVI SAD	FALSE	NOVI SAD	8.1.2015	2015	PRE	janjvar	I	četvrtak	TRUE	17:25 16-00h	19,7223	43,243289	Sa povredjenim	B
30	UŽICE	FALSE	BAJINA BAŠTA	8.1.2015	2015	PRE	janjvar	I	četvrtak	TRUE	6:00 00-08h	19,58876	43,970855	Sa mat.stetom	A
31	BEOGRAD	TRUE	MILADENOVAC	10.1.2015	2015	PRE	janjvar	I	subota	FALSE	17:40 16-00h	20,64424	44,421591	Sa mat.stetom	A
32	SOMBOR	FALSE	SOMBOR	11.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	0:30 00-08h	19,37558	43,789524	Sa mat.stetom	A
33	SOMBOR	FALSE	KULA	11.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	16:00 16-00h	19,50586	43,62803772	Sa mat.stetom	A
34	BEOGRAD	TRUE	NOVI BEOGRAD	11.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	10:00 08-16h	20,41165	44,825659	Sa mat.stetom	A
35	ZRENJANIN	FALSE	NOVA CRNIA	11.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	18:25 16-00h	20,59499	43,6634819	Sa povredjenim	B
36	SOMBOR	FALSE	SOMBOR	11.1.2015	2015	PRE	janjvar	I	nedelja	FALSE	10:30 08-16h	19,1108	43,773013	Sa mat.stetom	A
37	NOVI SAD	FALSE	NOVI SAD	12.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	6:45 00-08h	19,72368	43,311107	Sa povredjenim	B
38	SMEDEREVO	FALSE	SMEDEREVSKA PALANKA	12.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	9:34 00-08h	20,96371	44,369971	Sa povredjenim	B
39	SOMBOR	FALSE	SOMBOR	12.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	18:35 16-00h	19,10897	43,765617	Sa povredjenim	B
40	KRALJEVO	FALSE	RAŠKA	12.1.2015	2015	PRE	janjvar	I	ponedeljak	TRUE	7:30 00-08h	20,63635	43,340611	Sa povredjenim	B

Slika 2: Finalni izgled transformisanog skupa podataka

### 3. STABLO ODLUČIVANJA

Na osnovu transformisanog skupa podataka, odnosno uvedenih atributa i klase odlučivanja, u ovom odeljku biće opisan postupak *formiranja stabala odlučivanja* nad datim podacima. Kao što je već istaknuto, osnovni cilj jeste kreiranje modela odlučivanja koji predviđa vrednost ciljne promenljive, tj. vrste i težine saobraćajne nezgode na osnovu uvedenih atributa, kao osnovnih ulaznih varijabli. Kao što je već rečeno ranije, stabla odlučivanja predstavljaju veoma popularno sredstvo u tzv. *mašinskom učenju* (engl. *Machine Learning*), gde predstavljaju tzv. *algoritme nadgledanog učenja*. Sa praktičnog stanovišta, stabla odlučivanja sadrže niz pravila koja su jasno čitljiva i razumljiva, pri čemu ona praktično oslikavaju poznati koncept „dijagrama toka“ i daju hijerarhiju odlučivanja, sve dok se podaci potpuno ne klasifikuju.

Svako stablo odlučivanja predstavlja *klasifikator* koji funkciju klasifikacije predstavlja u obliku stabla odlučivanja, tj. *hijerarhijski*. Čvorovi na vrhu stabla imaju najveći uticaj na klasifikaciju, dok važnost promenljivih opada kretanjem ka dnu stabla. Pritom se svakom stablu odlučivanja može dodeliti skup određenih mera kojim se determiniše kvalitet stabla (greška predviđanja ili neka druga mera). Osnovni tip stabala odlučivanja koji ovde opisujemo jeste tzv. *ID3 stablo* (engl. *Iterative Dichotomiser Tree*) koje se jednostavno konstruiše algoritmom [5] u sledeća tri koraka (od kojih se drugi i treći iterativno ponavljaju):

1. računanje entropije skupa slučajeva;
2. izbor promenljive (atributa) za grananje;
3. izbor kriterijuma zaustavljanja.

*Entropija* predstavlja meru neuređenosti sistema i u praksi se najčešće izračunava *Šenonovim obrascem*:

$$H(S) = - \sum_{i=1}^n p_i \log_2 p_i, \quad (1)$$

gde je  $S$  skup ulaznih podataka,  $p_i$  verovatnoće da će biti doneta  $i$ -ta odluka, a  $n$  broj različitih odluka. Verovatnoće  $p_i$  se mogu izračunati kao tzv. *aposteriori verovatnoće*  $p_i = |C_i|/|S|$ , gde vrednosti  $|C_i|$  predstavljaju broj slučajeva  $C_i$  sa  $i$ -tom odlukom, dok je  $|S|$  ukupan broj elemenata skupa  $S$ . Entropija se kod ID3 algoritma koristi kao *kriterijum valjanosti*, tj. smatra se da najveći značaj ima atribut sa *najmanjom entropijom*. To znači da takav atribut daje *najveću količinu informacija* (tzv. *informacionu dobit*) i predstavlja najadekvatniji izbor promenljive u grananju. Informaciona dobit  $ID(A)$  je mera razlike u entropiji između stanja pre i nakon što je skup  $S$  podeljen atributom  $A$  na određen broj slučajeva i izračunava se formulom:

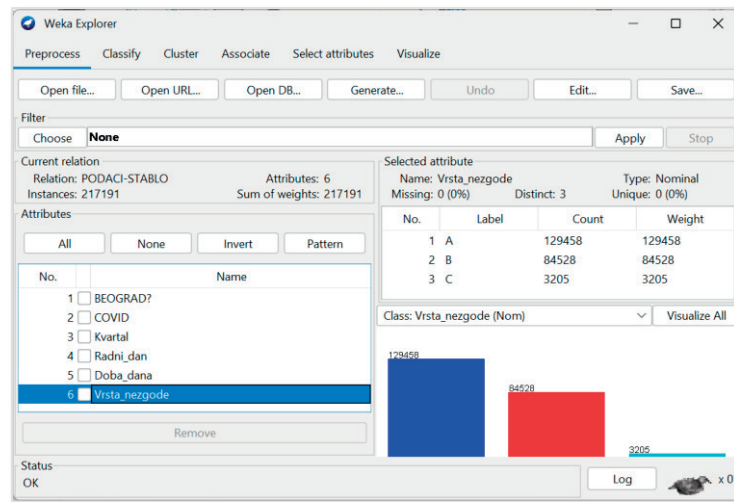
$$ID(A) = H(S) - \sum_{j=1}^m \frac{|S_j|}{|S|} \cdot H(S_j) = H(S) - H(S|A). \quad (2)$$

Ovde  $|S_j|$  predstavlja ukupan broj slučajeva, tj. elemenata sa  $j$ -tom osobinom atributa  $A$ , dok je  $H(S_j)$  entropija svih podskupova koji imaju  $j$ -tu osobinu.

U ID3 algoritmu se može izračunati informaciona dobit za svaki atribut, pri čemu se atribut sa najvećom informacionom dobiti koristi za podelu skupa u toj iteraciji. Na taj način, u svakoj iteraciji nastaje novi čvor za grananje, pri čemu taj čvor predstavlja atribut koji u datom koraku najviše smanjuje entropiju sistema. Najzad, *kriterijum zaustavljanja* ID3 algoritma, tj. prestanak grananja u stablu odlučivanja nastaje kada je ispunjen neki od sledeća tri uslova (tzv. *pravila zaustavljanja grananja*):

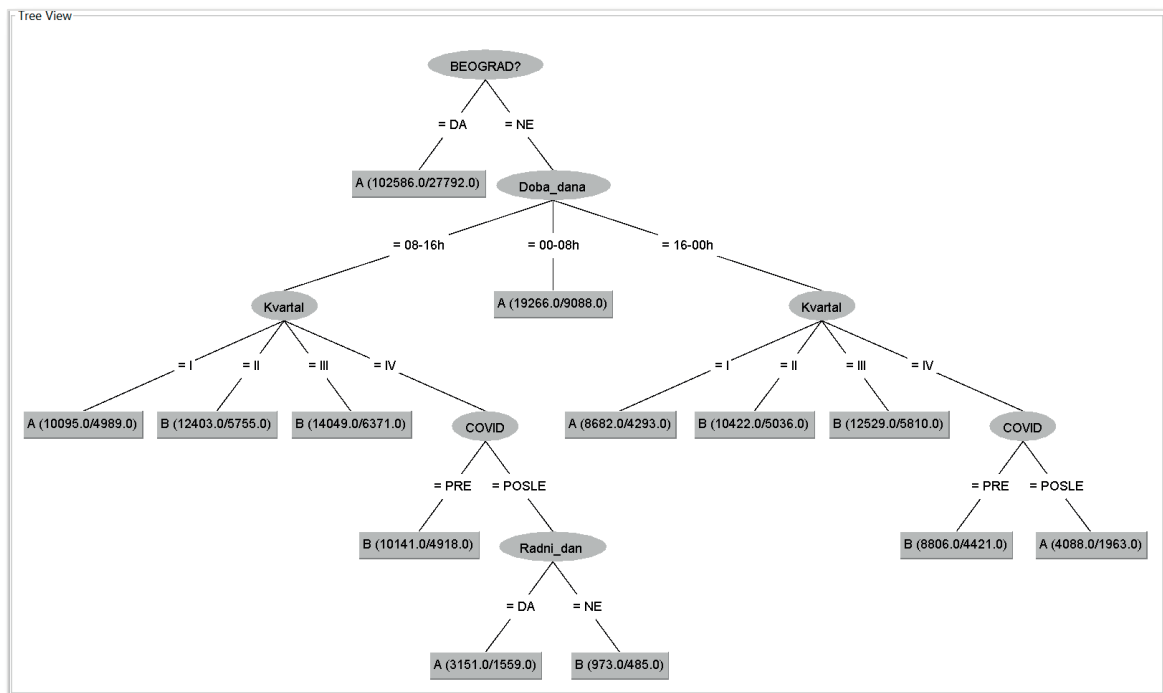
1. ako u čvoru stabla svi mogući slučajevi pripadaju istoj izlaznoj klasi, taj čvor se dalje ne deli;
2. ako je veličina čvora, tj. broj slučajeva u podstablu tog čvora manji od minimalne, korisnički definisane vrednosti, onda taj čvor nestaje;
3. ako atribut ima informacionu dobit manju od unapred zadate vrednosti, on neće biti izabran kao čvor.

Na osnovu prethodno formiranog skupa podataka o saobraćajnim nezgodama, konstrukcija stabla odlučivanja izvršena je primenom softverskog alata WEKA. Reč je o popularnom alatu za mašinsko učenje, koji sadrži veliki broj ugrađenih algoritama namenjenih upravo pretraživanju i učenju iz podataka [2]. Postupak rada obično počinje otvaranjem datoteke sa podacima, tj. izborom niza opcija *Preprocess* → *Open file...*, nakon čega se mogu koristiti osnovne funkcije softvera WEKA, kao što se može uočiti na slici 3. Primetimo da je naš skup podataka već prepoznat i klasifikovan po svim nominalnim osobinama unetih promenljivih (tj. atributa). Selektovanjem odgovarajućeg atributa u posebnom prozoru *Selected Attribute*, WEKA prikazuje osnovne podatke o njemu, odnosno njegovim grupisanim vrednostima. U našem slučaju, za klasifikacionu analizu uzeli smo pet atributa koje smo ranije formirali, a koji su od značaja za klasifikaciju ishoda saobraćajnih nezgoda: *BEOGRAD?*, *COVID*, *Kvartal*, *Radni dan* i *Doba dana*.



Slika 3: Unos podataka unutar alata WEKA Explorer

S druge strane, osnova klasifikacije i grananja u stablu odlučivanja biće vrednosti klase *Vrsta nezgode*, čije su osnovne karakteristike prikazane na slici 3. Nakon unosa podataka, izrada stabla odlučivanja zasniiva se na prethodno opisanom ID3 algoritmu, odnosno softverskoj implementaciji tog algoritma u alatu WEKA pod nazivom J48. Izborom niza opcija *Classify* → *Choose* → *Trees* → *J48* → *Visualize Tree* dobija se završni, vizuelni izgled stabla odlučivanja o uslovima i vrstama saobraćajnih nezgoda koji je dat na slici 4.



Slika 3: Stablo odlučivanja o uslovima nastanka i vrstama saobraćajnih nezgoda

#### 4. EVALUACIJA DOBIJENIH REZULTATA

Na osnovu dobijenog stabla odlučivanja moguće je izvršiti odgovarajuću analizu dobijenih rezultata, tj. njihovu kvalitativnu evaluaciju. Primetimo najpre da svi izlazni čvorovi stabla predstavljaju saobraćajne nezgode sa materijalnom štetom ili sa povređenim licima. Dakle, nema čvorova koji vode do saobraćajnih nezgoda sa smrtnim ishodom, što je očekivano jer, na svu sreću, udeo ovih saobraćajnih nezgoda je izrazito mali. Dalje, uočava se da atribut sa najvećom informacionom dobiti jeste promenljiva „*BEOGRAD?*“ koja predstavlja koren stabla. Ukoliko je izlaz grananja kod ovog atributa tačan, u najvećem broju slučajeva dolazi do saobraćajnih nezgoda sa materijalnom štetom (klasa *A*). Ovo je takođe očekivano, jer uslovi vožnje u gradskim uslovima, ograničenja brzine i gužve u kojima se učesnici kreću sporo uglavnom uzrokuju saobraćajne nezgode sa manjim materijalnim oštećenjima, nastalim usled slabijih kontakata dva vozila.

Međutim, zanimljivo je da u ostalim delovima Republike Srbije, tj. ukoliko je izlaz grananja atributa „*BEOGRAD?*“ netačan, dolazi do izraženog uticaja ostalih atributa i uslova pod kojima dolazi do saobraćajnih nezgoda. Najveći uticaj ovde ima atribut „*Doba dana*“, tj. vremenski interval tokom dana se saobraćajna nezgoda dogodila. Recimo, u jutarnjem periodu, tačnije od 00-08 sati ujutru, javljaju se uglavnom saobraćajne nezgode sa materijalnom štetom. Međutim, u ostalim periodima dana, dakle u periodu od 08-16 časova i 16-00 časova, izražen je uticaj kvartala u kojima se dešavaju saobraćajne nezgode. Tako je u prvom kvartalu izražen broj saobraćajnih nezgoda sa materijalnom štetom (klasa *A*), dok su u drugom i trećem kvartalu, dakle od aprila pa sve do septembra, dominantne saobraćajne nezgode sa povređenim licima kao ishodom (klasa *B*).

Najzad, u četvrtom kvartalu, dakle u oktobru, novembru i decembru, na ishod saobraćajne nezgode dolazi do uticaja vremenskog perioda pre ili nakon pandemije izazvane koronavirusom (opisan atributom „*COVID*“). Ovde je situacija malo kompleksnija. Ukoliko vreme u kome se dogodila saobraćajna nezgoda pripada dobu dana od 16 časova do ponoći, onda se kod saobraćajnih nezgode koje su se dogodile pre izbivanja pandemije, dominantno javljaju saobraćajne nezgode sa povređenim licima kao ishodom (klasa *B*). Suprotno tome, za saobraćajne nezgode koje su nastupile nakon pandemije dominantne su one sa materijalnom štetom (klasa *A*). U drugom slučaju, kada je doba dana u kome se dogodila saobraćajna nezgoda između 08 i 16 časova, takođe dolazi do uticaja atributa „*COVID*“, ali se stablo dalje grana na osnovu slučajeva (poslednjeg) atributa „*Radni dan*“. To znači da na dalji ishod saobraćajne nezgode utiče podatak o tome da li je dan u kome se ona dogodila bio radni dan ili ne. Ukoliko se saobraćajna nezgoda dogodila tokom radne nedelje, tj. bilo kog dana od ponedeljka do petka, dominantno se javljaju saobraćajne nezgode sa materijalnom štetom. Sa druge strane, ukoliko je dan u kome je nastupila saobraćajna nezgoda bio neradni, dakle subota ili nedelja, primetna je dominacija saobraćajnih nezgoda sa povređenim licima kao ishodom.

## 5. ZAKLJUČAK

Ovim radom obrađene su saobraćajne nezgode koje su se dogodile na teritoriji Republike Srbije u periodu od 01.01.2015. godine do 31.12.2021. godine. Cilj istraživanja bio je pronalaženje zakonitosti i obrazaca unutar naizgled slučajnih događaja, korišćenjem algoritama mašinskog učenja. Originalni podaci su, uvođenjem pojedinih atributa koji su od značaja za uslove pod kojima dolazi do saobraćajnih nezgoda, kao i na oblik i težinu samih nezgoda, segmentirani u oblik pogodniji za dalju analizu. Nakon ove transformacije podataka, pristupilo se izračunavanju entropija i informacione dobiti svakog atributa ponaosob. Pokazano je da saobraćajne nezgode nisu samo slučajni događaji, bez ikakve povezanosti i zakonitosti u njihovom pojavljivanju, već da se na bezbednost u saobraćaju može pozitivno uticati analizirajući svaki od definisanih atributa. Stoga, buduća istraživanja mogu podrazumevati proširivanje originalnog skupa novim podacima o evidentiranim saobraćajnim nezgodama na teritoriji Republike Srbije. Zatim, moguće je i uvođenje nekih novih atributa, drugačijih segmentacija od onih koje su autori ovde predložili, kao i novih izlaznih klasifikatora. Ovo se posebno odnosi na saobraćajne nezgode nastale na teritoriji Beograda, gde je predikativni ishod stabla samo materijalna šteta (klasa *A*), a to očito nije najverniji prikaz realne situacije. U tom smislu, smatramo da bi potencijalni značaj rada mogao da se ogleda i u davanju podsticaja nekim budućim istraživanjima iz oblasti nauke o podacima, tj. njenoj primeni u oblasti bezbednosti saobraćaja.

## LITERATURA

- [1] Aleksić A., Nedeljković S., Jovanović M., Randelović M., Vuković M., Stojanović V., Radovanović R., Randelović M., Randelović D. (2020) „Prediction of Important Factors for Bleeding in Liver Cirrhosis Disease Using Data Mining Approach“, *Mathematics*, Vol. 8(11), Article ID: 1887
- [2] Garner S. (1995) “WEKA: The Waikato Environment for Knowledge Analysis”, University of Waikato Machine Learning Project, New Zealand
- [3] Grzymala-Busse, J. W. (1993) "Selected Algorithms of Machine Learning from Examples". *Fundamenta Informaticae*. Vol. 18(2): pp. 193–207
- [4] Hatami B., Asadi F., Bayani A., Zali M. R., Kavousi K. (2022) “Machine learning-based system for prediction of ascites grades in patients with liver cirrhosis using laboratory and clinical data: design and implementation study“, *Clin. Chem. Lab.*, <https://doi.org/10.1515/cclm-2022-0454>
- [5] Quinlan, J. R. (1987). "Simplifying decision trees". *International Journal of Man-Machine Studies*. Vol. 27(3), pp. 221–234
- [6] Stojanović V. (2022) „Uvod u analitiku podataka“, Kriminalističko-policijski univerzitet, Beograd
- [7] <https://data.gov.rs/sr/datasets/podatsi-o-saobratshajnim-nezgodama-po-politsijskim-upravama-i-opshtinama/>

## PARETOV PRINCIP KAO MERA KORISNOSTI MODELA MAŠINSKOG UČENJA

### THE PARETO PRINCIPLE AS A UTILITY MEASURE OF MACHINE LEARNING MODELS

BORIS DELIBAŠIĆ<sup>1</sup>, SANDRO RADOVANOVIĆ<sup>1</sup>, ANDRIJA PETROVIĆ<sup>2</sup>, MILIJA SUKNOVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Fakultet organizacionih nauka, [boris.delibasic@fon.bg.ac.rs](mailto:boris.delibasic@fon.bg.ac.rs)

<sup>2</sup> Univerzitet Singidunum, [apetrovic@singidunum.ac.rs](mailto:apetrovic@singidunum.ac.rs)

---

**Rezime:** Pareto princip je jedan od osnovnih principa političke ekonomije, a može da se uoči u mnogim oblastima ljudskog delovanja. Po ovom principu 80% efekata u sistemu se ostvaruje preko 20% činioca. U mašinskom učenju ovaj princip može da se primeti kod faktorske analize, gde je očigledan, ali i kod fundamentalnih oblasti mašinskog učenja, tj. kod klasifikacije i klasterovanja. Naša tvrdnja jeste da upravo postojanje Pareto principa omogućava modelima klasifikacije i klasterovanja da budu primenljivi u praksi, tj. da odsustvo ovog principa u modelima klasifikacije i klasterovanja čine ove modele prilično neupotrebljivim. Predlažemo da u praksi mašinskog učenja počnu više da se koristi mere kvaliteta modela koji su zasnovane na Pareto principu, jer time mogu direktno da se povežu sa korišću koje mogu da ostvare ili ne ostvare u praksi..

**Ključne reči:** Pareto princip, Mašinsko učenje, Klasifikacija, Klasterovanje, Faktorska analiza.

**Abstract:** The Pareto principle is one of the basic principles of political economy, and can be observed in many areas of human activity. According to this principle, 80% of the effects in a system are realized by 20% of the factors. In machine learning, this principle can be applied in factor analysis, where it is obvious, but also in fundamental areas of machine learning, ie. in classification and clustering. Our claim is that the very existence of the Pareto principle allows the models of classification and clustering to be applicable in practice, ie. that the absence of this principle in classification and clustering models makes these models rather worthless. We suggest that in the practice of machine learning begin to use more quality measures of models that are based on the Pareto principle, because they can thus be directly related to the benefits that may or may not be achieved in practice.

**Keywords:** Pareto principle, Machine learning, Classification, Clustering, Factor Analysis.



## WIND POWER PREDICTION USING MACHINE LEARNING AND AI: A CASE STUDY

ISRAT HAQUE ZARIN<sup>1</sup>, AMZAD HOSSAIN<sup>2</sup>, HASNAIN AHMED<sup>3</sup>, MD ABU HELAL<sup>4</sup>, VISWANATH RAMAKRISHNA<sup>5</sup>

<sup>1</sup> University of Chittagong -Faculty of sciences, [israthaquezarini@gmail.com](mailto:israthaquezarini@gmail.com)

<sup>2</sup> University of Chittagong -Faculty of sciences, [amzadmthcu@gmail.com](mailto:amzadmthcu@gmail.com)

<sup>3</sup> Jahangirnagar University - Faculty of [mathematical](#) and [physical sciences](#), [hasnain.mathju@gmail.com](mailto:hasnain.mathju@gmail.com)

<sup>4</sup> Colorado State University - Warner College of Natural Resources, [abu.helal@colostate.edu](mailto:abu.helal@colostate.edu)

<sup>5</sup> The University of Texas at Dallas - Department of Mathematical Sciences, [vish@utdallas.edu](mailto:vish@utdallas.edu)

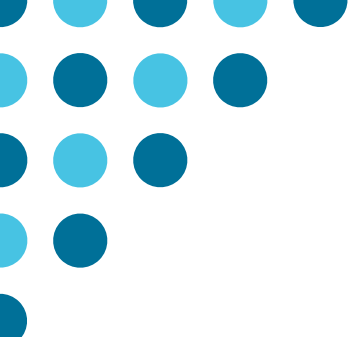
---

**Abstract:** *Wind power is a prominent renewable energy source because of its environmental and economic benefits. However, wind energy generation is sporadic due to the high volatility and random nature of wind energy, causing supply and demand issues. A high accuracy estimate of wind power is required to reduce the risk of a demand-supply imbalance. Machine learning ensemble techniques is a promising approach to achieve the better prediction accuracy than any single learning algorithm as it aggregates the predictions from various models. In a previous study, it has been shown that the random forest model outperformed the feedforward neural network, decision tree, and support vector machine. In this article, we use historical wind speed data to demonstrate that the XG-boost algorithms can predict wind power with a level of accuracy substantially superior to that of earlier algorithms. We used cutting-edge machine learning techniques throughout the course of our analysis.*

**Keywords:** *Power Forecasting, Wind Power, Machine Learning, XG-Boost Algorithm*







# POSLOVNA ANALITIKA

---

# BUSINESS ANALYTICS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business



## DETERMINING ENTREPRENEURS' PREFERENCES TOWARDS ONLINE MARKET RESEARCH USING CHOICE-BASED CONJOINT ANALYSIS

OGNJEN NIKOLIĆ<sup>1</sup>, MARIJA KUZMANOVIĆ<sup>2</sup>

<sup>1</sup> University of Belgrade – Faculty of Organizational Sciences, on20213655@student.fon.bg.ac.rs

<sup>2</sup> University of Belgrade – Faculty of Organizational Sciences, marija.kuzmanovic@fon.bg.ac.rs

**Abstract:** *A number of studies have shown that one of the top reasons why startups fail is due to their misreading of the market's need. The goal of this paper is to determine how entrepreneurs approach market research, how well they are informed and trained in those areas, and what would be their preferences regarding online market research. For that purpose, methodology that was used was choice-based conjoint analysis. Attributes that were addressed were: price, number of surveys that can be run, number of questions that they can ask in a survey, confidence interval of the results, quickness of results, hours of consultations with an expert and level of details provided in the report after the survey ends for the platform that does automatic collection of answers. Participants in this survey were entrepreneurs, product managers, UI/UX designers and students of the aforementioned studies.*

**Keywords:** *Choice-Based-Conjoint analysis, entrepreneurs, market research, preferences, willing-to-pay*

### 1. INTRODUCTION

As the global startup ecosystem grows and entrepreneurship is becoming an ubiquitous word, a pattern of problems that startups encounter can be observed, with one of the top problems and reasons for failure being misreading of the market's needs [1]. That alone can have a huge impact on the entrepreneurial team, if not detected early, since the budget for the initial development is mostly consisted of team's own finances [2]. That being said, development of the product that won't be used by the market can create big financial and motivational problems for the team and their future work.

In the preliminary phase, through interviews with startup founders, people from marketing and students doing their thesis about their way of market research, one aspect arose that is convenient for all of them – need for a primary market research, but a lack of resources to pay for one. Free online survey tools give the option of creating a survey, but don't help with finding the audience that would fill it out, which means a lot of work for promoting the survey, if there aren't enough resources to pay for the access to someone's panel of respondents. That usually means, devoting the time to find the place where the desired target group is located, creating connections that would help out with spreading the survey and finding the right incentive why would someone fill it out. Platforms that have the option of finding survey respondents and doing that job for the customer have prices ranging from 300-12000€, depending on the sample targeted, number of filters for target audience etc.

This research tested how respondents perceive the purpose and need for market research, how much they are willing to pay for its online version that guarantees sufficient responses from the target group, as well as their preferences for different attributes and levels of online market research service. To measure the preferences of the respondents and determine their willingness to pay for the service, a Choice-Based-Conjoint analysis (CBC) was used.

### 2. METHODOLOGY

Choice-Based-Conjoint analysis, also called Discrete choice analysis was created in order to measure importance of certain attribute and levels when they are in direct conflict with other attributes and their levels. It is grounded in Random Utility Theory.

CBC implies that respondents are shown different alternatives of products / services that consist of the same attributes, but differ in the levels of these attributes. Attributes can be qualitative or quantitative. Respondents choose between those alternatives and are expected to choose rationally in order to maximize the total value for them. Main output of the CBC analysis are utility scores, that correspond to the degree to which each attribute and level impact individuals' choices.

Let's assume that  $I$  respondents choose from the set of  $J$  mutually exclusive alternatives. Each respondent receives some utility from each of the alternatives. As stated, respondents are expected to behave rationally

and, an individual  $i$  ( $i = 1, \dots, I$ ) would choose alternative  $j$  ( $j = 1, \dots, J$ ) if and only if the utility corresponding to that alternative ( $U_{ij}$ ) is greater than or equal to the utility of all other alternatives. Utility  $U_{ij}$  is consisted of a deterministic component  $V_{ij}$  and stochastic component  $\varepsilon_{ij}$ , i.e.,

$$U_{ij} = V_{ij} + \varepsilon_{ij}$$

Deterministic component answers to the goal of a choice model to identify the attributes that affect the utility of individuals and estimate their importance values. For this reason, it is necessary to specify a functional form for  $V_{ij}$ , that is usually a linear in parameters, additive form [3]

$$V_{ij} = \sum_{k=1}^K \sum_{l=1}^{L_k} \beta_{ikl} x_{jkl}, \quad i = 1, \dots, I, \quad j = 1, \dots, J,$$

where  $K$  is the number of attributes;  $L_k$  is the number of levels of attribute  $k$ .  $\beta_{ikl}$  is respondent  $i$ 's utility with respect to level  $l$  of attribute  $k$  (so called part-worth utility).  $x_{ikl}$  is such a (0,1) variable that it equals 1 if alternative  $j$  has attribute  $k$  at level  $l$ , otherwise it equals 0.

To estimate the model parameters (part-worths), Markov Chain Monte Carlo Hierarchical Bayes (MCMC HB) can be used. This approach estimates part-worth utilities not for market as a whole, but rather for individuals, which makes it possible to further calculate the relative importance of each attribute for each respondent and for different groups of respondents. Relative importance scores are calculated by taking the utility range for each attribute separately and then dividing it by the sum of the utility ranges for all of the attributes [3].

### 3. STUDY DESIGN

The research addressed several research questions:

- Which methods and tools does the target group use for market research?
- Which methods and tools do they intend to use in the future?
- How much are they willing to pay for an online questionnaire that would be filled by their desired target group and a report of the results upon completion?
- What are respondents' preferences towards a key characteristic of the platform for online market research?

The first and crucial step in CBC is to identify the key attributes that influence the choice of respondents. In this study, attributes and levels were selected based on a literature review [1][4][5] and interviews with startup mentors and startup founders who highlighted their experience, problems, knowledge of data analytics, benefits they expect from market research and startup financial opportunities (see Table 1).

Table 1: The list of attributes and their levels used in research

Attributes	Levels			
Price	150 €	300 €	400 €	500 €
Number of surveys	1	3	5	
Number of questions	10	30	50	
Confidence interval	90%		95%	
Hours of consultations with an expert	0	2	4	6
Days for getting the report	3	5	7	
Level of details in a report	No report, just raw data	Basic*	Detailed**	

\*Just the basic results for every question, e.g., bar chart regarding the distribution of answers per option.

\*\*Statistics for every question, with key correlations between answers and segmentation

Conjoint.ly platform was used for creating the survey and collecting answers online [6]. Randomized block design [7] was used to generate a set of choice tasks that were presented to respondents, consisting of different attributes with different levels given in Table 1. A total of 10 different blocks of 8 tasks with 4 alternatives was generated. One example of those tasks can be seen in Figure 1. Alternatives represented various possible subscription packages to an online platform that would guarantee that the desired target group would complete the survey and that a report would be generated for the client. Only two prohibited pairs were set: 5 surveys at a price of 150 euros, and 50 questions per survey at a price of 150 euros.

	Package A	Package B	Package C	Package D
<b>Price</b>	150 €	300 €	500 €	400 €
<b>Surveys</b>	3	1	1	3
<b>Questions</b>	30	30	50	10
<b>Accuracy</b>	90% (standard)	90% (standard)	90% (standard)	95%
<b>Quickness of results</b>	5 workdays	5 workdays	7 workdays	3 workdays
<b>Consultations</b>	Total 2 hours	Don't exist	Total 2 hours	Total 4 hours
<b>Report</b>	Basic*	Basic*	Detailed**	Detailed**

Figure 1. Example of one choice task with four alternatives

Two main target groups were surveyed: entrepreneurs, freelancers and employees of smaller companies, and people with aspirations to become entrepreneurs, mainly consisted of students. In order to find more about their background, attitudes towards market research and general knowledge in the field, additional questions were posed.

## 4. RESULTS

One of the first hypothesis that was tested with this survey is that it is hard to find relevant target audience that would be willing to fill out an online survey. For this survey, with a slightly smaller population, the completion time was a little bit over 2.5 months and for 187 valid answers, there was 1090 entries to the survey, most of which exited upon entry (81.2%), some of them left after a couple of questions (1.3%) and some of them were invalid, due to low quality of answers (0.4%). For a full disclosure, one example doesn't make a rule, and it cannot be said that the hypothesis has been proven. But the comment must be made that after 2.5 months of posting in relevant channels and reaching more than 13000 people from the relevant population, whose members claim they love helping out, and with incentives in place for filling out the survey, percent of valid answers shows that this research is on the right track.

### 4.1. Demographic

Out of those 187 valid answers in the survey with 34.8% of the participants declaring themselves as entrepreneurs, 33.7% as product designers/managers, 16.9% as UI/UX designers 7.9% as members of R&D teams, 12.4% as researchers and 52.9% as students (participants could check more than one answer). As for their distribution in companies 37.5% work in startups, scaleups and middle companies, and 14.4% as freelancers, while others do not fall in the primary research group. It is interesting to conclude that most of the participants (73%) stated that their line of work is connected to the IT industry.

### 4.2. Attitude towards market research

Most of the participants (80.2%) rated their knowledge in the field of market research with the grade of 3 or below, on a scale 1 to 5. Only 19.8% rated their knowledge in the upper bounds, with the grades 4 and 5. Mean value was 2.8 with standard deviation of 0.9. That can also be observed in their answers about their knowledge of some of the methodologies of market research, where only 7.7% selected that they have heard about Van Westendorp Price Sensitivity analysis, 3.3% about Monadic test, and 15.8% selected that they have heard about Conjoint analysis. Based on the methodologies that they have heard of they were asked to rate how likely they are to use them in the future on a scale 1-5. Results can be observed in Table 2, where percentage of users likely to use it is based on all the answers with grades of 4 and 5.

It can be noted that participants share more interest towards qualitative and secondary research than quantitative, which can be seen by the general knowledge and plans for the future use of the methodologies listed in Table 2Table 2.

Although Conjoint analysis has proven to be one of the best quantitative methods for testing new ideas and concepts, predicting potential market share and determining respondents' willingness to pay, only a small percentage of respondents intend to use it in the future (5.99%). The main reason is the fact that the respondents have either not heard of the method at all or are not sufficiently familiar with its advantages.

Table 2. Answers regarding market research methodologies (in % of the total sample)

Method	Heard about	Intend to use
Simple questionnaire	79.8%	<b>51.95%</b>
A/B testing	45.4%	25.70%
Factor analysis	40.4%	21.29%
Van Westendorp Price Sensitivity Analysis	7.7%	3.30%
Monadic testing	3.3%	1.65%
Conjoint analysis	15.8%	5.99%
Focus groups	84.7%	<b>56.75%</b>
Customer interviews	79.2%	<b>53.54%</b>
Reading scientific articles	64.5%	38.83%
Reading business articles	73.2%	<b>53.51%</b>
Creating and releasing MVP	65.0%	46.99%
Talking with friends	73.8%	47.82%

### 4.3. Aggregated preferences

The utility parameters were estimated individually for each respondent in the sample (individual preferences), as well as for the total sample (aggregated preferences). McFadden's pseudo- $R^2$  of 59.3% implies medium goodness of fit. Table 3 shows both the averaged part-worth utilities for each level of each attribute and importance scores for each of the attributes. Higher part-worth utility values indicate greater preference.

*Price* and *Report* were found to be the two most important attributes, with their relative importance of 27.6% and 25.7%, respectively (Table 3). By analyzing the relative importance of attributes and their corresponding levels in Table 3, a conclusion can be made that after the price the most important attributes are those concerning the presence of help from an expert. That correlates with the answers regarding their confidence in the field of market research.

Table 3. Averaged part-worths and relative importance of attributes

Attribute	Level	Averaged preferences	Lower bound	Upper bound	Relative importance
			of 90% confidence interval		
Price	150 €	0.132	0.120	0.143	27.6%
	300 €	0.085	0.078	0.092	
	400 €	-0.057	-0.065	-0.049	
	500 €	-0.160	-0.170	-0.148	
Surveys	1	-0.083	-0.090	-0.076	13.3%
	3	0.027	0.023	0.030	
	5	0.056	0.049	0.063	
Questions	10	-0.031	-0.036	-0.026	7.8%
	30	0.009	0.004	0.015	
	50	0.021	0.015	0.027	
Accuracy	90% (standard)	-0.032	-0.036	-0.028	5.9%
	95%	0.032	0.028	0.036	
Quickness of results	3 workdays	0.003	0.000	0.006	5.1%
	5 workdays	0.008	0.004	0.012	
	7 workdays	-0.012	-0.016	-0.007	
Consultations	Don't exist	-0.095	-0.099	-0.091	14.6%
	Total 2 hours	0.019	0.013	0.026	
	Total 4 hours	0.036	0.029	0.042	
	Total 6 hours	0.040	0.035	0.045	
Report	No report, just raw data	-0.156	-0.165-	-0.146	25.7%
	Basic	0.012	0.007	0.018	
	Detailed	0.143	0.132	0.154	

As for the averaged part-worths, they are also presented in the Table 3. It can be observed that the *detailed* level of *report* has the highest relative part-worth, averaging across all respondents, closely followed by *price* level of 150€. On the other hand, *price* of 500€ and *lack of report* have a negative impact on the respondents, as well as *lack of consultations* and *just 1 survey* included in the package. None of the levels within the attributes *number of questions* and *quickness of results* has a significant impact on respondents' preferences.

One of the advantages of CBC analysis and part-worth utilities is the ability to calculate the marginal willingness to pay to upgrade to a certain attribute level if all other attributes in the alternative remain at existing levels [8]. First step is setting up the baseline alternative, and then analyzing the marginal willingness to pay for every level. For the baseline, lowest level of each attribute has been taken (Table 4).

Table 4. Marginal willingness to pay for levels

Attribute	Level	Marginal willingness to pay (in euros)	Lower bound of 90% confidence interval	Upper bound of 90% confidence interval
Surveys	1	0	0	0
	3	131.78	113.83	155.89
	5	158.20	140.40	188.20
Questions	10	0	0	0
	30	47.99	30.48	69.56
	50	49.91	36.16	69.50
Accuracy	90% (standard)	0	0	0
	95%	51.75	38.84	62.38
Quickness of results	3 workdays	0	0	0
	5 workdays	-10.98	-17.83	-3.53
	7 workdays	-17.72	-27.64	-5.30
Consultations	Don't exist	0	0	0
	Total 2 hours	99.88	87.93	116.93
	Total 4 hours	111.25	95.28	145.06
	Total 6 hours	137.27	117.32	166.79
Report	No report, just raw data	0	0	0
	Basic	160.68	140.59	200.44
	Detailed	313.63	269.03	361.26

Results show that respondents are highly prepared to pay for a *detailed report*, up to around 300€ if every other attribute has the lowest level which correlates with the part-worth for that level in the Table 3. It is interesting to note that the respondents are willing to pay around 160€ for the highest level within *expert* consultation attribute. For all other attributes there isn't a significant increase marginal willingness to pay between the levels themselves, except when compared with the baseline levels, which can be seen in Table 4.

## 5. CONCLUSION

The goal of the research was to find out how entrepreneurs, and those who want to become entrepreneurs perceive market research, their knowledge in the field and preferences towards using online tool for market research. Idea for it was based on a fact that one of the main reasons for failure of startups is the lack of market's need for what the startups were offering.

The research was conducted using CBC analysis and it helped determine respondents' preferences, but it also helped uncover barriers regarding doing market research online by oneself if the person is lacking skills in the field. As expected, most of the respondents show a lack of confidence in their skillset in the field of market research which resulted in their higher preferences towards expertise when creating and analyzing online surveys. On the other hand, they don't recognize value (or don't have enough money to pay) for quality online market research analysis, and in the conflict of lower price and higher expertise, lower price wins by a small margin. Given that founders usually don't have investors at the beginning, but use their own money, the question arises as to whether it would be cheaper for them to test their ideas through several different surveys, or spend months in development before finding that the idea is not market-fit.

When we look back at the beginning of this paper, the first sentence says that one of the main reasons why startups fail is due to misinterpretation of market needs. If they are aware of their lack of skills in market research and prefer the help of experts, a question can be raised why aren't they doing something about it, and why are they still continuing to work the same way, pouring money into their ideas, who in a lot of the cases, nobody will use. Moreover, should they change their perspective regarding market research and spare a little bit more money for the quality research, or should the market offer them a less expensive approach to quickly test their ideas.

Since this research had respondents with various backgrounds and McFadden's pseudo- $R^2$  was 59.3% one of the first steps in future research would be to find if there is a difference between different segments of the population, and to group respondents in clusters and explore main distinctions between them. Furthermore, since CBC analysis's main outputs are part-worths that gives the option to create simulations in various scenarios, so one of the future directions in the research would be to try to find optimal set of subscription packages for the online survey platform which would maximize value for the customer, as well as profit for the platform and test if people would actually want to use it.

## Acknowledgement

Platform used to create the survey, gather responses, and get analysis was Conjoint.ly with an academic license.

Big thanks to Hivemind, Željko Skenderović, Digital Serbia Initiative, Science Technology Park Belgrade, Impact Hub, Nova Iskra Workspace, *Udruženje mladih privrednika Srbije*, and all those who helped in spreading the survey.

## REFERENCES

- [1] Bednár, R., & Tarišková, N. (2017). Indicators of startup failure. *Industry 4.0*, 2(5), 238-240.
- [2] Digital Serbia Initiative. (2022). Startup Scanner. Retrieved from [www.preduzmi.rs:https://www.preduzmi.rs/wp-content/uploads/2022/04/Startup\\_Skener\\_ENG.pdf](http://www.preduzmi.rs:https://www.preduzmi.rs/wp-content/uploads/2022/04/Startup_Skener_ENG.pdf)
- [3] Kuzmanovic, M., Makajic-Nikolic, D., & Nikolic, N. (2019). Preference based portfolio for private investors: Discrete choice analysis approach. *Mathematics*, 8(1), 30.
- [4] Giardino, C., Wang, X., & Abrahamsson, P. (2014, June). Why early-stage software startups fail: a behavioral framework. In *International conference of software business* (pp. 27-41). Springer, Cham.
- [5] Cantamessa, M., Gatteschi, V., Perboli, G., & Rosano, M. (2018). Startups' roads to failure. *Sustainability*, 10(7), 2346.
- [6] Online Conjoint.ly platform. Link: <https://conjointly.com/guides/what-is-conjoint-analysis/>
- [7] Trochim, W. M., & Donnelly, J. P. (2001). *Research methods knowledge base* (Vol. 2). Macmillan Publishing Company, New York: Atomic Dog Pub..
- [8] Kojcic I., Kuzmanovic M. (2022). Conjoint Analysis of Green Consumer Preferences for Electronic Products, *International Journal for Quality Research*, 16(2), 559 - 575, doi: 10.24874/IJQR16.02-14



## EVALUACIJA UTICAJA PANDEMIJE COVID-19 NA POTRAŽNJU ZA PIVOM U REPUBLICI SRBIJI

### EVALUATION OF COVID-19 PANDEMIC IMPACT ON BEER DEMAND IN REPUBLIC OF SERBIA

OGNJEN ANĐELIĆ<sup>1</sup>\*, ZORAN RAKIĆEVIĆ<sup>1</sup>, ALEKSANDAR RAKIĆEVIĆ<sup>1</sup>

<sup>1</sup> Fakultet organizacionih nauka, Univerzitet u Beogradu, \*[ognjen.andjelic@fon.bg.ac.rs](mailto:ognjen.andjelic@fon.bg.ac.rs)

**Rezime:** *Pandemija prouzrokovana virusom COVID-19 uticala je na gotovo sve industrije širom sveta. Jedna od možda najviše pogođenih industrija, pored medicinske, je industrija hrane i pića, u okviru koje se nalazi i pivo kao alkoholno piće. Ova industrija trpi poremećaje prouzrokovane paničnim kupovinama, promenama ponašanja potrošača, zabranom okupljanja na javnim mestima i socijalnim distanciranjem. U radu je postavljeno istraživačko pitanje, da li je, na koji način i u kojoj meri pandemija COVID-19 imala uticaja na potražnju za pivom u Republici Srbiji? Kako bi se dobio odgovor na navedeno pitanje, metodama dekompozicije i predviđanja vremenskih serija analizirana je vremenska serija, koja predstavlja realne podatke sa srpskog tržišta, od 108 nedeljnih opservacija vezanih za potražnju za pivom u periodu pre i tokom pandemije. Podaci o prodaji piva dobijeni su na osnovu maloprodajnih POS izveštaja u nekoliko najvećih maloprodajnih lanaca u Republici Srbiji. Rezultati analize predstavljeni u radu sugerišu da je pandemija imala uticaja na potražnju za pivom, povećavajući istu.*

**Ključne reči:** *potražnja, predviđanje, pivo, COVID-19, pandemija, analiza vremenskih serija*

**Abstract:** *The pandemic caused by the COVID-19 virus has affected almost all industries around the world. One of perhaps the most affected industries, in addition to the medical industry, is the food and beverages industry, which includes beer as an alcoholic beverage. This industry suffers from disorders caused by panic buying, changes in consumer behavior, restriction of gatherings in public places and social distancing. This paper proposes a research question about whether the pandemic has had an effect on beer demand in the Republic of Serbia, and in what way and to what extent? In order to get an answer to this question, the time series with real data from the Serbian market consisting of 108 weekly observations related to the demand for beer pertaining to the period before and during the pandemic was analyzed by decomposition and time series prediction methods. Time series data on beer sales was obtained on the basis of retail POS reports in several of the largest retail chains in Republic of Serbia. The results of the analysis presented in the paper suggest that the pandemic has had an impact on the demand for beer, by increasing it.*

**Keywords:** *demand, forecasting, beer, COVID-19, pandemic, time series analysis*

## 1. UVOD

Pandemija COVID-19 značajno je uticala na sve aspekte globalnog poslovanja. Neometani procesi i operacije u proizvodnim sistemima i celokupnom lancu snabdevanja su od suštinskog značaja za uspešnost firme u veoma konkurentskim uslovima današnjice [27]. Međutim, kriza izazvana virusom COVID-19 otkrila je ranjivost i nisku otpornost sistema na iznenadne poremećaje. Pandemija kao događaj bez presedana je istovremeno izazvala problem koji utiče i na ponudu i na potražnju, što je otežalo poslovanje kompanijama i mogućnost istih da uspešno odgovore na postojeće izazove [9].

Većina istaknutih ekonomija širom sveta kao odgovor na pandemiju uvela je potpunu ili bar delimičnu obustavu privrednog i kulturnog života i pri čemu je od tada došlo i do skoka potražnje za osnovnim životnim namirnicama [17]. Kao rezultat širenja virusa, širom sveta su usvojene zdravstvene i bezbednosne mere koje su uticale na sve aspekte poslovanja i života uopšte kao što su [17]: delimično zatvaranje fabrika, zatvaranje granica, smanjenje i obustava međunarodnog i domaćeg vazdušnog saobraćaja, zatvaranje prodavnica, kafića i restorana, obustava turističkih aktivnosti, prelazak na vanrednu nastavu na daljinu u školama i na fakultetima, smanjenje broja zaposlenih na radnom mestu, uvođenje rada od kuće, ograničenja kretanja stanovništva, problemi sa nabavkom osnovnih dobara, ali i zatvaranje i izolacija. Virus COVID-19 uticao je i još uvek utiče na proizvodne sisteme i trgovinu na nacionalnom, pa čak i globalnom nivou.

Dostupnost i proizvodnja mnogih osnovnih artikala, kao što su hrana, namirnice i farmaceutski proizvodi, drastično je smanjena, a primećen je i veliki nesklad između ponude i potražnje [17].

Zbog trenutne situacije, globalizacija i efikasnost zasnovana na *lean* principima sve više se posmatraju u negativnom svetlu, a trenutno, barem kratkoročno, postoji određeni animozitet, nepoverenje i sumnja usmerena ka ovim principima [28]. Istovremeno, strategije i prakse održivosti ekonomije koje dolaze do izražaja su: podsticanje akcija kupovine lokalnih proizvoda i izgradnja poverenja u lokalne zajednice [25].

Pandemija COVID-19 takođe je značajno uticala na poslovanje indirektno kroz svoj uticaj na radnu snagu. Mnogi zaposleni su pogođeni niskim platama ili gubitkom prihoda. Sektori poput turizma, prehrambenih usluga, maloprodaje, proizvodnje i poslovnih i administrativnih aktivnosti su najviše izloženi riziku od nezaposlenosti [31].

Pandemija COVID-19 izazvala je ranjivost lanca vrednosti u različitim industrijama i negativno je uticala na proizvodne i uslužne operacije širom sveta [23, 1]. Svetska trgovina je opala u proseku za 13 do 32 odsto za razne proizvode [32]. Odluke koje su dovele do obustave proizvodnih aktivnosti imale su uticaj na globalni lanac snabdevanja, smanjenje prihoda i prekid saradnje, ali i delimičnu blokadu proizvodnje i isporuke proizvoda [10]. Međutim, i pored toga što je intenzitet globalne trgovine opao, kod pojedinih kategorija proizvoda tražnja je drastično porasla i ponuda nije mogla da se nosi sa tom situacijom, pa se postavilo pitanje opstanka snabdevanja konkretnih tržišta. Da bi izbegle ovu situaciju, proizvodne kompanije su preduzimale i preduzimaju korake da obezbede kontinuitet poslovanja. Mere uključuju međufunkcionalne kontrole i koordinaciju na globalnom i regionalnom nivou, povećane sigurnosne zalihe, poboljšane programe podrške korisnicima, mere oporavka radi poboljšanja potražnje i brze odgovore na promenljive izvore potražnje kako bi se osigurali izvori prihoda [13].

Jedan od lanaca snabdevanja koji je najviše pogođen je lanac snabdevanja hranom i pićem, u okviru koga je pozicionirano i pivo kao alkoholno piće. U [19] lanac snabdevanja hranom i pićem, pored lanaca snabdevanja koji su direktno vezani za zdravstvo, pominje se kao jedan od najteže pogođenih, uz upozorenje da panična kupovina određenih artikala može imati značajan uticaj na mnoge druge lance snabdevanja koji se bave sirovinama potrebnim za tu robu.

Postavlja se pitanje da li je i u kojoj meri pandemija COVID-19 uticala na potražnju za pivom? Kako bi se dobio odgovor na ovo pitanje, u radu je dalje analizirana vremenska serija koja predstavlja potražnju za pivom. Ostatak rada je strukturiran na sledeći način. U narednom poglavlju biće dat osvrt na teškoće predviđanja u uslovima vanrednih situacija kao što je pandemija COVID-19, kao i potencijalne pravce kretanja potražnje za alkoholnim pićima u ovakvim uslovima. U trećem poglavlju biće predstavljena metodologija kojom će se doći do odgovora na prethodno postavljeno istraživačko pitanje, a u četvrtom poglavlju biće prikazani rezultati analize. U okviru poslednjeg, petog poglavlja biće predstavljena diskusija dobijenih rezultata i zaključak samog rada.

## 2. POTRAŽNJA ZA ALKOHOLNIM PIĆIMA I PREDVIĐANJE U VREME PANDEMIJE

Pandemija je porementila obrasce potrošnje gotovo svih vrsta proizvoda i usluga. Jedan od prvih i teško pogođenih sektora bio je *HoReCa* sektor, što je bio rezultat mera u sprečavanju širenja virusa, među kojima su izolacija građana, zabrane javnih događaja i okupljanja i restrikcije kretanja. Navedene mere direktno su negativno uticale na potražnju u okviru navedenog sektora. Ovo je posledično stvorilo poremećaje na strani tražnje u industriji hrane i pića, a nije pomoglo ni panično ponašanje potrošača pri kupovini ovih proizvoda [11]. Model *just-in-time* koji se ranije pokazao veoma efikasnim, nije se pokazao kao dobro rešenje u uslovima ekstremnih poremećaja ponude i tražnje [11]. Kako bi se preduzećima moglo na adekvatan način upravljati u nestabilnim uslovima, u [8] se predlaže upotreba analitičkih alata i modela operacionih istraživanja. U oblasti operacionih istraživanja, predviđanje se ističe kao najčešća i najprisutnija disciplina i važan deo mnogih procesa koji se odnose na proces planiranja i donošenja odluka. Međutim, COVID-19 je značajno uticao na modele predviđanja umanjujući njihovu upotrebnu vrednost kroz uzrokovanje neadekvatnog snabdevanja sirovinama [27], menjajući obrasce potrošnje [18] i kroz svoj uticaj na globalnu trgovinu [17].

Brojni autori su se bavili problematikom predviđanja tražnje za alkoholnim pićima. U [15] predviđana je potražnja za votkom u SAD primenom naivne metode, tehnike vremenskih serija i metode veštačke inteligencije i mašinskog učenja, pokazujući da metode vremenskih serija i veštačke inteligencije i mašinskog učenja u velikoj meri nadmašuju naivne metode. Međutim, zaključak je bio da metode vremenskih serija imaju najbolje performanse za stabilne ili blago sezonske podatke, dok metode mašinskog učenja imaju bolje rezultate na podacima sa velikom varijansom. U [2] korišćeni su podaci od veleprodavaca alkoholnih pića u cilju predviđanja različitih tipova potražnje za alkoholom (votka i viski), fokusirajući se na

univarijantne modele vremenskih serija, kao i na multivarijantne i metode veštačke inteligencije. Zaključeno je da naprednije metode nadmašuju naivne metode [2].

Predviđanje potražnje za pivom i alkoholnim pićima uopšte tokom pandemije COVID-19 nije jednostavno, upravo zbog mera za suzbijanje virusa koje nameću različite zemlje, a koje se kreću od nametanja visokih poreza na alkohol, do zatvaranja maloprodajnih objekata i barova ili čak zabrane prodaje alkoholnih pića u potpunosti [6, 14]. Kako bi se procenio uticaj pandemije COVID-19 na potražnju piva, neophodno je sagledati društvene determinante pandemijske krize. Nekoliko kontradiktornih mehanizama vrše uticaj ili ka smanjenju ili ka povećanju konzumacije alkohola tokom pandemije. Ograničenja u društvenim interakcijama povezana su sa velikim i značajnim smanjenjem konzumacije piva u ugostiteljskim objektima, ali bez značajne promene u konzumaciji piva kod kuće [30]. Mlađa populacija koja najčešće konzumira alkoholna pića prilikom izlazaka značajno je smanjila konzumiranje alkohola [5], dok su, s druge strane, starije osobe koje više vole da piju kod kuće povećale svoju potrošnju [24]. Zanimljivo istraživanje sprovedeno je u [3] kako bi se objasnila četiri arhetipa potražnje koji su se pojavili tokom pandemije a koji opisuju prethodno pomenuto ponašanje. Pomenuti arhetipovi su: „skladištenje i potrošnja“, „skladištenje i čuvanje“, „sada kod kuće“, i „ne sada“. Za poslednja dva arhetipa koja sugerisu da ono što je ranije konzumirano napolju sada se ili uopšte ne konzumira ili se konzumira kod kuće, može se izneti argument da jedan ili oba imaju smisla kada je pitanju konzumacija piva. Postoje i ozbiljne indikacije da su mentalno zdravlje i ljudska psihologija uopšte igrali ulogu u povećanju konzumacije alkohola tokom ekonomski izazovnih i sveukupno stresnih vremena pandemije COVID-19 [5, 22]. Ljudi nisu objektivno procenjivali svoju situaciju, tvrdeći da su smanjili potrošnju kada to možda nije slučaj [4].

Različite nacionalnosti takođe pokazuju različita ponašanja u tom pogledu. U Nemačkoj i SAD ukupan prihod od prodaje alkoholnih pića je tokom pandemije povećan [3, 16]. Postoje i dokazi da se potrošnja u UK povećala [14, 24]. Međutim, u nekim zemljama kao što je Poljska, situacija nije u potpunosti jasna [7].

Javnim merama protiv COVID-19 u Republici Srbiji nije prekinuta prodaja alkoholnih pića. Međutim, odlaganje proslava, svečanosti i javnih skupova, kao i ograničenja za hotele, restorane i kafiće posredno su uticale na prodaju alkoholnih pića.

### 3. METODOLOGIJA

U ovom radu predstavljena je analiza vremenske serije o prodaji piva u Republici Srbiji u periodu pre i tokom pandemije COVID-19. Podaci su preuzeti sa tržišta glavnih maloprodajnih lanaca u Republici Srbiji na osnovu POS (engl. *Point of Sale*) izveštaja. Posmatrana vremenska serija predstavlja realne podatke koji sadrže 108 nedeljnih opservacija, od januara 2019. godine do kraja januara 2021. godine. Kako bi se detaljnije sagledale komponente vremenske serije (trend, sezonalnost i šum) prvo će biti urađena dekompozicija iste. Nakon toga, kako bi se simulirao proces odabira metode predviđanja vremenske serije, ona će biti podeljena u 2 seta, testni i validacioni, a na osnovu toga da li su opservacije pre ili nakon početka pandemije u Republici Srbiji. Za početak pandemije je uzet mesec mart 2020. godine.

Dalje, testni set je podeljen u 2 dela, gde će 8 različitih modela vremenskih serija biti istrenirano na prvom delu a validirano na drugom delu vremenske serije. Na taj način biće izabrana metoda koja će ponovo biti istrenirana na celokupnom testnom setu kako bi se videlo da li izabrana metoda može proizvesti dovoljno dobra predviđanja u okviru uslova pandemije, što su opservacije koje se odnose na validacioni set.

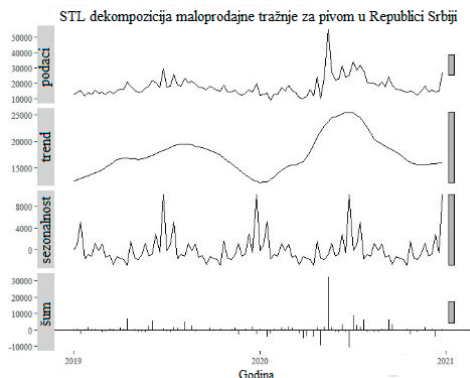
Kako bi se utvrdilo koja od 8 metoda predviđanja u normalnim uslovima daje najbolje rezultate, moguće je koristiti mnoge pokazatelje u vidu grešaka predviđanja [21], a u ovom radu je Tilova U statistika konkretno korišćena za odabir najbolje metode, uz navođenje drugih mera grešaka predviđanja koje će biti prikazane. Tilova U statistika je mera tačnosti koja naglašava kažnjavanje velikih grešaka, dok u isto vreme pruža uporediv indikator u odnosu na naivne metode predviđanja [29]. Uobičajeno, statistika U veća od 1 ukazuje na slabije performanse modela predviđanja u poređenju sa naivnom metodom i obrnuto, ako je statistika U manja od 1, model ima bolje performanse.

Prethodno su predstavljeni različiti scenariji koji dovode do različitih kretanja potražnje za pivom i alkoholnim pićima uopšte, tako da bi svaka pretpostavka o kretanju tražnje za vreme pandemije bila na nivou mišljenja i intuicije autora. Ipak, biće zanimljivo videti da li postoje određeni padovi, ili ono što je možda verovatnije, skokovi u potražnji u vreme početka i tokom pandemije u Republici Srbiji, zbog skupljanja zaliha u slučaju nestašica.

### 4. REZULTATI

Prethodno opisana vremenska serija najpre je analizirana primenom STL (engl. *Seasonal and Trend decomposition using Loess*) dekompozicije. U [12] dekompozicija vremenskih serija predlaže se kao korisno

sredstvo za poboljšanje razumevanja vremenske serije i za poboljšanje tačnosti predviđanja. Postoji nekoliko tipova dekompozicije vremenskih serija. U ovom radu korišćena je dekompozicija sezonskog trenda zasnovana na Loesu. STL dekompozicija ima mnoge prednosti [12] a neke od njih su sposobnost rada sa nedeljnim ili dvonedeljnim podacima i sezonskim karakteristikama, otpornost na ekstremne vrednosti i sposobnost da se nosi sa sezonskim promenama tokom vremena. STL dekompozicija vremenske serije o tražnji za pivom u Republici Srbiji je prikazana je na slici 1.



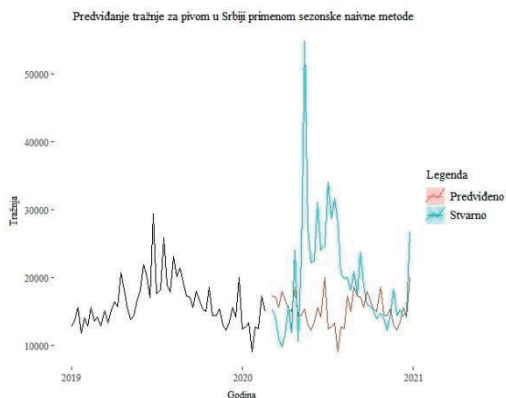
Slika 1: STL dekompozicija tražnje za pivom

Ispitivanje komponente trenda sa prethodne slike otkriva postojanje pikova u potrošnji u maloprodajnom segmentu srpskog tržišta piva. Dalje, rezidualna komponenta očigledno ima veće oscilacije u 2020. godini, što odgovara početku restriktivnih mera u Srbiji koje su se desile u martu 2020. Ovo sugeriše da su mere koje je nametnula vlada uticale na potražnju piva u Srbiji povećanjem iste. U tabeli 1 prikazane su greške predviđanja različitih metoda vremenskih serija.

Tabela 1: Analiza grešaka različitih metoda predviđanja

Metoda/Greška	ME	RMSE	MAE	MPE	MAPE	MASE	U
Naivna	<b>-21.090</b>	2723.049	2246.223	-4.843	18.310	1.817	0.592
Prosek	-3168.030	4177.435	3471.304	-29.548	31.305	2.807	0.913
Sezonska naivna	-271.782	<b>1674.270</b>	<b>1491.159</b>	-4.860	<b>12.404</b>	1.206	<b>0.372</b>
Eksponencijalno izravnjanje	-669.041	2803.956	2375.814	-9.929	20.168	1.921	0.559
Holt	-698.546	2804.875	2375.208	-10.149	20.200	1.921	0.558
Holt-Winters	-369.206	2879.321	2452.096	-7.838	20.660	<b>0.455</b>	0.528
ARIMA(0,0,0)	426.705	2724.507	2158.057	<b>-1.236</b>	16.930	1.745	0.653
ARIMA(0,1,1)(0,1,0)	2056.465	3562.349	3033.025	12.379	21.194	0.639	1.192

Na osnovu prethodne tabele dolazi se do zaključka da je za predviđanje tražnje za pivom u Srbiji u regularnim uslovima najbolja sezonska naivna metoda, zbog najmanje vrednosti U statistike. Analizom ostalih grešaka vidi se da je sezonska naivna metoda imala i najmanje RMSE, MAE i MAPE greške. Naivna metoda je imala najmanju ME, dok je ARIMA(0,0,0) imala najmanju MPE grešku. Stoga, sezonska naivna metoda dalje se koristi za predviđanje kretanja tražnje u uslovima pandemije, što je prikazano na slici 2.



Slika 2: Predviđanje tražnje primenom sezonske naivne metode

Prilikom predviđanja ukupne potražnje piva, vrednost U statistike je bila 1,008 što sugerise da je sezonska naivna metoda u uslovima nestabilnosti izazvane pandemijom predviđa tražnju sa jednakom greškom kao i naivna metoda. Ovo je vidljivo i na prethodno prikazanoj slici, gde se jasno vidi da je kretanje stvarne potražnje za pivom u vreme početka pandemije COVID-19 daleko veće od predviđene. Takođe moglo bi se reći i da se nakon nekog vremena stvarna tražnja vraća u okvire onoga što bi se moglo nazvati normalnim.

## 5. ZAKLJUČAK

Na osnovu prethodno prikazanih rezultata, može se reći da je postojao porast potražnje za pivom na srpskom tržištu od početka pandemije, a može se pretpostaviti i da je potrošnja alkohola, generalno, porasla.

Primenjene metode predviđanja vremenskih serija koje su davale dobre rezultate u uslovima normalnog poslovanja nisu uspele da odgovore na izazove poremećaja potražnje izazvanih pandemijom COVID-19. Međutim, dekompozicija vremenske serije pokazala se kao koristan alat za spoznaju vrednih informacija o osnovnim trendovima potražnje. Primena ovog alata bi mogla značajno povećati moć upravljanja procesom usklađivanja ponude i potražnje u budućnosti.

U daljem istraživanju, moglo bi se postaviti pitanje da li bi neke druge metode predviđanja, koje nisu obuhvaćene ovom studijom, bile bolje od odabranih metoda. Ova pitanja predstavljaju dobru osnovu za dalja istraživanja. Takođe, moglo bi se dalje istražiti da li određene vrste pakovanja imaju drugačije obrasce kretanja tražnje u uslovima pandemije. Treba napomenuti i da kada se pandemija COVID-19 smiri i svi aspekti poslovanja i života se vrate u normalu, stari obrasci kretanja potražnje mogu ponovo da se jave i utiču na potražnju za svim razmatranim kategorijama proizvoda. Međutim, novonastale okolnosti u vezi sa aktuelnim ratom u Ukrajini značajno utiču na disbalans ponude i tražnje, uzrokujući i globalnu inflaciju i energetska i ekonomsku krizu. Razlike u obrascima ponude, potražnje i potrošnje u okviru krize izazvane pandemijom sa jedne strane, i krize izazvane ratom sa druge, mogla bi biti zanimljiva tema za dalja istraživanja u primerima različitih proizvoda i usluga.

Nesumnjivo, kvantitativno predviđanje potražnje zajedno sa analizom velike količine podataka u realnom vremenu, *just-in-case* sistem poslovanja i povećanje otpornosti lanca snabdevanja predstavljaju ključne koncepte za FMCG kompanije u okruženju prinudno nametnutom zbog COVID-19 pandemije.

## LITERATURA

- [1] Ardolino, M., Bacchetti, A., & Ivanov, D. (2022). Analysis of the COVID-19 pandemic's impacts on manufacturing: a systematic literature review and future research agenda. *Operations Management Research*, 1-16.
- [2] Arora, T., Chandna, R., Conant, S., Sadler, B., & Slater, R. (2020). Demand Forecasting In Wholesale Alcohol Distribution: An Ensemble Approach. *SMU Data Science Review*, 3(1), 1–20.
- [3] Becdach, C., Brown, B., Halbardier, F., Henstorf, B., & Murphy, R. (2020). Rapidly forecasting demand and adapting commercial plans in a pandemic. *McKinsey & Company*, (April), 1–9.
- [4] Biddle, N., Edwards, A. Ben, & Gray, M. (2020). Alcohol consumption during the COVID- 19 period : May 2020 ANU Centre for Social Research and Methods, (May).
- [5] Callinan, S., Mojica-Perez, Y., ... Kuntsche, E. (2021). Purchasing, consumption, demographic and socioeconomic variables associated with shifts in alcohol consumption during the COVID-19 pandemic. *Drug and Alcohol Review*, 40(2), 183–191.
- [6] Callinan, S., Smit, K., Mojica-Perez, Y., D'Aquino, S., Moore, D., & Kuntsche, E. (2020). Shifts in alcohol consumption during the COVID-19 pandemic: early indications from Australia. *Addiction*.
- [7] Chodkiewicz, J., Talarowska, M., Miniszewska, J., Nawrocka, N., & Bilinski, P. (2020). Alcohol consumption reported during the COVID-19 pandemic: The initial stage. *International Journal of Environmental Research and Public Health*, 17(13), 1–11.
- [8] Choi, T.-M. (2021). Fighting against COVID-19: what operations research can help and the sense-and-respond framework. *Annals of Operations Research*, (0123456789), 1–17.
- [9] Fonseca, L. M., & Azevedo, A. L. (2020). COVID-19: outcomes for Global Supply Chains. *Management & Marketing*, 15(s1), 424-438.
- [10] Grida, M., Mohamed, R., & Zaied, A. N. H. (2020). Evaluate the impact of COVID-19 prevention policies on supply chain aspects under uncertainty. *Transportation Research Interdisciplinary Perspectives*, 8, 100240.

- [11] Hobbs, J. E. (2020). Food supply chains during the COVID-19 pandemic. *Canadian Journal of Agricultural Economics*, 68(2), 171–176.
- [12] Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- [13] Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. *International Journal of Production Research*, 58(10), 2904-2915.
- [14] Jacob, L., Smith, L., Armstrong, N. C., Yakkundi, A., Barnett, Y., Butler, L., ... Tully, M. A. (2021). Alcohol use and mental health during COVID-19 lockdown: A cross-sectional study in a sample of UK adults. *Drug and Alcohol Dependence*, 219 (November 2020).
- [15] Jiang, L., Rollins, K. M., Ludlow, M., & Sadler, B. (2020). Demand Forecasting for Alcoholic Beverage Distribution. *SMU Data Science Review*, 3(1).
- [16] Koopmann, A., Georgiadou, E., Kiefer, F., & Hillemacher, T. (2020). Did the General Population in Germany Drink More Alcohol during the COVID-19 Pandemic Lockdown? *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 55(6), 698–699.
- [17] Kumar, A., Luthra, S., Mangla, S. K., & Kazançoğlu, Y. (2020). COVID-19 impact on sustainable production and operations management. *Sustainable Operations and Computers*, 1, 1-7.
- [18] Maria del Rio-Chanona, R., Mealy, P., Pichler, A., Lafond, F., & Doyne Farmer, J. (2020). Supply and demand shocks in the COVID-19 pandemic: An industry and occupation perspective. *Oxford Review of Economic Policy*, 36, S94–S137.
- [19] Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., & Vasilakis, C. (2021). Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *European Journal of Operational Research*, 290(1), 99–115.
- [20] Qin, X., Godil, D. I., Khan, M. K., Sarwat, S., Alam, S., & Janjua, L. (2021). Investigating the effects of COVID-19 and public health expenditure on global supply chain operations: an empirical study. *Operations Management Research*, 1-13.
- [21] Rakićević, Z., & Vujošević, M. (2015). Focus Forecasting in Supply Chain: The Case Study of Fast Moving Consumer Goods Company in Serbia. *Serbian Journal of Management*, 10(1), 3–17.
- [22] Rehm, J., Kilian, C., Ferreira-Borges, C., Jernigan, D., Monteiro, M., Parry, C. D. H., ... Manthey, J. (2020). Alcohol use in times of the COVID 19: Implications for monitoring and policy. *Drug and Alcohol Review*, 39(4), 301–304.
- [23] Sahoo, P., & Ashwani. (2020). COVID-19 and Indian economy: Impact on growth, manufacturing, trade and MSME sector. *Global Business Review*, 21(5), 1159-1183.
- [24] Sallie, S. N., Ritou, V., Bowden-Jones, H., & Voon, V. (2020). Assessing international alcohol consumption patterns during isolation from the COVID-19 pandemic using an online survey: Highlighting negative emotionality mechanisms. *BMJ Open*, 10(11), 1–10.
- [25] Sarkis, J. (2020). Supply chain sustainability: learning from the COVID-19 pandemic. *International Journal of Operations & Production Management*, 41(1), pp. 63-73.
- [26] Shahed, K. S., Azeem, A., Ali, S. M., & Moktadir, M. (2021). A supply chain disruption risk mitigation model to manage COVID-19 pandemic risk. *Environmental Science and Pollution Research*, 1-16.
- [27] Sharma, M., Luthra, S., Joshi, S., & Kumar, A. (2022). Developing a framework for enhancing survivability of sustainable supply chains during and post-COVID-19 pandemic. *International Journal of Logistics Research and Applications*, 25(4-5), 433-453.
- [28] Singh, S., Kumar, R., Panchal, R., & Tiwari, M. K. (2021). Impact of COVID-19 on logistics systems and disruptions in food supply chain. *International Journal of Production Research*, 59(7), 1993-2008.
- [29] Small, G., & Wong, R. (2002, January). The validity of forecasting. In: A Paper for Presentation at the Pacific Rim Real Estate Society International Conference, Christchurch, New Zealand (pp. 1-14).
- [30] Vandenberg, B., Livingston, M., & O'Brien, K. (2020). When the pubs closed: beer consumption before and after the first and second waves of COVID-19 in Australia. *Addiction*, 1–7.
- [31] Verma, S., & Gustafsson, A. (2020). Investigating the emerging COVID-19 research trends in the field of business and management: A bibliometric analysis approach. *Journal of Business Research*, 118, 253-261.
- [32] World Trade Organization (WTO), 2020. Trade set to plunge as COVID-19 pandemic upends global economy. Link: [https://www.wto.org/english/news\\_e/pres20\\_e/pr855\\_e.htm](https://www.wto.org/english/news_e/pres20_e/pr855_e.htm)

## ANALITIKA TESTIRANJA

### TESTING ANALYTICS

DRAGAN AZDEJKOVIĆ<sup>1</sup>, SLAVICA MANIĆ<sup>2</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, dragan.azdejkovic@ekof.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, slavica.manic@ekof.bg.ac.rs

**Rezime:** Pre pokretanja velikog projekta ponekad postoji mogućnost probnog testiranja njegove profitabilnosti. Benefiti probnog testiranja u ovom radu analizirani su pomoću drveta odlučivanja, kontrolne table i Monte Karlo simulacije. Sastavni deo teksta su i radne tabele u kojima je impementirana jedna studija slučaja, koja bi trebalo da posluži kao primer upotrebe kombinovane metodologije u svrhu donošenja optimalne poslovne odluke u situacijama kad je probno testiranje izvodljivo.

**Ključne reči:** Poslovna analitika, drvo odlučivanja, Monte Karlo, kontrolna tabla, testiranje

**Abstract:** Before launching a large project, there is sometimes the possibility of a trial test of its profitability. The benefits of trial testing in this paper are analyzed using a decision tree, dashboard and Monte Carlo simulation. An integral part of the text are worksheets in which a case study is implemented. It is expected to serve as an example of the use of combined methodology for the purpose of making an optimal business decision in situations where trial testing is feasible.

**Keywords:** Business analytics, decision tree, Monte Carlo, dashboards, testing

### 1. UVODNA RAZMATRANJA

Pre pokretanja velikog projekta ponekad postoji mogućnost probnog testiranja njegove profitabilnosti. Problemi ovog tipa imaju različite varijetete i uglavnom se rešavaju primenom drveta odlučivanja. Glavni nedostatak ovog pristupa je to što su parametri problema fiksirani, što ne korespondira realnosti, budući da su za donosioca odluke neki parametri slučajne promenljive. Verovatno bi se složenim matematičkim izračunavanjem moglo doći do eksplicitnih formula za očekivani profit i optimalnu strategiju, ali ovakav pristup generiše dodatni problem jer svaka, čak i neznatna, promena modela dovodi do novih, složenijih, kalkulacija.

U literaturi postoji veći broj takozvanih klon studija slučaja ovog tipa sa fiksnim parametrima [1, 2, 3]. Uočava se, između ostalog, raznovrsnost slučajeva u kojima su korišćene Monte Karlo simulacije, samostalno ili u kombinaciji sa drvetom odlučivanja. Tako su, na primer, koristeći metodologiju teorije igara sa stohastičkim agentima, modeliranjem nesigurnih parametara metodom Monte Karlo, neki autori [5] rešavali problem predviđanja potražnje i vrednovanja poslovanja za industrije povezane sa računarima. Drugi su opisali metod za modeliranje verovatnoća i korisnosti u specifikaciji drveta odlučivanja primenom Monte Karlo simulacija u kontekstu rešavanja medicinskog problema [4, 6]. Takođe, u jednoj od studija [7] koristi se kombinacija metoda drveta odlučivanja i Monte Karlo metoda za procenu rudarskog projekta gde se razmatra odluka da li da se otvori rudnik ili ne.

Imajuću sve navedno u vidu, u ovom radu se testira mogućnost primene kombinacije metoda (drveta odlučivanja i Monte Karlo simulacija) i analiza rezultata preko eksel radnih tabela u svrhu donošenja optimalne poslovne odluke u situacijama kad je probno testiranje izvodljivo.

### 2. METODOLOGIJA I OPIS PROBLEMA

Klasa problema odlučivanja opisana u uvodu biće rešavana primenom drveta odlučivanja i simulacijama. Za analizu jedne studije slučaja iz klase problema biće korišćene unakrsne tabele eksela.

Zbog velikih fiksnih troškova i/ili rizika, pre pokretanja velikog projekta, kompanije često koriste strategiju da prvo testiraju profitabilnost projekta, pa tek potom odlučuju da li da nastave sa projektom ili da odustanu.

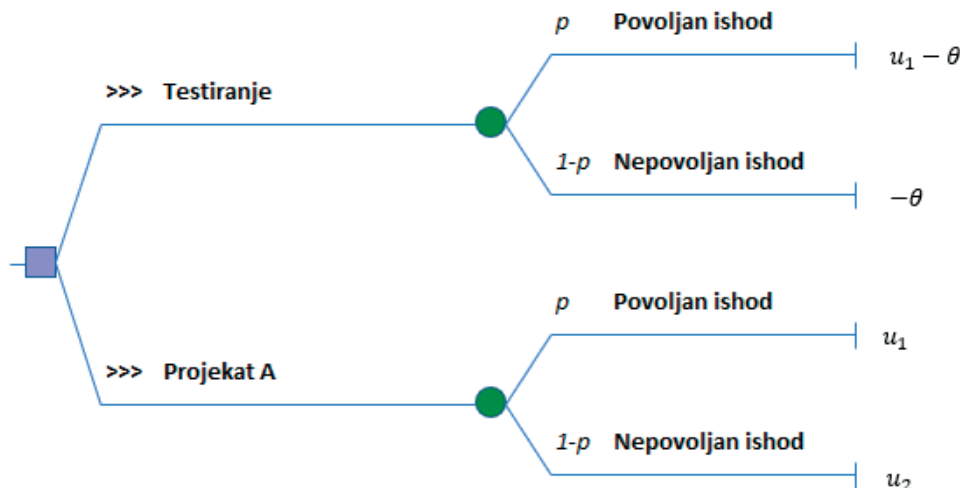
Elementarni problem ovog oblika prikazan je na Slici 1, gde su

$u_1$  – profit ako se ostvari povoljan ishod

$u_2$  – negativni profit (odnosno gubitak) ako se ostvari nepovoljan ishod

$p$  – verovatnoća povoljnog ishoda

$\theta$  – cena testiranja



Slika 1: Elementarni problem testiranja

Za elementarni problem testiranja ovog tipa moguće je naći eksplicitnu formulu kojom se opisuje optimalna očekivana korisnost:

$$E[A] = \max \{ pu_1 - \theta, pu_1 + (1 - p)u_2 \} = \max \{ -\theta, (1 - p)u_2 \} = \min \{ \theta, (p - 1)u_2 \} \quad (1)$$

Optimalna strategija zavisila bi od toga da li se minimum ostvaruje u  $\theta$  ili  $(p - 1)u_2$ , odnosno da li su manji troškovi probnog testiranja ili očekivani gubitak pri nepovoljnom ishodu.

Kompleksnost ovog problema može se povećati na sledeće načine:

- Pokretanje projekta bez testiranja može rezultirati većim brojem slučajnih ishoda.
- Testiranje može dati mnoštvo različitih, ali ne nužno pouzdanih, informacija.
- Testiranje se može odvijati u više etapa sa elementima iz prethodnog stava.

Tada bi bilo moguće izvesti sličnu formulu zatvorenog tipa, ali bi ona zbog kompleksnosti bila praktično neupotrebljiva.

Umesto toga, moguće je pokazati da se pitanje izbora optimalne strategije može pojednostaviti čak i u situaciji kad su neki podaci u modelu slučajne promenljive sa poznatim raspodelama (prirodna pretpostavka). U tu svrhu konstruiše se model koji će biti vizuelno jednostavan za korisnika, odnosno, isti se implementira kroz radne tabele eksela.

### 3. STUDIJA SLUČAJA

Proizvođač elektronskih igračaka (nazovimo ga ET), čija su specijalnost vozila na daljinsko upravljanje razmatra predstavljanje novog proizvoda, broda na daljinsko upravljanje, tzv. ET-1. Da bi se proizvod plasirao na tržište potrebno je izgraditi veliko postrojenje čija je cena 13 miliona dolara. Takođe, pre nego što proizvod bude pušten u prodaju, neophodno je napraviti prototip i testirati koliko je bezbedan. Prototip se može napraviti i bez postrojenja po ceni od 1,5 miliona dolara, ali ukoliko se prototip napravi nakon izgradnje postrojenja, njegovi troškovi su zanemarljivi (mogu se tretirati kao nula). Postoji nesigurnost u vezi



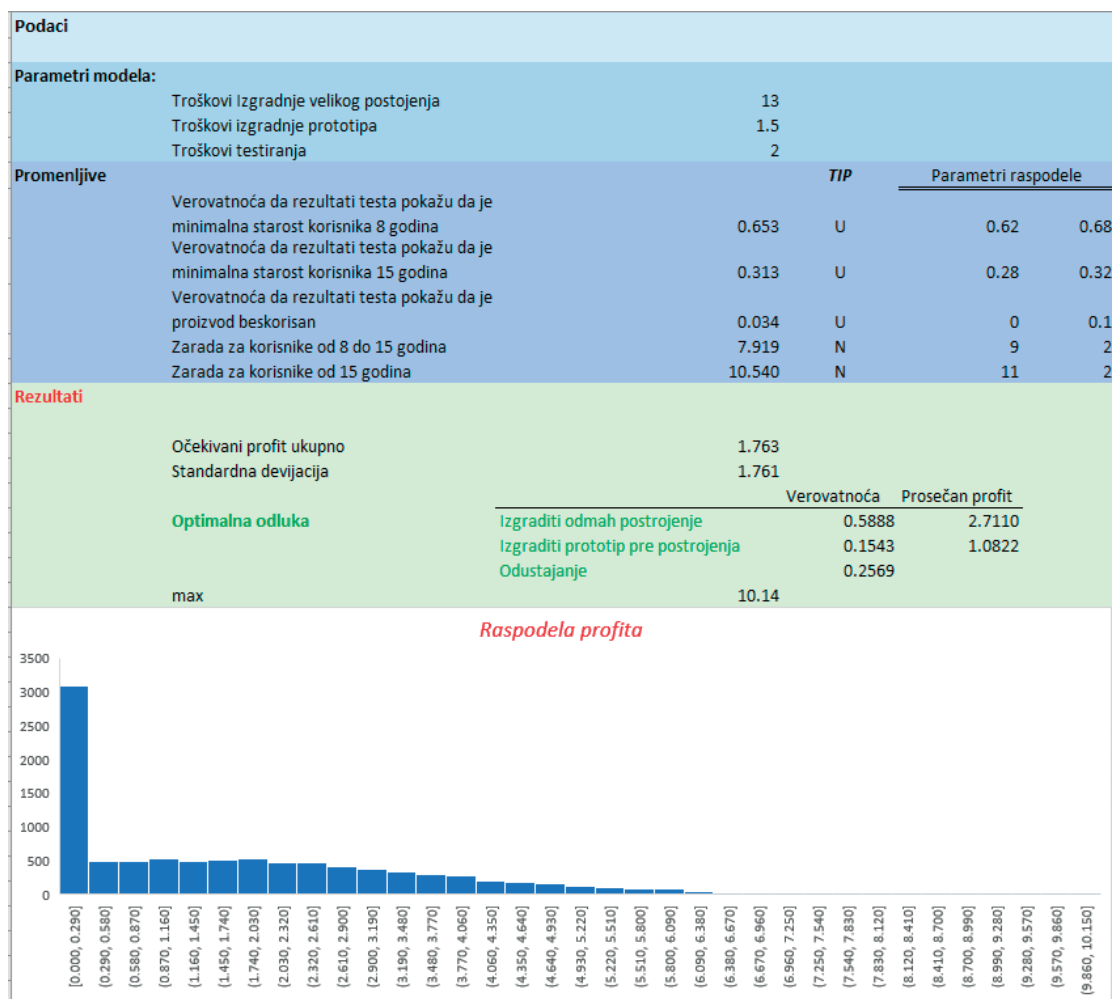
sa time koju će ocenu za bezbednost ET-1 dobiti. Ona bi mogla imati značajan uticaj na potražnju, jer će niža ocena za bezbednost povećati minimalnu starost korisnika. Troškovi testiranja bezbednosti iznose 2 miliona dolara. Procenjuje se da rezultati testa imaju  $P\%$  šanse da pokažu da je minimalna starost korisnika 8 godina,  $Q\%$  šanse da je minimalna starost 15 godina i  $(100 - P - Q)\%$  šanse da će proizvod biti proglašen nebezbednim – što znači da se uopšte neće prodavati. (Troškovi unapređenja nivoa bezbednosti već potpuno dizajniranog proizvoda smatraju se neprihvatljivim.) Nakon uspešnog testa bezbednosti proizvod se može pustiti na tržište po ceni od 2,5 miliona dolara. Pretpostavimo da su  $P$  i  $Q$  slučajne promenljive se uniformnim raspodelama, redom,  $U(62,68)$  i  $U(28,32)$ .

Postoji i nesigurnost oko potražnje, koja će imati presudan uticaj na eventualnu dobit. Ako proizvod bude dostupan korisnicima od 8 do 14 godina tada je zarada slučajna promenljiva sa normalnom raspodelom  $N(9, 2^2)$ . Slično, zarada za korisnike starije od 15 godina je slučajna promenljiva sa normalnom raspodelom  $N(11, 2^2)$ . (Rezultat potražnje je nezavisan među starosnim grupama i vrednosti su u milionima dolara.) U zaradi se u obzir ne uzimaju prethodno navedeni troškovi; ona se meri u odnosu na očekivane uslove trenutnih vrednosti, tako da se može direktno uporediti sa troškovima.

Koji je optimalan plan rada ET-a? Koja je trenutna očekivana ekonomska vrednost ET-ovog projekta?

### 3.1. Kontrolna tabla

Za ovu studiju slučaja, ali i generalno za probleme poslovne analitike, po našem mišljenju, najpogodniji alat su radne tabele eksela. Kompletno rešenje problema dostupno je na internet adresi [SYM-OP-IS-2022](#). Na Slici 2. je kontrolna tabla dizajnirana za ovu studiju slučaja.



Slika 2: Kontrolna tabla studije slučaja

### 3.2. Interpretacija dobijenih rezultata

Generalni rezultat možemo interpretirati u terminima verovatnoće. Pod datim pretpostavkama, u oko 59% situacija najbolje je sagraditi veliko postrojenje i napraviti prototip, jer je očekivani profit 2,7 miliona dolara, u približno 15% situacija treba prvo napraviti prototip (bez pomenutog postrojenja) i onda odlučivati u zavisnosti od rezultata testiranja šta dalje, što donosi očekivani profit od 1,07 milona dolara, dok je u ostalim slučajevima (26% situacija) najbolje odustati od proizvodnje. Donosilac odluke, na osnovu ovih informacija i bilansa kompanije treba da donese konačnu odluku.

### 4. ZAKLJUČAK

Poslovni problemi odlučivanja prirodno uključuju podatke koji nisu izvesni. U radu je pokazano da je moguće i izvodljivo uspešno se nositi sa tim problemom uključivanjem simulacija. Alati poslovne analitike, a posebno kontrolna tabla, pružaju dodatne mogućnosti - da se interaktivno menjaju ulazni parametri i prate izlazni podaci, kako vizuelno tako i analitički. Sve navedeno olakšava proces odlučivanja u smislu da donosilac odluke može vrlo brzo da analizira veliki broj scenarija i isto tako brzo donosi odgovarajuće odluke. U daljem radu na ovom problemu mogle bi se uvesti pretpostavke da su neki podaci rasplinuti (*fuzzy*) ili da se kombinuju podaci različitog tipa (slučajne i fazi promenljive).

### LITERATURA

- [1] Camm, J. D., Cochran, J. J., Fry, M. J., & Ohlmann, J. W. (2020). *Business analytics*. Cengage Learning.
- [2] Tadelis, S. (2013). *Game theory: an introduction*. Princeton University Press.
- [3] Hillier, F. S., & Lieberman, G. J. (2015). *Introduction to Operations Research*, McGraw Hill. Inc., New York, 4-15.
- [4] Critchfield, G. C., & Willard, K. E. (1986). Probabilistic analysis of decision trees using Monte Carlo simulation. *Medical Decision Making*, 6(2), 85-92.
- [5] Ikeda, Y., Kubo, O., & Kobayashi, Y. (2004). Forecast of business performance using an agent-based model and its application to a decision tree Monte Carlo business valuation. *Physica A: Statistical Mechanics and Its Applications*, 344(1-2), 87-94.
- [6] Dibra D, (2015). *Project Valuation and Decision Making under Risk and Uncertainty applying Decision Tree Analysis and Monte Carlo Simulation*. Books on Demand [Norderstedt].
- [7] Topal, E. (2008). Evaluation of a mining project using discounted cash flow analysis, decision tree analysis, Monte Carlo simulation and real options using an example. *International Journal of Mining and Mineral Engineering*, 1(1), 62-76.

## THE COVID-19 INFLUENCE ON SMES IN A DEVELOPING COUNTRY: A DELPHI METHOD

MIMOZA ARIFI<sup>1</sup>, VIOLETA CVETKOSKA<sup>2</sup>

<sup>1</sup> Ss. Cyril and Methodious University in Skopje, Faculty of Economics - Skopje, mimozzaarifi@gmail.com

<sup>2</sup> Ss. Cyril and Methodious University in Skopje, Faculty of Economics - Skopje, vcvetkoska@eccf.ukim.edu.mk

**Abstract:** *With the decisions that enterprises make today, they write their future whether they will work successfully or will be surpassed by the competition and in the worst case will disappear from the market. However, what shifted the normal course of operation of the enterprises in the first quarter of 2020 is the coping with the health crisis caused by the COVID-19 pandemic, when many enterprises worldwide laid off their employees, and millions of new layoffs are announced. The aim of this paper is to examine the judgments of panel experts on how SMEs in the Republic of North Macedonia are coping with the crisis, what measures can help SMEs overcome the negative consequences of the pandemic, and which factors affect their decision to lay off employees the most. The research is conducted using the Delphi method. The obtained results are presented and analyzed and will serve the policy makers in shaping the needed measures to improve the operating landscape of SMEs.*

**Keywords:** *SMEs, COVID-19, Expert's Judgements, Delphi Method, Developing Country.*

### 1. INTRODUCTION

The global COVID-19 pandemic has had vast consequences for business and consumer behavior. With technology at their disposal, businesses in a variety of industries have been compelled to close their doors to their offices and shops and adapt to more remote working circumstances.

[1] analyzes 69 SME crisis-related articles before COVID-19 and proposes ways to overcome the economic downturns in the finance, strategy, and institutional environment fields. To examine the challenges of returning to work and the policy requirements imposed by COVID-19, [2] studies 4807 SMES in Sichuan Province, China. Due to a shortage of epidemic mitigation materials, employees' inability to return to work, disrupted supply chains, and diminished market demand, most SMEs were unable to continue operations. [3] examine on the early effects of COVID-19 on 367 agrifood MSMEs in 17 countries. Their findings demonstrate that the pandemic has had an impact on 94.3 % of the sample's operations, first by reducing sales and then by limiting access to inputs and finance due to lack of financial reserves. Staffing issues are also frequently mentioned.

SMEs that use predictive analytics techniques can predict what will happen in the future. In cases when no historical data is available to make the prediction, a judgmental approach can be applied. Delphi is a qualitative method that is used for forecasting the likelihood of future events and their impact on the problem of interest, taking into account a panel of experts who need to respond to a series of questionnaires, each conducted in a separate round [4].

In this paper, we examine the judgments of panel experts on how SMEs in one developing country, the Republic of North Macedonia, are coping with the crisis; what measures can help SMEs overcome the negative consequences of the pandemic; and which factors affect their decision to lay off employees the most. The research was conducted using the Delphi method and is the first of its type in our country. The obtained results have been visualized and analyzed, and they will help policymakers shape the necessary measures to improve the operating landscape of SMEs.

The paper is organized as follows. Section 2 describes the methodology and data that were used. In Section 3 are presented and analyzed the obtained results, while the conclusion is given in Section 4.

### 2. METHODOLOGY AND DATA

The qualitative forecasting approach places emphasis on using experience and intuition and can be applied when there is no available historical data, or when the forecast refers to a longer future period [4]. One of the most used qualitative (judgmental) forecasting methods is Delphi. This method is used to achieve

consensus for the best solution to a complex issue, and it is appropriate when there is a geographical distance between the experts or when they have conflicting opinions. So, in the first round, the facilitator will send a questionnaire with more general questions in order to collect the panel experts' opinions. Then, the facilitator should compile the responses (those that are same and those that have the same meaning but are worded differently) and calculate the frequency of the responses. The questions for round two should be formulated based on the answers from round one. Therefore, the facilitator should send the new questionnaire in round two, where each respondent should see their own answer and the answers of the other respondents, but their identity will be kept confidential in all rounds. Preserving anonymity is crucial in applying this method in order to obtain unbiased solutions. In the second round, the respondents need to prioritize the answers and give a textual explanation. Again, the facilitator collects and summarizes the data and repeats the round until a consensus is achieved. Usually there are three rounds, but this number varies depending on the consensus achieved. The respondents are allowed to revise their responses in the next round, but if they make any revisions, they are required to explain the reason for that. The output of the conducted rounds in the Delphi method is the facilitator's final report, and it contains the respondents' answers.

In this paper we use a Delphi-based questionnaire to achieve our main objective. The questionnaire is distributed by e-mail to five panel experts in this area (the group is heterogeneous) and after answering the questions, they should send it by e-mail to the authors. Because in the use of this method, the participants must be anonymous in order to prevail in obtaining biased answers or solutions, each participant was coded with a different color that was only explained to him/her.

The first round consists of five open questions about the identification of sectors that will have the highest losses, opinions on the operation of small and medium enterprises in North Macedonia, measures that would help SMEs overcome the negative consequences of the pandemic, and factors that influence the most those enterprises that lay off employees. After obtaining the responses, we will compile them (those that are the same and those that have the same meaning but are worded differently).

The questions for round two are formulated based on the answers from round one. Therefore, we will send the new questionnaire in round two, where each respondent should see their own answer and the answers of the other respondents, but their identity will be kept confidential in all rounds. In this round, the respondents will need to answer "I agree" or "I do not agree" to all questions, except for the industry that will have the highest losses and the factors that influence employee layoffs, where they need to give a grade from 1–5 (1 being the least important, and 5 being the most important). Again, we will collect and summarize the data and repeat the round until a consensus is achieved. The number of rounds will depend on the consensus achieved.

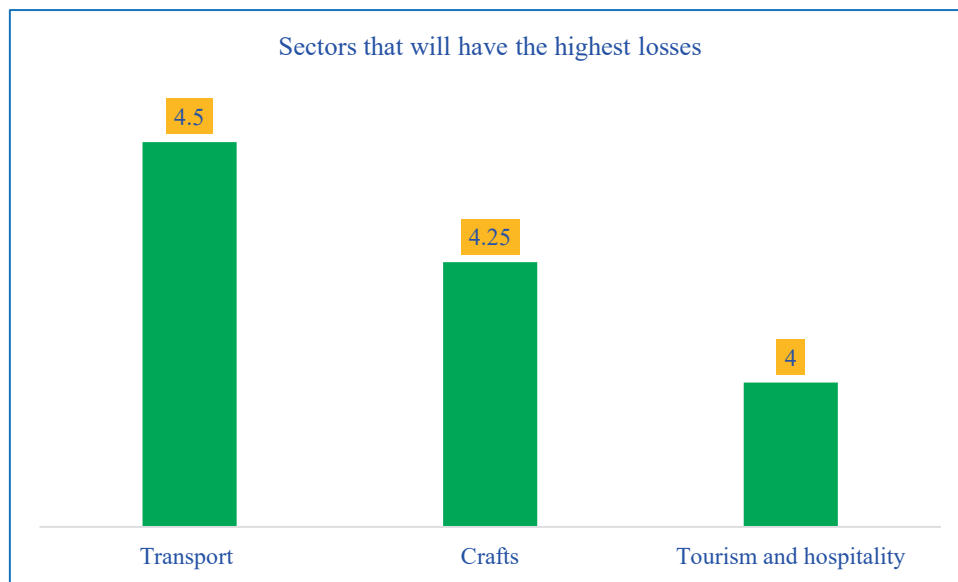
The first questionnaire was filled out in June 2020, and the second in the first half of July. In the third round, consensus was achieved, so the survey ended in July. It was answered by four panel experts. The obtained results are presented and analyzed in the Section that follows.

### 3. RESULTS AND ANALYSIS

According to the obtained results, the SMEs in the transport industry will have the most losses ( $\bar{x} = 4.50$ ), followed by crafts ( $\bar{x} = 4.25$ ) and the tourism and hospitality sector ( $\bar{x} = 4.00$ ) (Figure 1).

All respondents agree that how SMEs will cope with the crisis depends on how the neighboring countries, especially the countries with which North Macedonia has the largest trade exchange, will cope with the COVID-19 pandemic; from the capital investment projects by the state and the dynamics of their execution at the central and local levels; lack of liquidity; loss-making operations; and even if the coronavirus completely disappears, tourism and hospitality will have losses.

All of the experts agree that the reduction of work, which implies firms to be unable to pay their employees' salaries, had the greatest impact on the decision to lay off employees. In addition, three of the experts agreed that the state did not react quickly enough to help small and medium-sized enterprises with the set of measures adopted to cover the salaries of employees.



**Figure 1:** Average grade of importance for sectors that will have the highest losses

The measures that would help SMEs overcome the negative consequences of the crisis are: the establishment of fund to help to alleviate the negative effect of the coronavirus ( $\bar{x} = 4.75$ ), direct assistance from the state with interest-free credit lines through the Development Bank of North Macedonia for existing businesses ( $\bar{x} = 4.25$ ), the establishment of strategic subsidy packages ( $\bar{x} = 4.00$ ), the renewal of the campaign for purchasing domestic products to rise public awareness about the benefits of buying domestic products and their impact on the sustainability and development of domestic companies ( $\bar{x} = 4.00$ ), to encourage the processing industry of primary agricultural products ( $\bar{x} = 4.00$ ), the assistance from the state to ensure the placement of basic agricultural products in foreign markets ( $\bar{x} = 4.00$ ), etc. (Figure 2).

#	Measure	Average grade of importance
1	It is critical to establish a fund to help alleviate the negative effects of the coronavirus.	4.75
2	Direct assistance from the state with interest-free credit lines through the Development Bank of North Macedonia for existing businesses.	4.25
3	Establishing strategic subsidy packages.	4.00
4	Renewal of the campaign for purchasing domestic products to raise public awareness about the benefits of buying domestic products and their impact on the sustainability and development of domestic companies.	4.00
5	Encouraging the processing industry of primary agricultural products.	4.00
6	Assistance from the state to ensure the placement of basic agricultural products in foreign markets.	4.00
7	Encouraging domestic tourism.	3.75
8	Subsidies to return the population to rural areas by encouraging small and medium-sized businesses depending on the region.	3.75
9	Deferred payment of VAT by legal entities for the next six months.	3.50
10	Income tax exemption.	3.50
11	Encouraging cooperation among domestic SMEs by forming clusters.	3.25
12	Organizing fairs with the participation of only domestic companies.	3.25
13	Use of services from domestic companies.	3.25
14	Encouragement of infrastructure projects and other construction projects of public interest by the state in all regions at central and local levels.	3.25
15	Encouragement and planning of primary agricultural production and higher subsidies by the state for beginners without limitation on the age limit of entrepreneurs.	3.25
16	Free education and direct financial support from the state to restructure companies that will not be able to survive, especially those whose businesses are affected by the pandemic.	3.00

**Figure 1:** Average grade of importance for needed measures to overcome the negative consequences of the pandemic for SMEs

#### 4. CONCLUSION

In this study, for the first time in North Macedonia we explore the opinions of a panel of experts on how SMEs in the country are dealing with the COVID-19 crisis, what measures might help SMEs overcome the negative impacts of the pandemic, and which factors have the most influence on their decision to lay off staff. The research is being carried out using a three-round Delphi method. The obtained results provide important insights for economic policymakers.

The use of digital technology by SMEs to deal with the aftermath of extraordinary events like Covid-19 is investigated in [5]. Using a mediation analysis based on nonlinear probability models, [6] study the effects of SMEs' openness to Industry 4.0 on the perceived production recovery following the COVID-19 pandemic. The number of technologies adopted serves as a gauge of a businesses readiness for Industry 4.0. The study focuses on 2622 SMEs in Italy that manufacture goods. According to their research, openness to Industry 4.0 has both a short-term (by 2021) and medium-term (within 2022 and 2023) positive and significant impact on the perception of a production recovery. The perceived medium-term production recovery can be mediated by the classical reorganization, whereas the perceived short-term production recovery can be mediated by the digital reorganization. The Covid-19 pandemic mitigates positive impact on the utilization of foreign and domestic digital platforms on their international orientation through a sample of 372 observations from 250 Chinese SMEs [7]. The findings shed light on innovation as a tool for SMEs to survive both during and after the COVID-19 contingency, and it is determined that the utilization of digital resources is the major driver for networking and research-based product design in the setting of "social distance" [8]. According to findings in the developed theoretical model in [9] which was validated using a structural equation modeling (SEM) technique on data collected from 327 respondents from Indian SMEs, big data innovation aids in the sustainability of SME supply chains, as do internet of things (IoT) optimization and big data embedded supply chains. Additionally, SMEs leadership is essential for maintaining supply chain operations in times of crisis. The results of the statistical analysis conducted on data from 98 French and Algerian companies of various sizes reveal that Lean management tools and Industry 4.0 technologies tend to be strongly linked, and that understanding this relationship can enhance the organizational dimensions of leadership, strategy, operation, and production [10]. The best-practice representation of Japanese lean management concepts can be found in [11]. The 6Ps maturity model (MM), which [12] propose, intends to help manufacturing SMEs in the creation of their roadmaps toward Digitalized Manufacturing 4.0. The 6Ps maturity model is based on 36 Industry 4.0-related areas, a 5-level scale of digital maturity, and 6 primary socio-business and technical dimensions (i.e., Product-Services, Processes, Platform, People, Partnership, and Performance).

The COVID-19 pandemic is a "black swan" event, and SMEs who have embraced Industry 4.0 technologies like big data analytics, artificial intelligence, IoT, and blockchain technology can deal with it better, survive in the market over the long run, and gain a competitive advantage.

We propose to focus future study on the use of Industry 4.0 technology by SMEs in North Macedonia and its neighbors, as well as the impact on their operations and financial performance.

#### REFERENCES

- [1] Eggers, F. (2020). Masters of disasters? Challenges and opportunities for SMEs in times of crisis. *Journal of Business Research*, 116, 199–208. doi: 10.1016/j.jbusres.2020.05.025
- [2] Lu, Y., Wu, J., Peng, J., & Lu, L. (2020). The perceived impact of the Covid-19 epidemic: Evidence from a sample of 4807 SMEs in Sichuan Province, China. *Environmental Hazards*, 19(4), 323–340. doi: 10.1080/17477891.2020.1763902
- [3] Nordhagen, S., Igbeka, U., Rowlands, H., Shine, R. S., Heneghan, E., & Tench, J. (2021). COVID-19 and small enterprises in the food supply chain: Early impacts and implications for longer-term food system resilience in low- and middle-income countries. *World Development*, 141, 105405. doi: 10.1016/j.worlddev.2021.105405
- [4] Cvetkoska, V. (2022). Business Analytics. Skopje: Stobi Trejd Doel.
- [5] Papadopoulos, T., Baltas, K. N., & Balta, M. E. (2020). The use of digital technologies by small and medium enterprises during COVID-19: Implications for theory and practice. *International Journal of Information Management*, 55, 102192. doi: 10.1016/j.ijinfomgt.2020.102192
- [6] Cugno, M., Castagnoli, R., Büchi, G., & Pini, M. (2022). Industry 4.0 and production recovery in the covid era. *Technovation*, 114, 102443. <https://doi.org/10.1016/j.technovation.2021.102443>

- 
- [7] Lee, J. Y., Yang, Y. S., Ghauri, P. N., & Park, B. I. (2022). The impact of social media and digital platforms experience on SME international orientation: The moderating role of covid-19 pandemic. *Journal of International Management*, 100950. <https://doi.org/10.1016/j.intman.2022.100950>
- [8] Caballero-Morales, S.-O. (2021). Innovation as recovery strategy for SMEs in emerging economies during the COVID-19 pandemic. *Research in International Business and Finance*, 57, 101396. <https://doi.org/10.1016/j.ribaf.2021.101396>
- [9] Chatterjee, S., Chaudhuri, R., Shah, M., & Maheshwari, P. (2022). Big data driven innovation for sustaining SME supply chain operation in post COVID-19 scenario: Moderating role of SME technology leadership. *Computers & Industrial Engineering*, 168, 108058. <https://doi.org/10.1016/j.cie.2022.108058>
- [10] Najwa, E., Bertrand, R., Yassine, M., Fernandes, G., Abdeen, M., & Souad, S. (2022). Lean 4.0 tools and technologies to improve companies' maturity level: The COVID-19 context. *Procedia Computer Science*, 196, 207–216. <https://doi.org/10.1016/j.procs.2021.12.007>
- [11] Helmold, M., Yılmaz, A. K., Flouris, T., Winner, T., Cvetkoska, V., & Dathe, T. (2022). *Lean management, kaizen, kata and keiretsu: Best-practice examples and industry insights from japanese concepts* (1st ed. 2022 edition). Springer.
- [12] Spaltini, M., Acerbi, F., Pinzone, M., Gusmeroli, S., & Taisch, M. (2022). Defining the roadmap towards industry 4. 0: The 6ps maturity model for manufacturing smes. *Procedia CIRP*, 105, 631–636. <https://doi.org/10.1016/j.procir.2022.02.105>







# PRIMENE OI U ODBRANI

---

## OR APPLICATIONS IN MILITARY DEFENCE



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business





## DEFINISANJE KRITERIJUMA ZA IDENTIFIKOVANJE BEZBEDNOSNIH IZAZOVA RIZIKA I PRETNJI PRIMENOM MATEMATIČKOG MODELA VIŠEKRITERIJUMSKOG ODLUČIVANJA

### DEFINING CRITERIA FOR IDENTIFYING SECURITY CHALLENGES, RISKS AND THREATS USING THE MATHEMATICAL MODEL OF MULTICRITERIA DECISION MAKING

DRAGAN BOJANIĆ<sup>1</sup>, MARINA BOJANIĆ<sup>2</sup>, VLADIMIR RISTIĆ<sup>3</sup>

<sup>1</sup> Univerzitet odbrane u Beogradu – Institut za strategijska istraživanja, [dragan.bojanic@mod.gov.rs](mailto:dragan.bojanic@mod.gov.rs)

<sup>2</sup> Elektrotehnička škola „Nikola Tesla“ u Pančevu, [marina.bojanic@etsntesla.edu.rs](mailto:marina.bojanic@etsntesla.edu.rs)

<sup>3</sup> Univerzitet odbrane u Beogradu – Institut za strategijska istraživanja, [vladimir.ristic@mod.gov.rs](mailto:vladimir.ristic@mod.gov.rs)

**Rezime:** U radu je prikazan matematički model za tretiranje neizvesnosti u višekriterijumskom odlučivanju, koji je zasnovan na intervalnim grubim brojevima (IGB). Ovakav pristup omogućava donošenje odluka uz korišćenje isključivo internih znanja u podacima i operativnih podataka donosioca odluke. Određivanje težinskih koeficijenata kriterijuma (definisanje kriterijuma), jedan je od ključnih problema koji se javlja u modelima višekriterijumske optimizacije. Prikazani matematički model je u funkciji podrške odlučivanju pri identifikaciji činioca ugrožavanja bezbednosti u izradi Strategije nacionalne bezbednosti Republike Srbije.

**Ključne reči:** višekriterijumsko odlučivanje, izazovi, rizici, pretnje, ugrožavanje bezbednosti,

**Abstract:** The paper presents a mathematical model for treating uncertainty in multi-criteria decision-making, which is based on interval rough numbers (IGB). This approach enables decision-making with the use of exclusively internal knowledge in the data and operational data of the decision-maker. Determining the weight coefficients of the criteria (defining the criteria) is one of the key problems arising in multicriteria optimization models. The presented mathematical model is in the function of decision support in identifying the factors endangering security in developing the National Security Strategy of the Republic of Serbia.

**Keywords:** multi-criteria decision-making, challenges, risks, threats, endangering security,

## 1. UVOD

Utvrđivanje kriterijuma za donošenje odluka je težak i odgovoran zadatak. Danas se skoro uvek bavimo višekriterijumskom optimizacijom, odnosno donošenjem odluka u pogledu više od jednog kriterijuma. Razvijene su različite matematičke metode za rešavanje ovih problema. Pored fuzzy teorije, veoma pogodan alat za tretiranje neizvesnosti bez uticaja subjektivizma je teorija grubih skupova. Znajući prednosti teorije grubih skupova [7] i [8], danas je u modernoj praksi potpuno opravdano sprovođenje procesa odlučivanja uz primenu grubih skupova, kada su u njemu uključeni neodređeni i nedostupni podaci. Za razliku od teorije fuzzy skupova čija primena zahteva definisanje parcijalne funkcije pripadnosti bez jasnih granica skupa, u teoriji grubih skupova se koristi granična oblast skupa za izražavanje nejasnoća. Što se tiče primene intervalnih grubih brojeva u višekriterijumskom odlučivanju, autori su u literaturi pronašli svega nekoliko radova [11] koji su primenjivali intervalne grube brojeve za eksploataciju neizvesnosti prilikom određivanja težinskih koeficijenata atributa odluke. IGB su primenjivani i za izradu hibridnog IVFRN-MAIRCA modela [6] pristup koji kombinuje prednosti fuzzy i grubih koncepata.

Doprinos ovog rada je DEMATEL-ANP model koji je zasnovan na intervalnim ili klasičnim grubim brojevima (IGD'ANP), a koji obezbeđuje objektivniju ekspertsku evaluaciju kriterijuma u subjektivnom okruženju. Drugi značajan doprinos ovog rada je unapređenju MCDM tehnika. Treći doprinos rada je unapređenje metodologije evaluacije kriterijuma pri identifikaciji izazova, rizika i pretnji ili činioca ugrožavanja bezbednosti, kroz novi pristup u tretiranju neizvesnosti. Ovakav pristup Holandska vlada ima od 2007. godine, primenu MCDM modela, grupnog odlučivanja i ekspertsku analizu kriterijuma u izradi strategije nacionalne bezbednosti, kao i druge zemlje uključujući Nemačku, Englesku, Finsku i Kanadu [1].

## 2. METODOLOGIJA

U prvoj fazi, kroz primenu IG DEMATEL modela vrši se ekspertna evaluacija kriterijuma i formiraju se ulazni podaci za IG ANP model. Izlazni podaci iz IG DEMATEL-a dalje se obrađuju kroz algoritam IG ANP modela. Kao izlaz iz IG D'ANP modela dobijaju se intervalni grubi težinski koeficijenti kriterijuma.

U cilju sveobuhvatnijeg uvažavanja subjektivnosti koja postoji u grupnom donošenju odluka, u ovom radu je izvršena modifikacija DEMATEL metode (eng. Decision-Making Trial and Evaluation Laboratory) primenom intervalnih grubih brojeva. Primenom intervalnih grubih brojeva eliminiše se potreba za dodatnim informacijama kojima se određuju neizvesni intervali brojeva. Time se zadržava kvalitet postojećih podataka u grupnom odlučivanju i percepcije eksperata se na objektivni način iskazuju u agregiranoj matrici. Ovaj metod omogućava bolje razumevanje kompleksne strukture razmatranog problema i određuje veze između faktora, veze između nivoa strukture i jačine uticaja faktora [4]. *Korak 1: Ekspertna analiza faktora. Korak 2: Određivanje matrice prosečnih odgovora eksperata. Korak 3: Na osnovu matrice  $Z$  izračunava se normalizovana matrica prosečne percepcije, Korak 4: Izračunava se matrica ukupnih uticaja  $T$ , Korak 5: Izračunavanje sume redova i kolona matrice ukupnih uticaja  $T$  i Korak 6: Određivanje granične vrednosti.*

ANP metoda predstavlja generalizaciju AHP metode i za razliku od hijerarhijski strukturiranih problema uzima u obzir razne forme zavisnosti i povratnu spregu. Struktura povratne sprege nije linearnog karaktera i mnogo više podseća na mrežu na kojoj se učestalo pojavljuju petlje koje međusobno povezuju grupe. Matrice koje u slučaju postojanja mreže opisuju ove zavisnosti nazivaju se supermatrice i moraju posedovati osobinu da su kolonski stohastičke, što znači da je suma elemenata svake kolone jednaka jedinici [9]. Primenom tradicionalne ANP metode za proračun relativnih težina klastera/kriterijuma, nivoi međuzavisnosti faktora se tretiraju kao recipročne vrednosti. *Korak 1. Određivanje neotežane supermatrice, Korak 2. Normalizacija matrice ukupnog uticaja za kriterijume  $T_c^a$ , Korak 3. Proračun elemenata neotežane supermatrice  $W$ , Korak 4. Proračun elemenata otežane normalizovane supermatrice  $W^a$ , Korak 5. Pronalaženje granice otežane (ponderisane) supermatrice  $W^a$ .*

## 3. PRIMENA IGD'ANP MODELA: ODREĐIVANJE TEŽINSKIH KOEFICIJENATA KRITERIJUMA

Strategija nacionalne bezbednosti Republike Srbije (u daljem tekstu: SNB) jeste najviši strateški dokument čijim sprovođenjem se štite nacionalne vrednosti i interesi Republike Srbije od izazova, rizika i pretnji bezbednosti u svim oblastima društvenog života [10]. Nacionalna bezbednost je ugrožena ako su vitalni interesi države i/ili društva ugroženi do te mere da postoji – potencijalni – društveni poremećaj. Kako bi se država optimalno pripremila za suočavanje sa različitim bezbednosnim izazovima, rizicima i pretnjama, donosiocima odluka je potreban integrisan, koherentan i sistematičan pristup, koji će im omogućiti da klasifikuju i uporede različite činioce ugrožavanja bezbednosti, da odrede prioritete i donesu adekvatnu i efikasnu odluku u jačanju odgovarajuće pripreme i izgradnje potrebnih kapaciteta. Određivanje težinskih koeficijenata kriterijuma u postupcima identifikacije činioca ugrožavanja bezbednosti je višekriterijumski problem, pa je primena modela višekriterijumske optimizacije pogodan alat za donošenje odluka.

Tabela 1. Kriterijumi za identifikovanje činioca ugrožavanja bezbednosti

Vitalni interesi RS	sektori/kriterijumi
Ekonomski razvoj	<b>Ekonomski (D1)</b>
	Narušavanje energetske stabilnosti (C1)
	Narušavanje privrednog razvoja (C2)
	Troškovi (C3)
Očuvanje suverenosti teritorijalne nezavisnosti	<b>Vojni (D2)</b>
	Ugrožavanje teritorijalnog integriteta (C4)
	Ugrožavanje informacionog prostora (C5)
Očuvanje srpskog naroda	<b>Socijetalni (D3)</b>
	Narušavanje kolektivnog identiteta (C6)
	Narušavanje kulturnih vrednosti (C7)
Očuvanje životne sredine	<b>Ekološki (D4)</b>
	Ugrožavanje prirodnih resursa (C8)
	Narušavanje životne sredine (C9)
Očuvanje unutrašnje stabilnosti	<b>Politički (D5)</b>
	Narušavanje političke stabilnosti (C10)
	Narušavanje fizičke bezbednosti (C11)
	Narušavanje socijalne stabilnosti (C12)

U stvarnosti se u većini bezbednosnih problema prepliće dinamika iz više bezbednosnih sektora (politički, vojni, ekonomski, socijalni i ekološki). Prema Bojaniću, ovaj analitički okvir sektorskog pristupa kopenhavske škole može se uspešno primeniti u analizi i identifikovanju činioca ugrožavanja nacionalne bezbednosti [1]. Na osnovu vitalnih interesa iskazanih u SNB, sektorskog pristupa i analize holandske procene rizika u strategiji nacionalne bezbednosti [5], kriterijume možemo razvrstati u pet sektora (klastera) (Tabela 1).

**I faza: IG-DEMATEL model**

U prvom koraku IG-DEMATEL modela vrši se ekspertska analiza faktora. U ovom istraživanju učestvovalo je osam eksperata koji imaju minimum desetogodišnje iskustvo u analizi i vrednovanju bezbednosnih izazova rizika i pretnji. Prilikom evaluacije klastera/kriterijuma su koristili skalu: 0 - nema uticaja; 1 - mali uticaj; 2 - srednji uticaj; 3 - veliki uticaj; 4 - veoma veliki uticaj. Prilikom evaluacije, eksperti su odvojeno vršili evaluaciju klastera i kriterijuma. Nakon ekspertske evaluacije dobijeno je osam matrica poređenja u parovima kriterijuma dimenzija 12x12 (Tabela 2) i osam matrica poređenja u parovima klastera dimenzija 5x5.

Tabela 2. Ekspertsko poređenje kriterijuma evaluacije

E1												
	C1 <sub>1</sub>	C1 <sub>2</sub>	C1 <sub>3</sub>	C2 <sub>4</sub>	C2 <sub>5</sub>	C3 <sub>6</sub>	C3 <sub>7</sub>	C4 <sub>8</sub>	C4 <sub>9</sub>	C5 <sub>10</sub>	C5 <sub>11</sub>	C5 <sub>12</sub>
C1 <sub>1</sub>	(0;0)	(3;4)	(3;3)	(2;3)	(2;3)	(3;5)	(3;3)	(2;3)	(2;5)	(1;1)	(4;4)	(4;4)
C1 <sub>2</sub>	(3;4)	(0;0)	(3;5)	(1;5)	(2;3)	(2;3)	(3;4)	(3;3)	(3;5)	(2;3)	(4;4)	(4;4)
C1 <sub>3</sub>	(3;4)	(3;5)	(0;0)	(4;5)	(3;4)	(3;4)	(2;3)	(3;3)	(3;5)	(2;5)	(4;5)	(4;4)
C2 <sub>4</sub>	(3;4)	(5;5)	(5;5)	(0;0)	(2;3)	(2;3)	(3;3)	(3;4)	(4;5)	(3;4)	(3;4)	(5;5)
.....												
C4 <sub>9</sub>	(3;3)	(2;3)	(4;5)	(1;5)	(2;2)	(3;4)	(4;5)	(2;3)	(0;0)	(2;2)	(3;5)	(3;3)
C5 <sub>10</sub>	(4;4)	(2;3)	(2;5)	(5;5)	(1;1)	(2;3)	(4;5)	(2;3)	(2;3)	(0;0)	(3;4)	(3;4)
C5 <sub>11</sub>	(2;2)	(1;5)	(2;3)	(2;5)	(1;1)	(3;4)	(3;3)	(3;5)	(2;5)	(1;2)	(0;0)	(5;5)
C5 <sub>12</sub>	(3;3)	(2;5)	(1;3)	(2;3)	(4;5)	(4;4)	(3;3)	(3;3)	(3;5)	(4;5)	(4;4)	(0;0)
... E8												
	C1 <sub>1</sub>	C1 <sub>2</sub>	C1 <sub>3</sub>	C2 <sub>4</sub>	C2 <sub>5</sub>	C3 <sub>6</sub>	C3 <sub>7</sub>	C4 <sub>8</sub>	C4 <sub>9</sub>	C5 <sub>10</sub>	C5 <sub>11</sub>	C5 <sub>12</sub>
C1 <sub>1</sub>	(0;0)	(2;5)	(2;5)	(1;5)	(1;5)	(2;4)	(2;5)	(1;5)	(1;4)	(2;2)	(3;4)	(3;3)
C1 <sub>2</sub>	(2;4)	(0;0)	(3;4)	(1;4)	(1;5)	(2;5)	(4;4)	(2;4)	(2;4)	(2;2)	(3;5)	(3;5)
C1 <sub>3</sub>	(1;4)	(3;3)	(0;0)	(3;4)	(3;5)	(3;5)	(2;4)	(2;4)	(2;3)	(3;4)	(4;4)	(3;5)
C2 <sub>4</sub>	(2;4)	(4;4)	(4;4)	(0;0)	(2;5)	(2;4)	(3;4)	(4;4)	(4;4)	(2;5)	(2;5)	(4;5)
.....												
C4 <sub>9</sub>	(1;3)	(1;2)	(4;4)	(1;4)	(2;4)	(2;5)	(3;4)	(1;4)	(0;0)	(3;3)	(2;3)	(4;5)
C5 <sub>10</sub>	(3;5)	(1;2)	(1;4)	(4;4)	(1;3)	(1;4)	(4;4)	(1;5)	(1;4)	(0;0)	(4;4)	(4;4)
C5 <sub>11</sub>	(1;2)	(1;3)	(1;4)	(1;4)	(1;3)	(4;4)	(2;4)	(4;5)	(1;4)	(2;2)	(0;0)	(3;3)
C5 <sub>12</sub>	(1;2)	(2;4)	(1;4)	(1;4)	(4;4)	(4;5)	(2;5)	(5;5)	(4;5)	(3;4)	(4;5)	(0;0)

Na isti način dobijaju se intervalni grubi brojevi ostalih matrica poređenja u parovima kriterijuma (Tabela 3) i klastera.

Tabela 3. Intervalne grube matrice poređenja u parovima kriterijuma

E1						
	C1 <sub>1</sub>	C1 <sub>2</sub>	C1 <sub>3</sub>	C2 <sub>4</sub>	C2 <sub>5</sub>	..... C5 <sub>12</sub>
C1 <sub>1</sub>	[(0,0,0),(0,0,0)]	[(2,5,3),(3,67,4,33)]	[(2,5,3),(3,3,75)]		[(1,5,2),(4,25,5)]	[(3,4),(3,67,4,33)]
C1 <sub>2</sub>	[(3,4),(4,3,75)]	[(0,0,0),(0,0,0)]	[(3,3,5),(3,75,5)]		[(1,25,2),(3,4)]	[(3,75,4),(3,5,4,67)]
C1 <sub>3</sub>	[(3,5),(4,25,5)]	[(3,3,5),(3,5)]	[(0,0,0),(0,0,0)]		[(3,3,25),(3,4,25)]	[(3,25,4),(3,67,4,33)]
C2 <sub>4</sub>	[(3,4),(3,3,5)]	[(4,75,5),(4,5)]	[(4,75,5),(4,25,5)]		[(2,2,25),(3,5,4,67)]	[(4,75,5),(4,75,5)]
.....				....		...
C4 <sub>9</sub>	[(2,75,4),(2,75,3)]	[(1,25,2),(2,3)]	[(4,4,5),(4,5,5)]		[(1,5,2),(2,5,4)]	[(3,4),(3,4,25)]
C5 <sub>10</sub>	[(3,67,4,33),(4,5,5)]	[(1,5,2),(2,25,3)]	[(1,25,2),(3,5,5)]		[(1,1,25),(2,67,4,5)]	[(3,3,25),(3,5,4)]
C5 <sub>11</sub>	[(2,3),(1,1,5)]	[(1,1,25),(2,75,5)]	[(1,5,2),(2,33,3,5)]		[(1,1),(2,33,3,33)]	[(3,25,5),(3,5,5)]
C5 <sub>12</sub>	[(2,25,3),(1,67,2,33)]	[(2,2,5),(4,25,5)]	[(1,1,25),(3,3,25)]		[(3,75,4),(4,4,5)]	[(0,0,0),(0,0,0)]
... E8						
	C1 <sub>1</sub>	C1 <sub>2</sub>	C1 <sub>3</sub>	C2 <sub>4</sub>	C2 <sub>5</sub>	... C5 <sub>12</sub>
C1 <sub>1</sub>	[(0,0,0),(0,0,0)]	[(2,2,5),(4,5)]	[(2,2,5),(3,75,5)]		[(1,1,5),(3,5,4,67)]	[(2,67,3,33),(3,4)]
C1 <sub>2</sub>	[(2,67,3,3),(3,3,75)]	[(0,0,0),(0,0,0)]	[(3,3,5),(3,33,4,33)]		[(1,1,25),(4,5)]	[(3,3,75),(4,25,5)]
C1 <sub>3</sub>	[(2,3,3,67),(4,4,25)]	[(3,3,5),(2,33,3,67)]	[(0,0,0),(0,0,0)]		[(3,3,25),(4,25,5)]	[(3,3,25),(4,5)]

C2 <sub>4</sub>	[(2.67,3.33),(3,3.5)]	[(4,4.75),(3.67,4.33)]	[(4,4.75),(3.5,4.67)]	.....	[(2,2.25),(4.25,5)]	[(4,4.75),(4.75,5)]
C4 <sub>9</sub>	[(2.33,3.3),(2,2.75)]	[(1,1.25),(1.67,2.33)]	[(4,4.5),(4,4.5)]	.....	[(1.5,2),(3.5,5)]	[(3.67,4.33),(4.25,5)]
C5 <sub>10</sub>	[(4.0,5.0),(4.0,4.5)]	[(1,1.5),(1.5,2.67)]	[(1,1.25),(3,4.5)]		[(1,1.25),(3.25,5)]	[(3.25,4),(3.5,4)]
C5 <sub>11</sub>	[(1.67,2.33),(1,1.5)]	[(1,1.25),(2,4)]	[(1,1.5),(2.75,4)]		[(1,1),(2.75,4)]	[(2.67,3.67),(3,3.5)]
C5 <sub>12</sub>	[(1.5,2.67),(1.0,2.0)]	[(2,2.5),(3.5,4.67)]	[(1,1.25),(3.25,4)]		[(3.75,4),(4.5,5)]	[(0.0,0.0),(0.0,0.0)]

U drugom koraku IG-DEMATEL modela vrši se agregacija intervalnih grubih matrica odgovora eksperata. Na osnovu matrica odgovora klastera i kriterijuma (Tabela 3), primenom matematičkih izraza dobijamo osrednjeni intervalni grubi broj  $IRN(x_{ij}) = [RN(x_{ij}^L), RN(x_{ij}^U)]$ . Tako dobijamo osrednjene intervalne grube matrice prosečnih odgovora eksperata za klastera i kriterijume, (Tabela 4).

Tabela 4. Osrednjena intervalna gruba matrica kriterijuma

	C1 <sub>1</sub>	C1 <sub>2</sub>	C1 <sub>3</sub>	C2 <sub>4</sub>	...	C5 <sub>12</sub>
C1 <sub>1</sub>	[(0.00,0.00),(0.00,0.00)]	[(2.25,2.75),(3.83,4.67)]		[(1.25,1.75),(3.38,4.38)]		[(2.83,3.67),(3.33,4.17)]
C1 <sub>2</sub>	[(2.83,3.67),(3.50,3.75)]	[(0.00,0.00),(0.00,0.00)]		[(1.00,1.25),(3.54,4.67)]		[(3.38,3.88),(3.88,4.83)]
C1 <sub>3</sub>	[(2.67,4.33),(4.13,4.63)]	[(3.00,3.50),(2.67,4.33)]		[(3.13,3.63),(3.88,4.83)]		[(3.13,3.63),(3.83,4.67)]
C2 <sub>4</sub>	[(2.83,3.67),(3.00,3.50)]	[(4.38,4.88),(3.83,4.67)]	....	[(0.00,0.00),(0.00,0.00)]		[(4.38,4.88),(4.75,5.00)]
C4 <sub>9</sub>	[(2.54,3.67),(2.38,2.88)]	[(1.13,1.63),(1.83,2.67)]		[(1.00,1.50),(3.25,4.75)]	...	[(3.33,4.17),(3.63,4.63)]
C5 <sub>10</sub>	[(3.83,4.67),(4.25,4.75)]	[(1.25,1.75),(1.88,2.83)]		[(4.13,4.63),(4.25,4.75)]		[(3.13,3.63),(3.50,4.00)]
C5 <sub>11</sub>	[(1.83,2.67),(1.00,1.50)]	[(1.00,1.25),(2.38,4.50)]		[(1.13,1.63),(3.25,4.75)]		[(2.96,4.33),(3.25,4.25)]
C5 <sub>12</sub>	[(1.88,2.83),(1.33,2.17)]	[(2.00,2.50),(3.88,4.83)]		[(1.25,1.75),(2.54,3.75)]		[(0.00,0.00),(0.00,0.00)]

Osrednjavanje elemenata intervalnih grubih matrica poređenja u parovima kriterijuma na poziciji C1<sub>1</sub> - C1<sub>4</sub> vrši se primenom izraza (1) i (2):

$$RN(z_{1-2_4}^L) = RN(x_{1-2_4}^{1L}, x_{1-2_4}^{2L}, \dots, x_{1-2_4}^{8L}) = \begin{cases} z_{1-2_4}^L = \frac{1}{8} \sum_{e=1}^8 x_{1-2_4}^{eL} = \frac{1}{8} (1.5 + 1 + \dots + 1) = 1.25 \\ z_{1-2_4}^U = \frac{1}{8} \sum_{e=1}^8 x_{1-2_4}^{eU} = \frac{1}{8} (2 + 2 + \dots + 1.5) = 1.75 \end{cases} \quad (1)$$

$$RN(z_{1-2_4}^U) = RN(x_{1-2_4}^{1U}, x_{1-2_4}^{2U}, \dots, x_{1-2_4}^{8U}) = \begin{cases} z_{1-2_4}^L = \frac{1}{8} \sum_{e=1}^8 x_{1-2_4}^{eL} = \frac{1}{8} (3 + 2.5 + \dots + 3.75) = 3.38 \\ z_{1-2_4}^U = \frac{1}{8} \sum_{e=1}^8 x_{1-2_4}^{eU} = \frac{1}{8} (3.75 + 4.5 + \dots + 5) = 4.38 \end{cases} \quad (2)$$

gde  $e$  predstavlja  $e$ -tog eksperta ( $e = 1, 2, \dots, 8$ ),  $RN(z_{1-2_4}^L)$  i  $RN(z_{1-2_4}^U)$  predstavljaju donju i gornju granicu intervalnog grubog broja, respektivno. Predstavljene grube sekvence  $RN(z_{1-2_4}^L)$  i  $RN(z_{1-2_4}^U)$  predstavljaju osrednjeni intervalni grubi broj  $IRN(z_{1-2_4}) = [(1.25, 1.75), (3.38, 4.38)]$ .

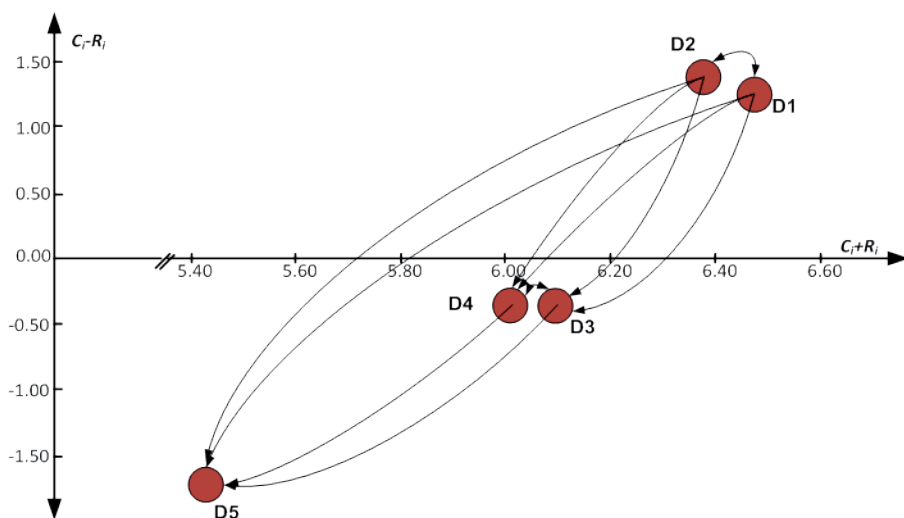
Nakon određivanja osrednjene matrice kriterijuma (Tabela 4) i klastera sprovodi se drugi korak IG-DEMATEL modela koji podrazumeva određivanje normalizovane matrice prosečne percepcije. Primenom matematičkih izraza proračunavaju se elementi IG početna matrica direktnog odnosa klastera i kriterijuma.

Početna matrica direktnog odnosa klastera/kriterijuma primenom matematičkih izraza transformišu se u matrice ukupnih uticaja klastera/kriterijuma (Tabela 5).

Tabela 5. Matrica ukupnih uticaja kriterijuma

	C1 <sub>1</sub>	C1 <sub>2</sub>	....	C2 <sub>4</sub>	..	C5 <sub>12</sub>
C1 <sub>1</sub>	[(0.22,0.35),(0.56,0.78)]	[(0.29,0.36),(0.62,0.94)]		[(0.17,0.24),(0.72,1.02)]		[(0.34,0.46),(0.76,1.01)]
C1 <sub>2</sub>	[(0.34,0.49),(0.65,0.86)]	[(0.27,0.34),(0.55,0.87)]		[(0.19,0.26),(0.74,1.05)]		[(0.41,0.53),(0.78,1.04)]
C1 <sub>3</sub>	[(0.39,0.57),(0.68,0.88)]	[(0.41,0.47),(0.62,0.95)]		[(0.28,0.35),(0.77,1.05)]		[(0.47,0.58),(0.8,1.04)]
C2 <sub>4</sub>	[(0.43,0.6),(0.66,0.85)]	[(0.48,0.55),(0.65,0.94)]	....	[(0.22,0.29),(0.69,0.95)]		[(0.55,0.67),(0.83,1.03)]
C4 <sub>9</sub>	[(0.32,0.49),(0.56,0.79)]	[(0.29,0.38),(0.52,0.85)]		[(0.19,0.27),(0.65,0.97)]		[(0.39,0.54),(0.68,0.97)]
C5 <sub>10</sub>	[(0.36,0.51),(0.62,0.83)]	[(0.3,0.38),(0.55,0.86)]		[(0.27,0.34),(0.71,0.98)]		[(0.4,0.52),(0.72,0.97)]
C5 <sub>11</sub>	[(0.26,0.42),(0.51,0.76)]	[(0.25,0.33),(0.53,0.88)]		[(0.16,0.24),(0.64,0.97)]		[(0.34,0.48),(0.67,0.96)]
C5 <sub>12</sub>	[(0.35,0.52),(0.61,0.79)]	[(0.37,0.44),(0.63,0.9)]		[(0.22,0.3),(0.73,0.98)]		[(0.37,0.49),(0.7,0.9)]

Sumiranjem elemenata matrice ukupnih uticaja klastera/kriterijuma po redovima i po kolonama, određuju se vrednosti ukupnih direktnih i indirektnih efekata koje je kriterijum  $j$  dobio od ostalih kriterijuma i predao ostalim kriterijumima. Te vrednosti se, zajedno sa graničnom vrednošću ( $\alpha$ ), koriste za definisanje dijagrama uzročno posledičnih odnosa (engl. cause and effect relationship: CER). CER dijagram (Slika 1) se izrađuje da bi se vizuelno predstavili složeni odnosi i pružile informacije radi donošenja zaključka koji su faktori najvažniji i kako utiču jedni na druge.



Slika 1. CER dijagram

Vrednosti elemenata matrice  $T$  (Tabela 5), koji imaju višu vrednost u odnosu na graničnu vrednost  $\alpha$  izdvajaju se i unose u dijagram (Slika 1) na kome  $x$  - osa predstavlja  $IGB(R_i+C_i)$ , a  $y$  - osa  $IGB(R_i-C_i)$ . Te vrednosti služe za predstavljanje veze između dva faktora. Prilikom predstavljanja veza između faktora strelica uzajamno-posledične veze se usmerava od faktora koji ima nižu vrednost u odnosu na  $\alpha$  ka elementu koji ima višu vrednost u odnosu na  $\alpha$ .

## II faza: IG-ANP model

U radu je korišćena IG-ANP metoda za određivanje težina 12 kriterijuma i pet klastera na osnovu matrice ukupnih uticaja klastera/kriterijuma ( $T$ ) koja je dobijena primenom IG-DEMATEL modela (I faza). Na osnovu CER dijagrama, u II fazi modela, proračunavaju se težinski koeficijenti klastera/kriterijuma. U prvom koraku IG-ANP modela na osnovu matrice ukupnih uticaja klastera/kriterijuma proračunavaju se elementi intervalne grube neponderisane i ponderisane supermatrice.

Tabela 6. Težinske vrednosti klastera/kriterijuma

Klasteri/kriterijumi	Težinski koeficijent	Rang
<b>Ekonomski (D1)</b>	[(0.039,0.918),(0.047,1.017)]	<b>1</b>
Ugrožavanje energetske stabilnosti (C1 <sub>1</sub> )	[(0.015,0.289),(0.017,0.331)]	6
Narušavanje privrednog razvoja (C1 <sub>2</sub> )	[(0.013,0.255),(0.018,0.298)]	10
Troškovi C1 <sub>3</sub> )	[(0.011,0.374),(0.012,0.388)]	1
<b>Vojni (D2)</b>	[(0.104,0.585),(0.023,0.623)]	<b>3</b>
Ugrožavanje teritorijalnog integriteta (C2 <sub>4</sub> )	[(0.013,0.287),(0.014,0.308)]	8
Ugrožavanje informacionog prostora (C2 <sub>5</sub> )	[(0.090,0.298),(0.010,0.315)]	3
<b>Socijetalni (D3)</b>	[(0.018,0.502),(0.020,0.578)]	<b>5</b>
Narušavanje kolektivnog identiteta (C3 <sub>6</sub> )	[(0.008,0.285),(0.009,0.311)]	9
Narušavanje kulturnih vrednosti (C3 <sub>7</sub> )	[(0.010,0.217),(0.011,0.267)]	12
<b>Ekološki (D4)</b>	[(0.018,0.597),(0.017,0.655)]	<b>4</b>
Ugrožavanje prirodnih resursa (C4 <sub>8</sub> )	[(0.09,0.288),(0.011,0.324)]	7
Narušavanje životne sredine (C4 <sub>9</sub> )	[(0.007,0.309),(0.008,0.331)]	5
<b>Politički (D5)</b>	[(0.021,0.935),(0.025,1.031)]	<b>2</b>
Narušavanje političke stabilnosti (C5 <sub>10</sub> )	[(0.008,0.364),(0.009,0.379)]	2
Narušavanje fizičke bezbednosti (C5 <sub>11</sub> )	[(0.003,0.314),(0.005,0.373)]	4
Narušavanje socijalne stabilnosti (C5 <sub>12</sub> )	[(0.010,0.257),(0.011,0.279)]	11

U tabeli 6 pored težinskih koeficijenata prikazana je i prioritizacija klastera/kriterijuma. Uočavamo da su klasteri D1 i D5 najuticajni i da se u njima nalaze kriterijumi sa najvećim težinskim koeficijentima, Troškovi ( $C1_3$ ) i Politička stabilnost ( $C5_{10}$ ). Zatim po uticaju slede klasteri D4 i D3. Prikazanom prioritizacijom klastera/kriterijuma potvrđene su preporuke o značajnosti kriterijuma [2], [3].

## 5. ZAKLJUČAK

Uvažavanje neizvesnosti u višekriterijumskom odlučivanju je veoma značajan aspekt za objektivno i nepristrasno donošenje odluka. Osnovna ideja primene algoritama za donošenje odluka koji su bazirani na intervalnom pristupu podrazumeva primenu intervalnih brojeva za prezentovanje vrednosti atributa odluke. Primena IGB u višekriterijumskom donošenju odluka prikazana je kroz kombinovani model koji se sastoji od IGD'ANP modela. Primena IG-D'ANP modela prikazana je kroz određivanje težinskih koeficijenata kriterijuma za potrebe identifikovanja činioca ugrožavanja bezbednosti pri izradi Strategije nacionalne bezbednosti. Ovo istraživanje pokazuje da se IGB mogu efikasno primenjivati u modelima višekriterijumskog odlučivanja uz uvažavanje nedoumica u procesu donošenja odluka. Time se eliminišu subjektivnosti i pretpostavke koje u značajnoj meri mogu da utiču na vrednosti atributa i konačano na rangiranje alternative

Drugi važan segment ovog rada je primena IG-DEMATEL i IG-ANP modela koji su prikazani od strane autora, a predstavljaju doprinos literaturi višekriterijumskog donošenja odluka i primeni matematičkog modela u oblasti društvenih nauka. Predloženi modeli omogućavaju evaluaciju alternativa uprkos neizvesnosti u procesu donošenja odluka i nedostatku kvantitativnih informacija. Dalja integracija IG pristupa u postojeće modele višekriterijumskog odlučivanja u značajnoj meri omogućila bi eksploataciju neodređenosti i subjektivnosti koje postoje u procesu razdvajanja činioca ugrožavanja bezbednosti na izazove, rizike i pretnje u SNB.

## LITERATURA

- [1] Bojanić D., (2022), The theoretical and methodological analysis of challenges, risks and threats in modern theory of national security, *Vojno delo*, Institut za strategijska istraživanja 2/2022, DOI broj: 10.5937/vojdelo2202001B
- [2] Chai J., Lui N.K. J., Ngai W.T. E., (2013), Application of decision-making techniques in supplier selection: A systematic review of literature, *Expert Systems with Applications*, Volume 40, Issue 10, pp 3872-3885, <https://doi.org/10.1016/j.eswa.2012.12.040>
- [3] Dobi, J. K., Gugić, D., Kancijan, (2010) AHP As a Decision Support Tool in the Multi criteria Evaluation of Bids in Public Procurement, *Proceedings of the ITI 2010 Int. Conf. on Information Technology Interfaces*, 32, 447- 452.
- [4] Gigović, LJ., Pamučar, D., Božanić, D., Ljubojević, S. (2017). "Application of the GIS-DANP-MABAC multi-criteria model for selecting the location of wind farms: A case study of Vojvodina, Serbia". *Renewable Energy*, 103, pp 501-521.
- [5] Mennen M.G. & van Tuyll M.C. Dealing with future risks in the Netherlands: the National Security Strategy and the National Risk Assessment, *Journal of Risk Research*, Vol. 18, No. 7, 2015 860–876, <http://dx.doi.org/10.1080/13669877.2014.923028>
- [6] Pamučar D., Ćirović G., Božanić D., (2019), Application of interval valued fuzzy-rough numbers in multi-criteria decision making: the IVFRN-MAIRCA model, *Yugoslav Journal of Operations Research*, 29, Number 2, 221-247 DOI: <https://doi.org/10.2298/YJOR180415011P>.
- [7] Pawlak, Z. (1982). Rough sets. *International Journal of Computer & Information Sciences*, 11(5), 341–356.
- [8] Pawlak, Z. (1991). *Rough sets: Theoretical aspects of reasoning about data* (Vol. 9). Berlin: Springer.
- [9] Saaty, T. L., Vargas, L. G. (2012). *Models, methods, concepts and applications of the analytic hierarchy process*, Vol. 175. Springer Science and Business Media
- [10] Strategija nacionalne bezbednosti Republike Srbije „Službeni glasnik RS”, broj 94/2019.
- [11] Wang, J., Wang, J.Q, Zhang, H.Y, Chen, X.H. (2015). Multi-criteria decision-making based on hesitant fuzzy linguistic term sets: An outranking approach. *Knowledge-Based Systems*, 86, pp. 224–236.



## MODEL RECENZIJE NAUČNOISTRAŽIVAČKIH PROJEKATA U MINISTARSTVU ODBRANE I VOJSCI SRBIJE

### MODEL OF REVIEWING SCIENTIFIC RESEARCH PROJECTS IN THE MINISTRY OF DEFENSE AND THE SERBIAN ARMY

SRĐAN DIMIĆ<sup>1</sup>, SRĐAN LJUBOJEVIĆ<sup>2</sup>, DRAGAN KOSTADINOVIĆ<sup>3</sup>

<sup>1</sup> Univerzitet odbrane u Beogradu – Vojna akademija, srdjan.dimic@mod.gov.rs

<sup>2</sup> Univerzitet odbrane u Beogradu – Vojna akademija, srdjan.ljubojevic@va.mod.gov.rs

<sup>3</sup> Univerzitet odbrane u Beogradu – Vojna akademija, kotoza@beotel.net

**Rezime:** *Ministarstvo odbrane i Vojska Srbije radi stručne verifikacije naučnoistraživačkih projekata i unapređenja naučnoistraživačke delatnosti, svake godine, organizuju recenziju pojedinih naučnoistraživačkih projekata. S obzirom na to da se u Ministarstvu odbrane i Vojsci Srbije realizuje veći broj naučnoistraživačkih projekata, iz različitih naučnih oblasti, pored izbora projekata koji će biti recenzirani, problem predstavlja i način vršenja recenzija. U radu je prikazan modifikovani način vršenja recenzija naučnoistraživačkih projekata u Ministarstvu odbrane i Vojsci Srbije, kojim se čitav proces kvantifikuje i dobijaju ocene na osnovu kojih je moguće porediti recenzirane projekte.*

**Ključne reči:** *Model, recenzija, naučnoistraživački projekat.*

**Abstract:** *For the purpose of professional verification of scientific research projects and improvement of scientific research activities, the Ministry of Defense and the Serbian Army organize a review of certain scientific research projects every year. Considering that a large number of scientific research projects are being implemented in the Ministry of Defense and the Serbian Army, from various scientific fields, in addition to the selection of projects to be reviewed, the problem is how to conduct reviews. The paper presents a modified way of reviewing scientific research projects in the Ministry of Defense and the Serbian Army, which quantifies the entire process and obtains estimates based on which it is possible to compare peer-reviewed projects from various scientific fields.*

**Keywords:** *Model, Review, Scientific research projects.*

## 1. UVOD

Recenziranje je kamen temeljac u procesu objavljivanja naučnih radova, ali i često zanemaren faktor razvoja celokupne nauke. Zasnovano je na ideji da se rezultati svakog naučnog istraživanja moraju stručno oceniti od strane eksperata – recenzenata, pre nego se prezentuju naučnoj zajednici i javnosti [1].

Radi unapređenja naučnoistraživačke delatnosti u Ministarstvu odbrane i Vojsci Srbije, svake godine, vrši se recenzija pojedinih naučnoistraživačkih projekata iz Plana naučnoistraživačke delatnosti. Imajući u vidu da se u Ministarstvu odbrane i Vojsci Srbije realizuje veći broj naučnoistraživačkih projekata, iz različitih naučnih oblasti, pored izbora projekata koji će biti recenzirani, problem predstavlja i način vršenja recenzija.

U skladu sa popisima kojima se uređuje naučnoistraživačka delatnost u Ministarstvu odbrane i Vojsci Srbije, ministar odbrane formira Savet za naučnoistraživačku delatnost, kao savetodavno telo, koje između ostalih zadataka, razmatra i predlaže projekte za recenziju. Na osnovu predloga Saveta, ministar odbrane donosi rešenje kojim odobrava zaključenje ugovora sa recenzentima i recenziju naučnoistraživačkih projekata. Recenzente predlaže Stalno radno telo Saveta za naučnoistraživačku delatnost u Ministarstvu odbrane o recenziji i realizaciji naučnoistraživačkih projekata. Za recenzente naučnoistraživačkih projekata u Ministarstvu odbrane i Vojsci Srbije biraju se kompetentna lica iz naučnoistraživačkih organizacija i visokoškolskih ustanova u Republici Srbiji.

Prema postojećem načinu, recenzije naučnoistraživačkih projekata u Ministarstvu odbrane i Vojsci Srbije vrše se po devet elemenata, koji su utvrđeni Rešenjem o obrazovanju Stalnog radnog tela, dodeljivanjem opisnih ocena, tako da svaki element, a u krajnjem i čitav naučnoistraživački projekat mogu biti ocenjeni pozitivno ili negativno.

U cilju dobijanja ocene recenzije koja omogućava poređenje naučnoistraživačkih projekata, predložena je modifikacija postojećeg načina vršenja recenzija.

## 2. POSTOJEĆI NAČIN RECENZIJJE NAUČNOISTRAŽIVAČKIH PROJEKATA

Proces recenzije naučnoistraživačkih projekata započinje tako što Stalno radno telo razmatra projekte iz Plana naučnoistraživačke delatnosti, koji bi bili recenzirani. Plan naučnoistraživačke delatnosti, iz koga se biraju projekti za recenziju, obuhvata projekte iz osnovnih, primenjenih i razvojnih istraživanja, iz različitih naučnih oblasti (tehničko-tehnoloških, prirodno-matematičkih, medicinskih i društveno-humanističkih nauka). U skladu sa tim, Stalno radno telo, prilikom razmatranja projekata za recenziju, teži da obuhvati projekte iz svih naučnih oblasti.

U skladu sa [4], izbor naučnoistraživačkih projekata iz Plana naučnoistraživačke delatnosti koji se recenziraju, vrši se na osnovu jednog od sledećih kriterijuma:

- 1) da je naučnoistraživački projekat završen;
- 2) da je naučnoistraživački projekat bio u skladu sa prioritetima utvrđenim Planom naučnoistraživačke delatnosti;
- 3) da je planirana realizacija naučnoistraživačkog projekta tri i više godina.

Stalno radno telo, razmatrane naučnoistraživačke projekte dostavlja Savetu za naučnoistraživačku delatnost, koji usvaja konačan predlog naučnoistraživačkih projekata i predlaže ministru odbrane da ih odobri za recenziju. Na osnovu predloga Saveta, ministar odbrane donosi rešenje kojim odobrava recenziju naučnoistraživačkih projekata. Za recenziju svakog naučnoistraživačkog projekta određuju se najmanje po dva nezavisna recenzenta iz naučne oblasti iz koje su predloženi naučnoistraživački projekti.

Poslovi sprovođenja recenzije naučnoistraživačkih projekata povereni su Upravi za stratezijsko planiranje, organizacionoj jedinici koja je nosilac poslova naučnoistraživačke delatnosti u Ministarstvu odbrane i Vojsci Srbije. Nakon donošenja rešenja ministra odbrane o odobrenju zaključenja ugovora o recenziji naučnoistraživačkih projekata, Uprava za stratezijsko planiranje, organizuje potpisivanje ugovora sa recenzentima i dostavlja im sav potreban materijal za objektivnu recenziju naučnoistraživačkih projekata.

Nakon potpisivanja ugovora o recenziji naučnoistraživačkih projekata, recenzenti imaju rok od 30 dana da urade recenziju. U skladu sa [5], recenzenti ocenjuju projekat po devet elemenata, i to:

- E1 - naučna opravdanost istraživanja,
- E2 - naučna zasnovanost istraživanja,
- E3 - primenljivost rezultata projekta,
- E4 - nivo originalnosti projekta,
- E5 - cena projekta,
- E6 - kompetentnost rukovodioca projekta,
- E7 - kompetentnost istraživačkog tima,
- E8 - opremljenost realizatora projekta i
- E9 - rizik primene rezultata projekta.

Recenzenti svaki element ocenjuju pozitivno ili negativno, na osnovu čega na kraju daju zaključnu ocenu o recenziranom projektu koja je takođe opisna, pozitivna ili negativna. U slučaju da jedan recenzent da pozitivnu, a drugi negativnu ocenu, angažuje se i treći recenzent. Nakon toga donosi se konačna ocena o uspešnosti recenziranog naučnoistraživačkog projekta.

Radi unapređenja naučnoistraživačkog rada, izvršene recenzije naučnoistraživačkih projekata dostavljaju se nosiocima planiranja i realizacije zadataka naučnoistraživačke delatnosti, na koje se naučnoistraživački projekti odnose.

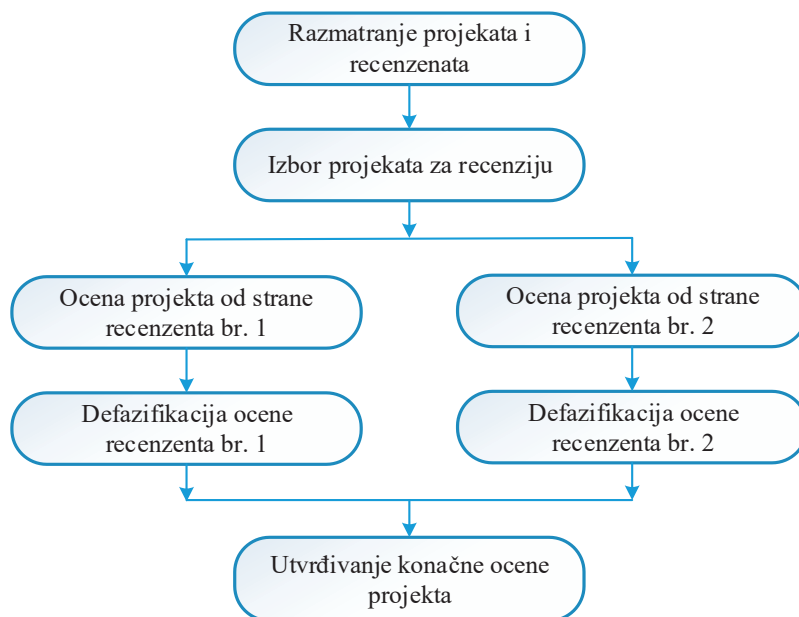
Nakon završenih svih recenzija, Uprava za stratezijsko planiranje, izrađuje Izveštaj o realizovanim recenzijama naučnoistraživačkih projekata, koji se prezentuje Savetu za naučnoistraživačku delatnost u Ministarstvu odbrane, koji taj izveštaj i usvaja. Na taj način završava se proces recenzija naučnoistraživačkih projekata u jednoj kalendarskoj godini.

## 3. PREDLOG MODIFIKOVANOG NAČINA RECENZIJJE NAUČNOISTRAŽIVAČKIH PROJEKATA

Imajući u vidu da recenzenti prilikom recenzije naučnoistraživačkih projekata daju pozitivnu ili negativnu ocenu, na osnovu zahteva Saveta za naučnoistraživačku delatnost u Ministarstvu odbrane, javila se potreba da se razmotri nov način vršenja recenzija, kako bi se projekti ocenjivali brožčanim ocenama, koje bi bile uporedive.

U skladu sa tim zahtevom, izrađen je predlog kojim se proces recenzije naučnoistraživačkih projekata kvantifikuje, na način da projekti dobijaju brojčanu ocenu, koju je moguće uporediti.

Proces modifikovanog načina recenzije naučnoistraživačkih projekata u Ministarstvu odbrane i Vojsci Srbije, prikazan je na slici 1.

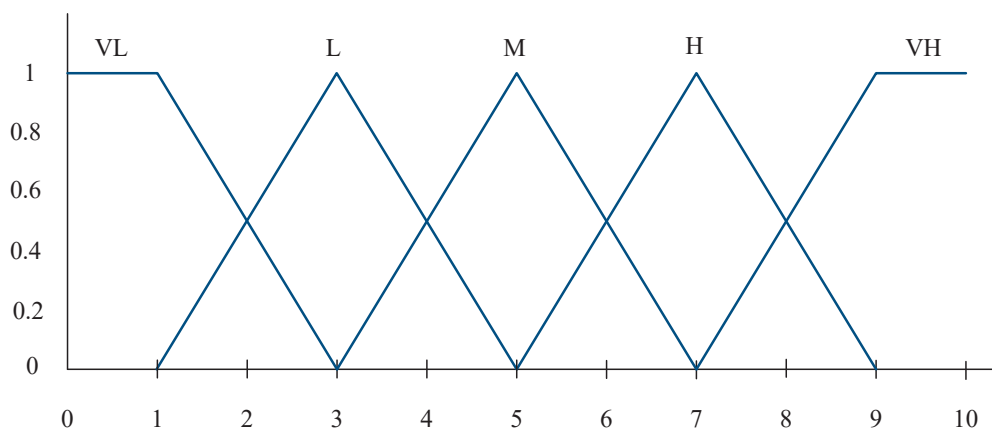


**Slika 1:** Proces modifikovanog načina recenzije naučnoistraživačkih projekata

Prvi korak u primeni prikazanog modela predstavlja razmatranje naučnoistraživačkih projekata od strane Stalnog radnog tela Saveta za naučnoistraživačku delatnost u Ministarstvu odbrane i izbor recenzenata. Nakon usaglašavanja predloga projekata koji bi bili recenzirani, na sednicu Stalnog radnog tela pozivaju se rukovodioci naučnoistraživačkih projekata da iste prezentuju, kako bi članovi Stalnog radnog tela u potpunosti sagledali razmatrane naučnoistraživačke projekte. Nakon diskusije o svakom projektu, Stalno radno telo formuliše predlog koji dostavlja Savetu na razmatranje i usvajanje. U skladu sa važećim odukama nadležnih tela, biraju se po dva recenzenta za svaki projekat, s tim da je, načelno, jedan iz akreditovanih naučnoistraživačkih organizacija Republike Srbije, a jedan iz Ministarstva odbrane i Vojske Srbije.

U drugom koraku naučnoistraživačke projekte, koje je izabralo Stalno radno telo, razmatra i usvaja Savet za naučnoistraživačku delatnosti u Ministarstvu odbrane. Savet izabrane projekte predlaže ministru odbrane koji ih u konačnom i odobrava za recenziju, donošenjem odgovarajućeg rešenja.

U trećem koraku svaki od recenzenata vrši nezavisnu recenziju dodeljenog naučnoistraživačkog projekta tako što ocenjuje svaki od devet napred navedenih elemenata. Imajući u vidu da je ocene elemenata složeno numerički iskazati, modelom je predviđeno da se one iskazuju sledećim skupom lingvističkih izraza: izuzetno niska (very low – VL), niska (low – L), srednja (medium – M), visoka (high – H) i izuzetno visoka (very high – VH). Skala za fazifikaciju lingvističkih izraza prikazana je na slici 2.



**Slika 2:** Skala za fazifikaciju lingvističkih izraza

Sve ocene se mogu predstaviti fuzzy brojem, čija funkcija pripadnosti u načelu može biti proizvoljnog oblika, ali se prema preporukama [2], iz praktičnih razloga, koriste trouglasti fuzzy brojevi  $A=(a,b,c)$ , gde je  $a$  – donja ili pesimistička vrednost,  $b$  – srednja ili najverovatnija vrednost i  $c$  – gornja ili optimistička vrednost ocene. Funkcija pripadnosti trouglastog fuzzy broja može se prikazati izrazom (1):

$$\mu_A(x) = \begin{cases} 0; & x \leq a \\ \frac{x-a}{b-a}; & a < x \leq b \\ \frac{c-x}{c-b}; & b \leq x < c \\ 0; & x \geq c \end{cases} \quad (1)$$

Srednja ocena recenzije naučnoistraživačkog projekta, od strane jednog recenzenta,  $A_{sr} = (a_{sr}, b_{sr}, c_{sr})$ , dobija se agregacijom pojedinačnih ocena u vidu srednje vrednosti dobijenog seta fuzzy brojeva (ocena) za svaki element, prema izrazu (2):

$$A_{sr} = (a_{sr}, b_{sr}, c_{sr}) = \left( \frac{1}{n} \cdot \sum_{i=1}^n a_i, \frac{1}{n} \cdot \sum_{i=1}^n b_i, \frac{1}{n} \cdot \sum_{i=1}^n c_i \right) \quad (2)$$

Kada se utvrdi srednja vrednost ocene recenziranog projekta, od strane jednog recenzenta, primenom izraza (3) koji je opisan u [3], vrši se defazifikacija fuzzy broja kako bi se dobila brojčana ocena.

$$A_{def} = \frac{(c-a) + (b-a)}{3} + a \quad (3)$$

Poslednji korak podrazumeva da se izračuna srednja vrednost ocena oba dva recenzenta, kako bi se dobila konačna ocena recenziranog projekta.

$$O = \frac{A_{def1} + A_{def2}}{2} \quad (4)$$

Ovako dobijena ocena recenzije omogućava lakše poređenje projekata, jer je ocena prikazana kao brojčana vrednost, a ne izrazom pozitivna, odnosno negativna.

#### 4. PRIMENA MODIFIKOVANOG MODELA

Primena modifikovanog načina recenzije naučnoistraživačkih projekata u Ministarstvu odbrane i Vojsci Srbije prikazana je na primeru recenzije projekata iz 2021. godine, s tim da su ocene recenzenta date hipotetički.

U skladu sa opisanim modelom, prvi korak predstavlja izbor naučnoistraživačkih projekata za recenziju. Za recenziju u 2021. godini bilo je izabrano pet naučnoistraživačkih projekata i za svaki projekat angažovana su po dva recenzenta.

Recenzent prilikom recenzije projekta, nakon opisnog dela, svakom od devet utvrđenih elemenata dodeljuje lingvistički izraz po petostepenoj skali prikazanoj na slici 2. U tabeli 1 prikazane su hipotetičke ocene recenzenta, za jedan naučnoistraživački projekat, iskazane preko lingvističkih izraza.

**Tabela 1:** Ocene recenzenta za jedan projekat

Naziv projekta: PROJEKAT xxxx									
Recenzent	Elementi za ocenjivanje								
	E1	E2	E3	E4	E5	E6	E7	E8	E9
Recenzent 1	H	VH	H	M	H	VH	VH	H	H
Recenzent 2	VH	H	H	L	H	VH	H	H	M

U sledećem koraku, lingvistički izrazi se transformišu u fuzzy brojeve prema već utvrđenoj skali za defazifikaciju lingvističkih izraza. U skladu sa opisanim postupkom, primenom izraza (2), izračunava se srednja ocena recenzije za svakog recenzenta ponaosob, a zatim se ta ocena primenom izraza (3) defazifikuje u konačnu ocenu recenzije naučnoistraživačkog projekta, od strane tog recenzenta. U tabeli 2 prikazane su ocene recenzenta broj 1, za Projekat XXX, izračunate na opisani način.

**Tabela 2:** Ocene jednog recenzenta

Recenzent	Element za ocenu	Lingvistički izraz	Intenzitet uticaja		
			a	b	c
Recenzent 1	E1	H	5	7	9
	E2	VH	7	9	10
	E3	H	5	7	9
	E4	M	3	5	7
	E5	H	5	7	9
	E6	VH	7	9	10
	E7	VH	7	9	10
	E8	H	5	7	9
	E9	H	5	7	9
<b>Srednja ocena recenzenta (<math>A_{sr}</math>)</b>			<b>5,44</b>	<b>7,44</b>	<b>9,11</b>
<b>Ocena projekta (<math>A_{def1}</math>)</b>			<b>7,33</b>		

Nakon izračunavanja defazifikovane vrednosti ocene svakog recenzenta, primenom izraza (4) vrši se proračun srednje vrednosti ocene recenziranog projekta. Konačne ocene izvršenih recenzija, prikazane su u tabeli 3.

**Tabela 3:** Konačne ocene izvršenih recenzija

Naučnoistraživački projekat	Recenzent	Ocena recenzenta	Konačna ocena projekta (O)
Projekat br. 1	Recenzent 1	7,11	6,98
	Recenzent 2	6,85	
Projekat br. 2	Recenzent 1	8,67	8,48
	Recenzent 2	8,30	
Projekat br. 3	Recenzent 1	7,33	7,02
	Recenzent 2	6,70	
Projekat br. 4	Recenzent 1	5,63	5,54
	Recenzent 2	5,44	
Projekat br. 5	Recenzent 1	5,85	6,28
	Recenzent 2	6,70	

Na opisani način dobijaju se brođane ocene recenziranih projekata koje je lakše porediti, nego ocene koje se dobijaju na postojeći način, a koje se iskazuju kao pozitivne, odnosno negativne.

## 5. ZAKLJUČAK

Zbog izuzetnog značaja recenziranja za razvoj nauke i društva uopšte, od recenzenata se očekuje da budu kompetentni, eksperti u odgovarajućoj oblasti i da poseduju dovoljno znanja i iskustva da procene da li su u opisanom istraživanju korišćene prikladne metode i/ili eksperimenti, da li su dobijeni rezultati tačni i pouzdani, a interpretacija rezultata razumna i zasnovana na postojećim znanjima [1].

Postojeći način recenzije naučnoistraživačkih projekata u Ministarstvu odbrane i Vojsci Srbije daje samo opisnu ocenu koja može biti pozitivna ili negativna, bez bilo kakve gradacije među projektima. Ovakav način prezentovanja rezultata recenzija nije bio prihvatljiv za jedan nivo menadžmenta u Ministarstvu odbrane i Vojsci Srbije, pa je ovim radom napravljen pokušaj da se čitav proces na neki način kvantifikuje, kako bi bilo lakše porediti ocene izvršenih recenzija naučnoistraživačkih projekata.

Predloženi model predstavlja pokušaj, da se u određenoj meri, način recenzija naučnoistraživačkih projekata u Ministarstvu odbrane i Vojsci Srbije kvantifikuje, primenom fuzzy logike. Na bazi pretpostavke da je recenzentima lakše da putem određenih lingvističkih izraza ocene neki element, nego da daju brođanu ocenu, primenjena je fuzzy logika. Prezentovanim modelom predviđeno je da recenzenti svoje ocene daju opisno izborom nekog od ponuđenih lingvističkih izraza.

Za razliku od postojećeg načina recenzije naučnoistraživačkih projekata u Ministarstvu odbrane i Vojsci Srbije, primenom opisanog modela dobija se kvantifikovana ocena izvršene recenzije koja omogućava određenu vrstu poređenja uspešnosti recenziranih projekata. Takođe, dobijene ocene recenziranih projekata moguće je izraziti i u vidu procenata uspešnosti, računajući u odnosu na maksimalno moguću ocenu.

Modifikovani model opisan u radu ostavlja mogućnost da se kao lingvistički izrazi, kojima će se ocenjivati elementi po kojima se recenziraju projekti, koristi npr. sedmostepena skala kojom bi se napravila još detaljnija gradacija ocena svakog elementa. Takođe, model se lako može prilagoditi eventualnim promenama broja elemenata za ocenjivanje ili broja recenzenata.

## LITERATURA

- [1] Dekanski, A., Drvenica, I., & Nedić, O. (2017). Kako recenzirati naučni rad. *Zaštita materijala*, 58(3), 259-270.
- [2] Ghazinoory S., Esmail Zadeh A. & Memariani A. (2007). Fuzzy SWOT analysis, *Journal of Intelligent & Fuzzy Systems*, 18, 99-108.
- [3] Pamučar, D., Božanić, D. & Kurtov, D. (2016). Fazifikacija Saaty-jeve skale i prikaz hibridnog modela fuzzy AHP - TOPSIS - primer izbora vatrenog položaja brigadne artiljerijske grupe u odbrambenoj operaciji, *Vojnotehnički glasnik*, vol. 64, no. 4, pp. 966-986.
- [4] Pravilnik o načinu organizovanja, planiranja, sprovođenja i razvoja naučnoistraživačke delatnosti u Ministarstvu odbrane i Vojsci Srbije, „Službeni vojni list”, broj 1/21.
- [5] Rešenje o obrazovanju Stalnog radnog tela Saveta za naučnoistraživačku delatnost u Ministarstvu odbrane o recenziji i realizaciji naučnoistraživačkih projekata, „Službeni vojni list”, br. 13/21 i 2/22.

## MULTI-CRITERIA ANALYSIS IN DEVELOPMENT OF ACADEMIC STUDY PROGRAMS FOR EXPERTISE ON HAZARDOUS MATERIALS SAFETY AND PROTECTION

JOVICA BOGDANOV<sup>1</sup>, ZORAN BAJIĆ<sup>1</sup>, ZLATE VELIČKOVIĆ<sup>1</sup>, MIHAEL BUČKO<sup>1</sup>

<sup>1</sup> University of Defense in Belgrade, Military Academy, [jovica.bogdanov@va.mod.gov.rs](mailto:jovica.bogdanov@va.mod.gov.rs), [zoran.bajic@va.mod.gov.rs](mailto:zoran.bajic@va.mod.gov.rs), [zlate.velickovic@va.mod.gov.rs](mailto:zlate.velickovic@va.mod.gov.rs), [mihael.bucko@va.mod.gov.rs](mailto:mihael.bucko@va.mod.gov.rs)

**Abstract:** *Very significant part of military materiel is classified as hazardous material. Modern regulations stipulate specialized expertise for personnel that are supposed to work with hazardous materials, which requires appropriate education and training. According to its traditions in higher education of military officers and engineers for work with different hazardous materials, University of Defense in Belgrade has developed and accredited study programs “Chemical Engineering of Materials and Protection” for education of experts that fulfill the requirements of Ministry of Defense and Army of Serbia. These programs can provide appropriate expertise for other organizations also, primarily from defense industry and security sectors. The method of multi-criteria analysis using Analytical Hierarchy Process (AHP), that was developed for analysis of previous bachelor and master academic study programs in the field of chemical engineering since 1993, was used during the development of new study programs in order to achieve better quality of education and expertise of personnel in the future. The results of the analysis are presented in this paper, where new study programs are best ranked.*

**Keywords:** *multi-criteria analysis, academic study program, hazardous materials, chemical engineering.*

### 1. INTRODUCTION

Serbia has long tradition of military institutions of higher education. Since 1850 these institutions have been educating not only military officers for military units but also engineers for domestic defense industry. Such combined educational goal was established from its early beginnings, where military was tasked to be the main driving force of domestic industrialization [1–2]. The best example is the first Statute of Artillery School, where cadets were supposed to be educated for artillery officers, but with engineering curriculum, which provided “work in cannon and ammunition production workshops” [1].

Such approach proved to be successful. New educational institutions and curriculums were established, accompanied by many organizational reforms according to social and political changes. Modern education of military officers and engineers for defense industry is concentrated in Military Academy in Belgrade and University of Defense in Belgrade since 2000 and 2012, respectively. Different study programs are developed and accredited according to domestic legislation. Some programs were developed in cooperation with other universities, in order to improve quality of education and cooperation between institutions. Current study programs cover three educational levels (bachelor, master and doctorate) within several different scientific fields and branches.

All military activities include hazardous materials, either their storage, transport and usage, protection from them or their safe and environmentally acceptable disposal. Education for work with hazardous materials has its own specific characteristics. First, there are significantly different hazardous materials thus many different methods and techniques are used with them. Also, quality of education is a key factor of safety considerations. In current perspective, academic studies in the field of chemical engineering offer appropriate expertise for work with hazardous materials. Such higher education has its traditions in Serbian military educational institutions for more than 50 years. Similar approach can be seen in foreign military universities and academies as well. For the last accreditation cycle, new bachelor, master and doctorate study programs “Chemical Engineering of Materials and Protection” were developed by Military Academy in Belgrade and were successfully accredited in 2021. Primarily, purpose of programs is education of military personnel for three different groups of hazardous materials: explosive ordnance, chemical-biological-radiological-nuclear (CBRN) protection and flammable substances (fuels, lubricants etc.).

Multi-criteria decision making can be useful tool in evaluation of a curriculum [3–4]. Development of a modern study program is stepwise and complex process, since many limitations are imposed by legislature and accreditation procedures. Other issue is the need for quality of offered knowledge, practical skills and other

expertize for work with hazardous materials. Also, it is worth mentioning that engineering curriculums constantly need more changes because of their nature [3].

During the development phase, it was possible to do multi-criteria analysis of bachelor and master study programs “Chemical Engineering of Materials and Protection” and compare the results with previous study programs with same or similar purpose [5]. Based on the results, the factors can be adjusted in order to achieve better performance of a study program.

## 2. METHODS OF ANALYSIS

Development of study programs “Chemical Engineering of Materials and Protection” was based on general requirements of Army of Serbia, which constituted framework for further development of curriculums. Also, study programs must fulfill all regulations on higher education, in order to be accredited. Four different study programs for education of engineers of chemical engineering in the period 1993–2018 were previously analyzed [5] using Analytic Hierarchy Process (AHP) method [6]:

- Study program “Technical Service”, specialty “Ammunition and Energetic Materials” (Military Technical Academy, Belgrade, 1993);
- Study program “Technical Service”, specialties “Explosive Ordnance” and “Propulsion and Protection” (Military Technical Academy, Belgrade, 1997);
- Study programs of bachelor and master academic studies “Military Chemical Engineering”, specialties “Explosive Ordnance” and “CBRN Protection” (Military Academy, Belgrade, 2009);
- Study programs of bachelor and master academic studies “Military Chemical Engineering”, specialties “Explosive Ordnance”, “Propulsion and Protection” and “CBRN Protection” (Military Academy, Belgrade, 2009);

Since study programs from 1993 and 1997 were based on previous legislative, where higher education in engineering lasted five years, study programs after 2009 are comparable only if considered integrally (bachelor and master).

Following criteria were chosen for analysis [5]:

- hours of scientific vocational courses (C1);
- hours of applied vocational courses (C2);
- hours of scientific and applied vocational exercises and practice (C3);
- hours of military education (C4);
- hours of vocational and military practice (C5);
- hours of other technical and natural science courses (C6);
- hours of foreign language courses (C7);
- hours of other social science courses (C8);
- elective courses factor (C9);
- hours of elective courses (C10);
- average hours of active lessons in a semester (C11);
- average number of exams in a semester (C12).

Each criterion was assigned a number on a Saaty’s comparison scale [6] and the eigenvalue vector for weights  $w_i$  was determined, presented in Table 1.

**Table 1:** Comparison matrix and relative weights of criterions

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	$w_i$
C1	1	1/3	1/3	1	1/3	1	3	3	1/3	1/3	1/3	1/5	0,0434
C2	3	1	1	1	1	3	5	9	1	5	3	3	0,1439
C3	3	1	1	3	1	3	7	7	1	3	5	5	0,1588
C4	1	1	1/3	1	1/3	1	5	9	1/2	1	1	3	0,0791
C5	3	1	1	3	1	3	7	9	1	1	5	5	0,1517
C6	1	1/3	1/3	1	1/3	1	5	7	1/2	1	3	3	0,0737
C7	1/3	1/5	1/5	1/5	1/7	1/5	1	3	1/3	1/3	1/3	1/3	0,0235
C8	1/3	1/9	1/7	1/9	1/9	1/7	1/3	1	1/5	1/3	1/5	1/5	0,0143
C9	3	1	1	2	1	2	3	5	1	3	5	3	0,1313
C10	3	1/5	1/3	1	1	1	3	3	1/3	1	3	1	0,0692
C11	3	1/3	1/5	1	1/5	1/3	3	5	1/5	1/3	1	1/5	0,0453
C12	5	1/3	1/5	1/3	1/5	1/3	3	5	1/3	1	5	1	0,0657



Such judgement has following consistency parameters:  $\lambda_{\max} = 13,6303$ , consistency index  $CI = 0,1482$  and consistency ratio  $CR = 9,62\%$ .

Using this method, the performance of study programs of bachelor and master study programs “Chemical Engineering of Materials and Protection” was constantly monitored and curriculum structure and hours of appropriate courses were stepwise modified during the development phase. Criterion values  $x_i$  for previous and final version of new study programs are presented in Table 2.

**Table 2:** Criterion values for considered study programs

Criterion	Technical Service, 1993	Technical Service, 1997	Military Chemical Engineering, 2009	Military Chemical Engineering, 2017	Chemical Engineering of Materials and Protection, 2021
C1	705	735	435	555	892
C2	825	885	840	975	915
C3	630	705	555	615	757
C4	1050	915	705	495	555
C5	910	910	652	1034	1034
C6	1410	1335	825	1055	802
C7	120	120	240	360	240
C8	195	90	90	120	45
C9	0	2	2	2,72	3,42
C10	0	990	905	1620	1665
C11	460	432	342	410	398
C12	5,5	6,1	4,8	5,6	5,7

In order to compare study program, ratio of individual criterion values  $X_i$  was calculated using following expression:

$$X_i = \frac{x_i}{(x_i)_{\max}} \quad (1)$$

Using these results and AHP weights, values of Weighted Linear Combination (WLC) were calculated for all considered study programs (Table 3):

$$S_j = \sum(w_i \cdot X_i) \quad (2)$$

**Table 3:** Analysis results for considered study programs

Technical Service, 1993	Technical Service, 1997	Military Chemical Engineering, 2009	Military Chemical Engineering, 2017	Chemical Engineering of Materials and Protection, 2021
70,12%	82,72%	67,56%	84,89%	89,28%

After the development, new study programs were approved by Ministry of Defense and Army of Serbia and during the accreditation process their quality was highly graded by external evaluation committee.

### 3. CONCLUSION

The previously developed method was used for performance evaluation of new study programs during their development phase. It proved to be a valuable tool during the development phase, where curriculums and courses were stepwise modified in order to achieve better performance. The analysis results presented in this paper show that new study programs of bachelor and master study programs “Chemical Engineering of Materials and Protection” from 2021 have better performance than all previous study programs with similar

purpose. It can be concluded that new study programs offer high quality of specialized expertise for work with hazardous materials, which can positively contribute to the safety of individuals and whole social environment.

Quantitative methods are considered to have many limitations and there can be noticed strong skepticism around their use in the field of higher education [7]. It can be said that presented method can be used for comparison and performance evaluation of study programs with similar purpose or in the same field or branch of science.

### **Acknowledgement**

The paper is part of scientific research project “VA-TT/1/22-24” of University of Defense in Belgrade and funded by Ministry of Defense of Republic of Serbia.

### **REFERENCES**

- [1] Maksimović, V. (1925). Artillery School and Military Academy (in Serbian). In Memorial of 75<sup>th</sup> Anniversary of Military Academy 1850–1925. Belgrade: Štamparska radionica Ministarstva Vojske i Mornarice.
- [2] Pršić, M. (2000). Military schools in Serbia (in Serbian). In B. Ljušić, S. Bojković, M. Pršić, & B. Jovović, Officers in Serbian higher education 1804–1918, 201-340. Belgrade: Vojnoizdavački zavod.
- [3] Erkan, T. E., & Rouyendegh, B. D. (2013). Curriculum change parameters determined by multi criteria decision making (MCDM). *Procedia – Social and Behavioral Sciences*, 116, 1744–1747.
- [4] Kawser, M. A. (2014). Curriculum Development in Higher Education. University of Oslo, Oslo: Representralen.
- [5] Bogdanov, J., Bajić, Z., & Bučko, M. (2018). Analysis of study programs for education of chemical technology engineers in military academies between 1993 and 2018 (in Serbian). In Proceeding of XLV Symposium on Operational Research “SYM-OP-IS 2018”, Zlatibor, September 16-19, 325-330.
- [6] Saaty, T. (1980). *The Analytical Hierarchy Process*. Pittsburgh: RWS Publications.
- [7] Katharaki, M., & Katharakis, G. (2010). A comparative assessment of Greek universities’ efficiency using quantitative analysis. *International Journal of Educational Research*, 49, 115–128.

## ODREĐIVANJE VELIČINE POPREČNOG PRESEKA ELASTIČNE GREDE BEZLEŽAJNOG ROTORA BESPILOTNOG HELIKOPTERA

### OPTIMIZING THE SIZE OF CROSS-SECTION OF ELASTIC BEAM OF THE BEARINGLESS ROTOR OF AN UMANNED HELICOPTER

DALIBOR PETROVIĆ<sup>1</sup>, VLADIMIR STANKOVIĆ<sup>2</sup>, IVAN MUDRI<sup>3</sup>

<sup>1</sup>Univerzitet odbrane – Vojna akademija, daliborpetrovic140@gmail.com

<sup>2</sup>Vojska Srbije – RV i PVO, vladastankovic1995@gmail.com

<sup>3</sup>Vojska Srbije – RV i PVO, ivan.mudri2015@gmail.com

**Rezime:** Glavni element bezležajnih rotora helikoptera je elastična greda, koja zamenjuje sva tri zgloba (zglob mahanja, zabacivanja i promene koraka), i mora da bude dobro dizajnirana kako bi obezbedila potrebna kretanja lopatice rotora, bez prekoračenja dozvoljenih napona u materijalu. Cilj rada je odrediti optimalnu veličinu poprečnog preseka glave bezležajnog rotora bespilotnog helikoptera u odnosu na funkcije mahanja i promene koraka. U radu je razmatran uticaj veličine poprečnog preseka elastičnog elementa na karakteristike mahanja i promene koraka, analizom rezultata dobijenih metodom konačnih elemenata. Optimalna veličina poprečnog preseka određena je kombinacijom interpolacije funkcije dve promenjive sa podeljenim razlikama i metodom najmanjih kvadrata za funkcije mahanja i promene koraka. Analizirani su poprečni preseki elastične grede u opsegu od 7 do 11mm, a kao optimalno rešenje dobijen je poprečni presek debljine od 9.19mm.

**Ključne reči:** Bezležajni rotor helikoptera, elastična greda, konačni elementi, interpolacija funkcije dve promenjive, podeljene razlike, metoda najmanjih kvadrata.

**Abstract:** The main element of bearingless rotor of a helicopter is elastic beam, which replaces all three hinges (flapping hinge, lead-lag hinge, and feathering hinge), and it has to be well designed in order to provide necessary movements of the rotor blade, without exceeding the permitted stresses in the material. The goal of this study is to find optimal size of cross-section of the head of the bearingless rotor of an unmanned helicopter in respect to both flapping and feathering functions. The study considers the influence of the size of the cross-section to flapping and feathering properties, by analyzing the results obtained by finite elements method. The optimal size of cross-section is found by approximating flapping and feathering functions with corresponding polynomials with divided differences and applying least square method to find the size of cross-section that leads to minimal error. In that scope, the data obtained for the size of the cross-section ranging from 7 to 11 mm were used and it was shown that the optimal value for the cross-section is equal to 9.19mm.

**Keywords:** Bearingless rotor, Elastic beam, Finite elements, Interpolation of function with two variables, Finite differences, Least square method.

## 1. UVOD

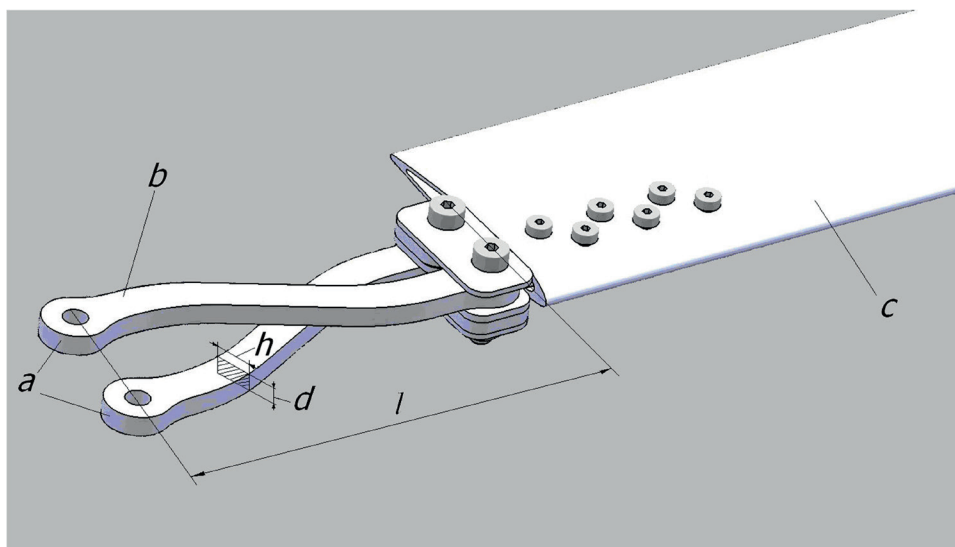
Stalna težnja proizvođača helikoptera da se uprosti konstrukcija glavnog rotora helikoptera, uspela je sa primenom elastične grede (elastičnog elementa) u konstrukciji rotora. Elastična greda je svojim oblikom, veličinom i kompozitnim materijalom od koga je sačinjena zamenila dotadašnja tri zgloba (zglob mahanja, zabacivanja i promene koraka). U cilju postizanja što boljih karakteristika glavnog rotora, to jest, dobijanja karakteristika sličnih šarnirnim rotorima, u praksi se mogu videti različiti oblici elastičnih greda sa različitim poprečnim presecima i dužinama [4]. Da bi se dobile što bolje karakteristike torzije, mahanja i zabacivanja lopatice rotora, veoma je bitno što bolje dizajnirati samu gredu, jer je ona u ovom slučaju glavni element od koga zavise karakteristike upravljanja. Zbog visoko povezanih dinamičkih i strukturnih zahteva elastične grede koji se odnose na njenu dužinu, veličinu poprečnog preseka elemenata, razmak elastičnih elemenata u korenu i na lopatici, itd, neophodno je da se, koristeći matematičke metode optimizacije, dođe do optimalnih vrednosti geometrijskih karakteristika elastične grede koje zadovoljavaju tražene dinamičke karakteristike (mahanje, zabacivanje i promenu koraka). Iako su svi strukturalni zahtevi međusobno povezani, veoma teško ih je sve zajedno posmatrati i analizirati, imajući u vidu da je svaki pojedinačni zahtev veoma kompleksan. U

vezi sa tim, u radu se razmatra uticaj veličine poprečnog preseka elastičnog elementa na karakteristike mahanja i promene koraka.

## 2. TEHNIČKI PROBLEM

Glavni tehnički problem kod bezležajnih rotora je ostvarivanje potrebnog stepena slobode zabacivanja, mahanja i promene smeštajnog ugla lopatice helikoptera, deformacijom elastične grede (slika 1) koja spaja lopaticu sa glavom rotora. Elastična greda sastoji se od dva specijalno oblikovana elastična elementa postavljena jedan iznad drugog na međusobnom rastojanju. Elementi su izrađeni od laminarnog kompozita, koga čine čelični limovi međusobno spojeni gumom. Čelik elastičnim gredama daje potrebnu krutost i žilavost, a guma nelinearno ponašanje prilikom deformacije.

Za razliku od savremenih konstrukcijskih rešenja greda glavnih rotora helikoptera kod kojih su prisutni različiti geometrijski oblici poprečnih preseka (počev od pravougaonih, krstastih „+“, višestrukih „H“, zatvorenih do višestrukih otvorenih poprečnih preseka), kod analizirane grede primenjen je pavougaoni poprečni presek. Pravougaoni poprečni presek grede je izabran zbog njegove velike sposobnosti nosivosti, pojednostavljenog dizajna, a samim tim i proizvodnog procesa, što kao rezultat ima niže troškove izrade i kontrole. Međutim, da bi elastični element imao zadovoljavajuće karakteristike pomeranja, u ovom slučaju mahanja i torzije, potrebno je definisati veličinu poprečnog preseka elastične grede [5].



**Slika 1:** Elastična greda i lopatica: a) greda, b) elastični element, c) lopatica rotora, d) visina elementa, h) širina elementa i l) dužina grede

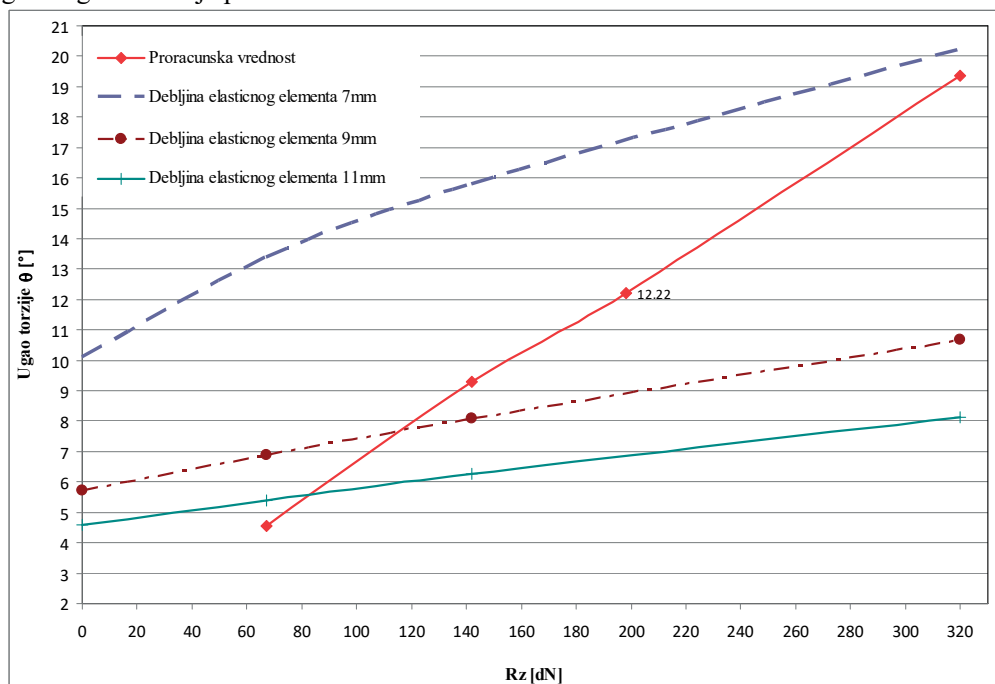
## 3. ZAHTEVI OPTIMIZACIJE I ULAZNI PODACI

Primarni cilj dimenzionisanja elastične grede rotora helikoptera je postizanje potrebne dinamike kretanja bez prekoračenja dozvoljenih napona u materijalu. Realizacija ovog cilja izvršena je optimizacijom površine poprečnog preseka elastičnog elementa, a samim tim i kompletne grede [2]. Prilikom određivanja veličine poprečnog preseka elastične grede varirani su visina elementa ( $d$ ) dok je širina ( $h$ ) ostala ista.

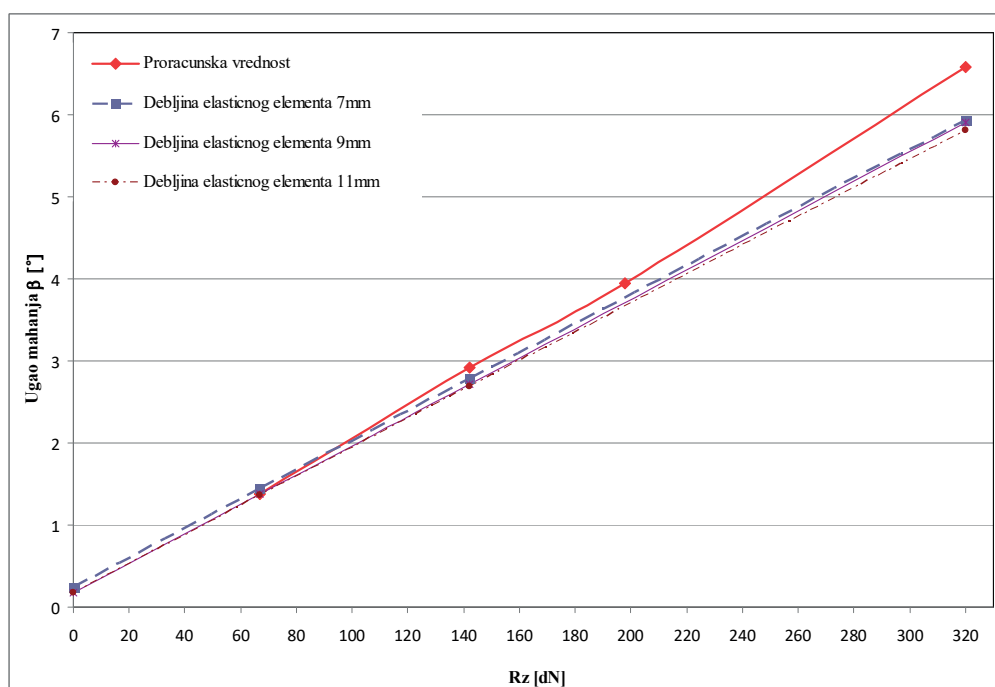
Imajući u vidu da relativno mekane elastične grede rotora helikoptera mogu imati značajna pomeranja u smislu savijanja i uvijanja, kao i da mogu dovesti do povećanog nivoa vibracija u lopaticama (što bi za posledicu imalo ugradnju sistema za suzbijanje vibracija), u radu je za minimalnu analiziranu debljinu elastičnog elementa izabrana debljina  $d=7\text{mm}$ .

Ulazni podaci dobijeni su analizom elastične grede metodom konačnih elemenata. Metodom konačnih elemenata [8] simulirano je ponašanje elastične grede dužine  $l=240\text{mm}$  i debljine  $d=7, 9$  i  $11\text{mm}$ . Analiza je izvršena pri sledećim uslovima. Elastični elementi su stegnuti u korenu za glavu vratila, to jest, površine koje dolaze u dodiru sa vijkom glave vratila su definisane tako da je njihovo pomeranje u svim pravcima 0. Centrifugalna sila  $F_c=3000\text{dN}$  koja deluje na elastični element definisana je da deluje u centru mase aeroprofila lopatice rotora, a aerodinamička sila u aerodinamičkom centru potiska aeroprofila, na  $0.75R$  od ose obrtanja rotora.

Analizirana je promena postavnog ugla (torzija elastične grede) i ugla mahanja (savijanje elastične grede) za različite sile uzgona i to  $R_z=67, 142$  i  $320\text{dN}$ . Sila uzgona od  $67\text{dN}$  odgovara auto rotaciji,  $142\text{dN}$  uzgonskoj sili na azimutu  $0$  pri brzini  $120\text{km/h}$  i  $320\text{dN}$  na azimutu  $0$  pri brzini  $180\text{km/h}$ . Dobijeni podaci za postavni ugao i ugao mahanja prikazani su na slikama 1 i 2.



Slika 1: Promena postavnog ugla (ugla torzije) lopatice rotora u odnosu na silu uzgona



Slika 2: Promena ugla mahanja lopatice rotora u odnosu na silu uzgona

#### 4. INTERPOLACIJA FUNKCIJA VIŠE PROMENLJIVIH

Aproksimacija funkcije mahanja  $\beta$  kao funkcije dve promenljive (debljina poprečnog preseka elastične grede  $d$  i veličina uzgonske sile  $R_z$ ), odnosno  $\beta(d, R_z)$ , izvršena je korišćenjem interpolacionog polinoma sa podeljenim razlikama [3] na osnovu zadatih vrednosti  $\beta(d_i, R_{z_j})$ ,  $i = 0, \dots, n$ ,  $j = 0, \dots, n$ . Na isti način izvršena je aproksimacija funkcije promene koraka lopatice rotora  $\theta$  (ugla torzije) koja takođe zavisi od promenljivih  $d$  i

Rz. U nastavku će ukratko biti izložena interpolacija funkcije dve promenljive polinomom sa podeljenim razlikama.

Neka je funkcija  $\beta(d,Rz)$  zadata svojim vrednostima u  $(n+1)^2$  čvorova  $\beta_{ij}=\beta(d_i,Rz_j)$ ,  $i = 0,\dots,n$ ,  $j = 0,\dots,n$  prikazanih sledećom tabelom:

**Tabela 1:** Promena ugla mahanja u odnosu na silu uzgona i debljinu elastične grede

$\beta_{ij}$	$Rz_0$	$Rz_1$	...	$Rz_n$
$d_0$	$\beta(d_0, Rz_0)$	$\beta(d_0, Rz_1)$	...	$\beta(d_0, Rz_n)$
$d_1$	$\beta(d_1, Rz_0)$	$\beta(d_1, Rz_1)$	...	$\beta(d_1, Rz_n)$
...	...	...	...	...
$d_n$	$\beta(d_n, Rz_0)$	$\beta(d_n, Rz_1)$	...	$\beta(d_n, Rz_n)$

Na isti način su zadate vrednosti funkcije  $\theta(d, Rz)$ .

**Tabela 2:** Promena ugla torzije u odnosu na silu uzgona i debljinu elastične grede

$\theta_{ij}$	$Rz_0$	$Rz_1$	...	$Rz_n$
$d_0$	$\theta(d_0, Rz_0)$	$\theta(d_0, Rz_1)$	...	$\theta(d_0, Rz_n)$
$d_1$	$\theta(d_1, Rz_0)$	$\theta(d_1, Rz_1)$	...	$\theta(d_1, Rz_n)$
...	...	...	...	...
$d_n$	$\theta(d_n, Rz_0)$	$\theta(d_n, Rz_1)$	...	$\theta(d_n, Rz_n)$

Podeljene razlike za funkciju  $\beta(d,Rz)$  definišu se na sledeći način [6]:

$$\text{Podeljena razlika reda } 0+0: \quad \beta[d_0; Rz_0] = \beta(d_0, Rz_0) \quad (1)$$

$$\text{Podeljena razlika reda } 1+0: \quad \beta[d_0, d_1; Rz_0] = \frac{\beta[d_1; Rz_0] - \beta[d_0; Rz_0]}{d_1 - d_0} \quad (2)$$

$$\text{Podeljena razlika reda } 0+1: \quad \beta[d_0; Rz_0, Rz_1] = \frac{\beta[d_0; Rz_1] - \beta[d_0; Rz_0]}{Rz_1 - Rz_0} \quad (3)$$

⋮

$$\text{Podeljena razlika reda } i+j: \quad \beta[d_0, d_1, \dots, d_i; Rz_0, Rz_1, \dots, Rz_j] = \frac{\beta[d_1, \dots, d_i; Rz_0, \dots, Rz_j] - \beta[d_0, \dots, d_{i-1}; Rz_0, \dots, Rz_j]}{d_i - d_0} \quad (4)$$

$$= \frac{\beta[d_0, \dots, d_i; Rz_1, \dots, Rz_j] - \beta[d_0, \dots, d_i; Rz_0, \dots, Rz_{j-1}]}{Rz_j - Rz_0}$$

Označimo sa  $a_{ij} = \beta[d_0, \dots, d_i; Rz_0, \dots, Rz_j]$ . Tada polinom sa podeljenim razlikama [7] koji aproksimira funkciju  $\beta(d,Rz)$  ima sledeći oblik:

$$P_\beta(d, Rz) = \sum_{k=0}^n \sum_{i+j=k} a_{ij} (d-d_0) \dots (d-d_{i-1}) \cdot (Rz - Rz_0) \dots (Rz - Rz_{j-1}) \quad (5)$$

$$P_\beta(d, Rz) = \sum_{k=0}^n \sum_{i+j=k} \beta[d_0, \dots, d_i; Rz_0, \dots, Rz_j] (d-d_0) \dots (d-d_{i-1}) \cdot (Rz - Rz_0) \dots (Rz - Rz_{j-1}) \quad (6)$$

Slično, ako uvedemo oznaku  $b_{ij} = \theta[d_0, \dots, d_i; Rz_0, \dots, Rz_j]$ , tada polinom sa podeljenim razlikama koji aproksimira funkciju  $\theta(d,Rz)$  ima sledeći oblik:

$$P_\theta(d, Rz) = \sum_{k=0}^n \sum_{i+j=k} b_{ij} (d-d_0) \dots (d-d_{i-1}) \cdot (Rz - Rz_0) \dots (Rz - Rz_{j-1}) \quad (7)$$

$$P_\theta(d, Rz) = \sum_{k=0}^n \sum_{i+j=k} \theta[d_0, \dots, d_i; Rz_0, \dots, Rz_j] (d-d_0) \dots (d-d_{i-1}) \cdot (Rz - Rz_0) \dots (Rz - Rz_{j-1}) \quad (8)$$

U našem konkretnom slučaju je  $n=2$  i tabele 1 i 2 imaju sledeći oblik:

**Tabela 3:** Vrednost ugla mahanja u odnosu na silu uzgona i debljinu elastične grede

$\beta_{ij}$		Rz <sub>0</sub>	Rz <sub>1</sub>	Rz <sub>2</sub>
		67dN	142dN	320dN
d <sub>0</sub>	7mm	1.44	2.78	5.93
d <sub>1</sub>	9mm	1.38	2.72	5.91
d <sub>2</sub>	11mm	1.36	2.68	5.8

**Tabela 4:** Vrednost postavnog ugla u odnosu na silu uzgona i debljinu elastične grede

$\theta_{ij}$		Rz <sub>0</sub>	Rz <sub>1</sub>	Rz <sub>2</sub>
		67dN	142dN	320dN
d <sub>0</sub>	7mm	13.37	15.8	20.25
d <sub>1</sub>	9mm	6.88	8.07	10.7
d <sub>2</sub>	11mm	5.39	6.26	8.11

Koristeći vrednosti date u tabelama 3 i 4 možemo konstruisati polinome sa podeljenim razlikama  $P_\beta(d, Rz)$  i  $P_\theta(d, Rz)$  koji imaju sledeće oblike:

Polinom po dve promenljive koji aproksimira funkciju  $\beta(d, Rz)$  ima izraz:

$$P_\beta(d, Rz) = 1.44 - 0.0300 \cdot (d - 7) + 0.0179 \cdot (Rz - 67) - 0.627 + 0.000001 \cdot (Rz - 67) \cdot (Rz - 142) + 2.68 \cdot (d - 7) \cdot (d - 9) \quad (9)$$

Polinom po dve promenljive koji aproksimira funkciju  $\theta(d, Rz)$  ima izraz:

$$P_\theta(d, Rz) = 3.37 - 3.245 \cdot (d - 7) + 0.0324 \cdot (Rz - 67) + 0.0083 \cdot (d - 7) \cdot (Rz - 67) - 2.92 \cdot 0.0000 \cdot (Rz - 67) \cdot (Rz - 142) + 0.626 \cdot (d - 7) \cdot (d - 9) \quad (10)$$

Kako su proračunske krive  $\beta(Rz)$  i  $\theta(Rz)$  poznate, funkciju greške možemo definisati kao sumu kvadrata odstupanja [1,7] polinoma  $P_\beta(d, Rz)$  do  $\beta(Rz)$  i polinoma  $P_\theta(d, Rz)$  do  $\theta(Rz)$  za vrednosti  $Rz_0, \dots, Rz_n$  (u našem slučaju  $n=2$ ). Ovako definisana funkcija greške nam omogućava da odredimo optimalnu vrednost za  $d$  i u odnosu na funkciju mahanja i u odnosu na funkciju promene koraka. Funkcija greške će biti funkcija po promenljivoj  $d$ . Preciznije, važi:

$$G(d) = \sum_{i=0}^n \left[ (P_\beta(d, Rz_i) - \beta(Rz_i))^2 + (P_\theta(d, Rz_i) - \theta(Rz_i))^2 \right] \quad (11)$$

$$G(d) = (P_\beta(d, Rz_0) - \beta(Rz_0))^2 + (P_\beta(d, Rz_1) - \beta(Rz_1))^2 + (P_\beta(d, Rz_2) - \beta(Rz_2))^2 + (P_\theta(d, Rz_0) - \theta(Rz_0))^2 + (P_\theta(d, Rz_1) - \theta(Rz_1))^2 + (P_\theta(d, Rz_2) - \theta(Rz_2))^2 \quad (12)$$

Ovde su  $\beta(Rz_0) = \beta(67) = 1.38$ ,  $\beta(Rz_1) = \beta(142) = 6.58$ ,  $\beta(Rz_2) = \beta(320) = 6.58$  vrednosti za proračunsku krivu.

Takođe, važi:

$$P_\beta(d, Rz_0) = (2.68 \cdot d - 18.76) \cdot (d - 9) - 0.03 \cdot d + 1.02 \quad (13)$$

$$P_\beta(d, Rz_1) = (2.68 \cdot d - 18.76) \cdot (d - 9) - 0.03 \cdot d + 2.37 \quad (14)$$

$$P_\beta(d, Rz_2) = (2.68 \cdot d - 18.76) \cdot (d - 9) - 0.03 \cdot d + 5.6 \quad (15)$$

Ovde su  $\theta(Rz_0) = \theta(67) = 4.57$ ,  $\theta(Rz_1) = \theta(142) = 9.29$ ,  $\theta(Rz_2) = \theta(320) = 19.35$  vrednosti za proračunsku krivu. Takođe, važi:

$$P_{\theta}(d, Rz_0) = (0.626 \cdot d - 4.382) \cdot (d - 9) - 3.245 \cdot d + 36.085 \quad (16)$$

$$P_{\theta}(d, Rz_1) = (0.626 \cdot d - 4.382) \cdot (d - 9) - 2.623 \cdot d + 34.158 \quad (17)$$

$$P_{\theta}(d, Rz_2) = (0.626 \cdot d - 4.382) \cdot (d - 9) - 1.145 \cdot d + 29.583 \quad (18)$$

Kad se sračuna, dobija se:

$$G(d) = (0.03 \cdot d - (2.68 \cdot d - 18.76) \cdot (d - 9) + 0.98)^2 + (0.03 \cdot d - (2.68 \cdot d - 18.76) \cdot (d - 9) + 0.357)^2 + \\ + (0.03 \cdot d - (2.68 \cdot d - 18.76) \cdot (d - 9) + 0.555)^2 + ((0.626 - 4.382) \cdot (d - 9) - 3.245d + 31.515)^2 + \quad (19) \\ + ((0.626 - 4.382) \cdot (d - 9) - 2.62d + 24.87)^2 + ((0.626 - 4.382) \cdot (d - 9) - 1.145d + 10.233)^2$$

Naš zadatak je da odredimo minimum funkcije  $G(d)$ , odnosno optimalni korak za koje je odstupanje i u odnosu na funkciju mahanja i funkciju koraka najmanje. Kako je u pitanju nenegativna funkcija, definisana kao suma kvadrata,  $G(d)$  ima jedinstveni globalni minimum koji se dobija kao nula prvog izvoda, odnosno kao rešenje nelinearne jednačine  $G'(d) = 0$ .

Rešavajući nelinearnu jednačinu  $G'(d) = 0$  primenom Njutnove metode [1,7], dobija se  $d_{opt} = 9.19$  mm.

## 5. ZAKLJUČAK

U radu je pokazana mogućnost optimizacije poprečnog preseka elastične grede u odnosu na funkciju mahanja i promene koraka lopatice rotora primenom kombinacije interpolacije funkcije dve promenjive sa podeljenim razlikama i metode najmanjih kvadrata. Ulazni podaci dobijeni su analizom ponašanja elastične grede debljine  $d=7, 9$  i  $11$ mm, metodom konačnih elemenata. Prilikom analize poprečnog preseka širina elementa je bila konstantna dok se menjala samo visina. Cilj rada je odrediti optimalnu veličinu poprečnog preseka glave bezležajnog rotora bespilotnog helikoptera koja će zadovoljiti zadate dinamičke karakteristike rotora. Optimizacijom dobijenih rezultata, u odnosu na funkciju mahanja i promene koraka dobijena je vrednost debljine poprečnog preseka od  $9.19$ mm.

## LITERATURA

- [1] Burden R.L & Faires J.D. Numerical Analysis. Brooks/ Cole: USA, 2001; 104–163.
- [2] Hamilton, B.K., & Peters, J.R. (1989). Multi-objective/loading optimization for rotating composite flexbeams. *NASA. Langley Research Center, Recent Advances in Multidisciplinary Analysis and Optimization, Part 1*, 235-256.
- [3] Kunkle T. Multivariate differences, polynomials, and splines. *J. Approx. Theory* 1996; 84: 290–314.
- [4] Manoj K.D., Sung N.J. & Tae J.K. (2014). Evolutionary Shape Optimization of Flexbeam Sections of a Bearingless Helicopter Rotor. *Composites Research*, 27(6),207-212.
- [5] Manoj K.D., Kyu B.L., Sung N.J. & Tae J.K. (2013). Particle Swarm Assisted Genetic Algorithm for the Optimal Design of Flexbeam Sections. *International Journal of Aeronautical & Space Sciences*,14(4), 341–349.
- [6] Neamtu M. Multivariate divided difference. I. Basic properties. *SIAM J. Numer.Anal.*1992; 29: 1435–1445.
- [7] Radunović D., "Numeričke metode", Akademska misao, Beograd, 2004.
- [8] TaeY.C., Han Y.R., Hae C.S., Sang J.S., Young J.K. & Deog K.K. (2013). Structural Analysis of a Bearingless Rotor using an Improved Flexible Multibody Model. *Journal of Aircraft*, 50(2), 539-550.



## VARIRANJE REZULTATA VIŠEKRITERIJUMSKOG IZBORA U ZAVISNOSTI OD PRIMENJENE METODE VIŠEKRITERIJUMSKOG ODLUČIVANJA

### VARIATION OF THE RESULTS OF MULTI-CRITERIA SELECTION DEPENDING ON APPLIED MULTI-CRITERIA DECISION MAKING METHOD

SRĐAN LJUBOJEVIĆ<sup>1</sup>, SRĐAN DIMIĆ<sup>2</sup>, DRAGAN KOSTADINOVIĆ<sup>3</sup>, ALEKSANDAR GOŠIĆ<sup>4</sup>

<sup>1</sup> Univerzitet odbrane u Beogradu – Vojna akademija, [srdjan.ljubojevic@va.mod.gov.rs](mailto:srdjan.ljubojevic@va.mod.gov.rs)

<sup>2</sup> Univerzitet odbrane u Beogradu – Vojna akademija, [srdjan.dimic@mod.gov.rs](mailto:srdjan.dimic@mod.gov.rs)

<sup>3</sup> Univerzitet odbrane u Beogradu – Vojna akademija, [kotoza@beotel.net](mailto:kotoza@beotel.net)

<sup>4</sup> Akademija tehničko-vaspitačkih strukovnih studija, Niš – Odsek Vranje, [golic.aleksandar@gmail.com](mailto:golic.aleksandar@gmail.com)

**Rezime:** U radu je izvršena analiza variranja rezultata višekriterijumskog izbora uzrokovanog promenama ulaznih parametara u modelu odlučivanja. Analiza je izvršena na primeru izbora luke za potrebe strategijskog transporta u Vojsci Srbije i ukazano je da pažnju treba usmeriti pre na izbor odgovarajuće metode višekriterijumskog odlučivanja, nego na izbor „najbolje“ metode.

**Ključne reči:** Višekriterijumsko odlučivanje, strategijski transport, SAW, COPRAS, TOPSIS; VIKOR

**Abstract:** Variation of the results of multi-criteria selection caused by changes in input parameters in the decision-making model is analysed in this paper. The analysis is performed on the case study of a port selection problem in strategic transport in the Army of Serbia and it is pointed out that attention should be focused on the choice of an appropriate multicriteria decision-making method, rather than choosing the "best" method.

**Keywords:** Multi-criteria decision making, Strategic transport, SAW, COPRAS, TOPSIS; VIKOR

## 1. UVOD

U današnjim uslovima, pojam strategijski transport, u vojskama većine država, podrazumeva transport koji se obavlja na velikim udaljenostima i van teritorije matične države. U Vojsci Srbije (VS) strategijski transport je određen kao prevoz ljudi i sredstava van granica Republike Srbije, za potrebe učestvovanja u multinacionalnim operacijama, vežbama i drugim oblicima međunarodne vojne saradnje. Potrebe i zahtevi koje diktiraju konkretni zadaci u domenu strategijskog transporta uslovljavaju da odlučivanje bude zasnovano na principima racionalnosti, uz korišćenje pređašnjeg iskustva, a takođe impliciraju i potrebu da se neizvesnost i rizik ispravno tretiraju pri odlučivanju. Primena metoda višekriterijumskog odlučivanja (VKO) u takvim situacijama je nezaobilazna.

S obzirom na raznolikost i specifičnosti metoda i tehnika VKO, a naročito s obzirom na raznolikost njihovog matematičkog aparata, mogu se očekivati različiti rezultati primene različitih metoda i tehnika na istom problemu odlučivanja (paradoks odlučivanja koji su formulisali Triantaphyllou i Mann [14]: „Koju metodu treba koristiti za izbor najbolje metode odlučivanja?“). Zbog slučajeva u kojima rezultati primene više metoda VKO nisu međusobno podudarni, najčešće je u zaključcima različitih studija isticana potreba da se ispita pouzdanost dobijenih rešenja, odnosno da se analizira osetljivost metode VKO primenjene u konkretnom problemu.

U literaturi prisutni pristupi analizi osetljivosti metoda VAO uglavnom se zasnivaju na analizi promene rezultata rangiranja u odnosu na promenu pojedinih ulaznih parametara u modelu odlučivanja. U tom smislu, i primenu metoda VKO u zadacima strategijskog transporta treba da prati odgovarajuća analiza osetljivosti izabranih metoda, odnosno analiza stabilnosti dobijenih rezultata rangiranja.

## 2. PROBLEM IZBORA TERMINALA U STRATEGIJSKOM TRANSPORTU

Problemi odlučivanja u zadacima strategijskog transporta javljaju se u različitim formama, shodno konkretnom zadatku i situacionim uslovima. Neki od najčešćih problema su: problem izbora vida/grane i načina realizacije strategijskog transporta, problem izbora lokacija (terminala) utovara/istovara i ukrcavanja/iskrcavanja,

problem izbora rute za realizaciju transporta, problem izbora provajdera transportne usluge, problem izbora tima za kontrolu kretanja i transporta itd.

U ovom radu analiza rezultata primene VKO metoda izvršena je na problemu izbora terminala za potrebe strategijskog transporta. Pojavni oblik problema izbora terminala u strategijskom transportu zavisi od toga o kojoj vrsti terminala se radi (železnički terminali, avioterminali – aerodromi, rečne ili morske luke i sl.). U svakom slučaju, zajednička karakteristika problema izbora u svim ovim slučajevima jeste njihov višekriterijumski karakter.

Analiza primene VKO pri izboru terminala sprovedena je praćenjem promena u rezultujućim rangovima alternativa izazvanim promenama tri različita faktora: 1) težinskih koeficijenata kriterijuma odlučivanja (pri čemu je poželjna osetljivost rezultata na promenu težinskih koeficijenata kriterijuma), 2) primenjenih mernih skala za merenje vrednosti alternativa po kriterijumima (poželjna je stabilnost rezultata po uzoru na tzv. *uslov nezavisnosti od vrednosne skale* koji se primenjuje u normativnoj teoriji odlučivanja u uslovima rizika i neizvesnosti [7]) i 3) načina formulacije kriterijuma (poželjna je stabilnost rezultata po uzoru na tzv. *uslov deskriptivne invarijantnosti*, koji je u bihevioralnoj teoriji odlučivanja poznat i kao *uslov racionalnosti* izbora individualnog donosioca odluka [8]).

Dakle, najpre je izvršena postavka problema izbora i primenjene su različite metode VKO. Svaki od dobijenih rangova alternativa predstavlja rešenje problema prema određenoj metodi VKO. Ovi početni rangovi predstavljaju referentne vrednosti rangova sa kojima su, u drugom koraku, upoređeni rezultati primene istih metoda u scenarijima sa različitim vrednostima ulaznih parametara, pri čemu je analizirana stabilnost rezultata rangiranja.

U jednom od svojih pojava oblika, problem izbora terminala u strategijskom transportu, odnosno lokacije na kojoj se iz nacionalne faze transporta prelazi u strategijsku fazu, može se definisati kao problem izbora pomorske luke iz koje će se glavna oprema transportovati u zonu mirovne misije.

Kriterijumi koje pri tome treba razmatrati, mogu biti kvantitativni i kvalitativni, i u načelu, mogu biti grupisani u tri kategorije: kriterijume cene (troškova), usluge i rute transporta [6]. Često se najvažnijim smatraju kriterijumi rute transporta (najviše su opredeljeni *povoljnošću položaja luke* iz aspekta transportnih potreba klijenta). Sa druge strane, kriterijumi usluge se percipiraju kao značajniji od kriterijuma troškova, ali ipak ne i značajniji od kriterijuma rute. Interesantno je da se, prema [13], ovaj skup kriterijuma izbora ne razlikuje mnogo kada odluku donosi pošiljalac sredstava u odnosu na slučajeve kada o tome odlučuje provajder transportne usluge. Naime, nekada pošiljalac sredstava samostalno vrši izbor luke iz koje će sredstva biti transportovana brodom, a nekada on tu odluku prepusti prevozniku ili špediteru. U praksi strategijskog transporta u VS mogu biti prisutna oba slučaja, iako je češći slučaj samostalnog izbora luke.

Na osnovu analize stručne literature i frekvencije pojavljivanja kriterijuma odlučivanja pri izboru luke, može se izdvojiti 11 dominantnih kriterijuma, sa odgovarajućim težinskim koeficijentima ( $w$ ):  $K_1$  – lokacijska pogodnost luke ( $w=0,25$ ; vrednosti alternativa su bezdimenzionalne i predstavljaju odnos između udaljenosti od posmatrane luke do određene destinacije i od izvorišta – lokacije čuvanja/prikupljanja sredstava za transport do posmatrane luke),  $K_2$  – dužina ukupnog transportnog puta ( $w=0,15$ , put od izvorišta do određene destinacije),  $K_3$  – operativna efikasnost luke ( $w=0,07$ ; vrednosti alternativa se određuju na osnovu koeficijenta efikasnosti, prema metodologiji navedenoj u [11]),  $K_4$  – tehnološki nivo lučkih procesa ( $w=0,07$ ; vrednosti alternativa se subjektivno, lingvistički izražavaju),  $K_5$  – stanje slobodnih skladišnih kapaciteta u luci ( $w=0,03$ ; lingvistički se vrednuju uslovi u kojima će se skladištiti sredstva u luci),  $K_6$  – pogodnost skladištenja tereta u luci ( $w=0,03$ ; lingvistički se vrednuju mogućnosti, troškovi i administrativni aspekti skladištenja),  $K_7$  – zastupljenost tzv. dodatnih usluga ( $w=0,06$ ; lingvistički se vrednuje kvalitet usluga praćenja tereta i informisanja pošiljaoca, pomoć u ispunjavanju administrativnih procedura, odnos prema žalbama klijenata, odnos prema specifičnim zahtevima klijenata itd.),  $K_8$  – posebni uslovi i naknade za skladištenje i manipulisanje opasnim teretom ( $w=0,08$ ; lingvistički se vrednuju složenost specifičnih administrativnih procedura i posebne naknade),  $K_9$  – bezbednost tereta i pratećeg osoblja ( $w=0,10$ ; lingvistički se vrednuju zaštićenost tereta od krađe, oštećenja ili gubitaka tokom transfera kroz luku, kao i bezbednost osoblja),  $K_{10}$  – troškovi lučkih taksi ( $w=0,12$ ; *odnosi se na osnovne naknade za transfer tereta kroz luku* - reprezent ovih troškova je taksa za transfer jednog standardnog kontejnera 20 ft - stopa kroz lučki terminal, izražena u monetarnoj jedinici *evro - EUR* ili *dolar - USD*) i  $K_{11}$  raznovrsnost ponude prevozne usluge u luci ( $w=0,04$ ; lingvistički se vrednuje mogućnostima izbora logističkih provajdera i cena usluge prevoza u posmatranoj luci).

Evidentno je da se deo kriterijuma može kvantifikovati, dok je drugi deo kriterijuma odlučivanja kvalitativnog karaktera, pa su oni vrednovani lingvističkim izrazima sa fazifikovane Likert-ove skale [3], u kojoj su lingvistički izrazi predstavljeni trouglastim fuzzy brojevima prema sledećem: veoma dobro (VD = (4.5, 5, 5)), dobro (D = (3.5, 4, 4.5)), skromno (S = (2.5, 3, 3.5)), loše (L = (1.5, 2, 2.5)) i veoma loše (VL = (1, 1, 1)).

Defazifikacija lingvističkih izraza, odnosno njihovih fuzzy ekvivalenata izvršena je prema izrazu (1), [9], gde su  $a^{(l)}$  i  $a^{(u)}$  leva i desna granica (respektivno) distribucije intervala poverenja trougaonog fuzzy broja, a  $a^{(m)}$  vrednost u kojoj trougaona funkcija pripadnosti dostiže maksimum.

$$\text{defuzzy}A = \frac{(a^{(l)} + 4 \cdot a^{(m)} + a^{(u)})}{6} \quad (1)$$

U konkretnom primeru, pretpostavka je da su sredstva za transport locirana u Nišu, a da je krajnje određište luka Bejrut (Liban). Pitanje formulisanja alternativa odnosi se na uslov da svaka od alternativnih morskih luka mora biti iz R. Srbije dostupna železnicom. U tom smislu, kao moguće alternative nameću se luke: Bar (A1), Solun (A2), Drač (A3), Ploče (A4), Varna (A5), Konstanca (A6), Burgas (A7), i Pirej (A8).

Što se tiče izbora metoda VKO za rangiranje izdvojenih alternativa, odabrane su neke od jednostavnijih i najčešće korišćenih metoda VKO – metode SAW, TOPSIS, COPRAS i VIKOR.

### 3. ANALIZA PRIMENE VKO METODA NA PROBLEM IZBORA LUKE

Matrica odlučivanja za inicijalno rangiranje morskih luka, formirana na osnovu dostupnih podataka i subjektivnih preferencija baziranih na tim podacima i prikupljenim informacijama o lukama, prikazana je u tabeli 1.

**Tabela 1:** Matrica odlučivanja pri izboru morske luke

Kriterijum (K <sub>i</sub> )	Alternativne morske luke							
	A1 Bar	A2 Solun	A3 Drač	A4 Ploče	A5 Varna	A6 Konstanca	A7 Burgas	A8 Pirej
K <sub>1</sub> [–] (max)	4,12	4,05	3,19	2,85	3,41	1,77	4,06	1,39
K <sub>2</sub> [km] (min)	2891	2312	2849	3519	3042	3732	2896	2500
K <sub>3</sub> [–] (max)	0,277	0,379	0,208	0,223	0,263	0,697	0,078	0,183
K <sub>4</sub> [lingv] (max)	D	VD	D	L	D	VD	D	VD
K <sub>5</sub> [lingv] (max)	D	D	D	D	S	D	D	VD
K <sub>6</sub> [lingv] (max)	VD	S	S	L	D	D	D	S
K <sub>7</sub> [lingv] (max)	S	D	L	L	D	VD	S	VD
K <sub>8</sub> [lingv] (max)	VD	S	S	D	S	L	S	L
K <sub>9</sub> [lingv] (max)	VD	S	VL	VL	D	D	D	S
K <sub>10</sub> [EUR] (min)	179	161	197	200	200	152	200	167
K <sub>11</sub> [lingv] (max)	S	VD	L	VL	D	VD	S	VD

\* [lingv] – lingvistički izraz iz fazifikovane Likert-ove skale

Primenom izabranih metoda VKO na matricu odlučivanja dobijeni se rangovi alternativa koji ukazuju na očigledna neslaganja u preferencijama alternativa, tabela 2.

**Tabela 2:** Rezultat rangiranja morskih luka različitim metodama VKO

Metoda VKO	Rang alternativa
SAW	A2 > A1 > A7 > A5 > A6 > A8 > A3 > A4
TOPSIS	A1 > A2 > A7 > A5 > A6 > A3 > A4 > A8
COPRAS	A1 > A2 > A5 > A6 > A7 > A8 > A3 > A4
VIKOR	A2 > A1 > A7 > A5 > A3 > A6 > A4 > A8

Oznaka  $A_i > A_j$  predstavlja izraz: alternativa  $A_i$  je bolje rangirana od alternative  $A_j$ .

U drugom koraku analize izvršena je promena ulaznih parametara (težinskih koeficijenata kriterijuma) i ponovo su utvrđeni rangovi. U tu svrhu definisano je sedam scenarija sa različitim težinskim koeficijentima, zavisno od toga da li se prednost u značaju kriterijuma daje grupi kriterijuma rute, grupi kriterijuma usluge ili grupi kriterijuma troškova.

Prva dva scenarija predstavljaju slučajeve u kojima su dominantni kriterijumi rute transporta (K<sub>1</sub> i K<sub>2</sub>). U oba scenarija suma težinskih koeficijenata tih kriterijuma iznosi 0,6, dok sume težinskih koeficijenata kriterijuma usluge (od K<sub>3</sub> do K<sub>9</sub>) iznose 0,3 u prvom, odnosno 0,1 u drugom scenariju, kao što i suma težinskih koeficijenata za kriterijume troškova (K<sub>10</sub> i K<sub>11</sub>) iznosi 0,1 za prvi scenario i 0,3 za drugi scenario). Scenario 3 i scenario 4 predstavljaju slučajeve u kojima su dominantni kriterijumi usluge (sume težinskih koeficijenata za kriterijume usluge su 0,6, dok sume težinskih koeficijenata za kriterijume rute i kriterijume troškova

naizмениčno uzimaju vrednosti 0,1 i 0,3). U scenariju 5 i scenariju 6 kriterijumima troškova je dat najveći značaj (sume njihovih težinskih koeficijenata su 0,6, a sume ostale dve grupe kriterijuma naizмениčno uzimaju vrednosti 0,1 i 0,3). U poslednjem scenariju (scenario 7) nema favorizovanih kriterijuma (svim kriterijumima dodeljena je ista vrednost težinskog koeficijenta).

Promene težinskih koeficijenata kriterijuma, prema navedenim scenarijima, rezultovale su promenama rangova alternativa u svakoj od primenjenih metoda VKO (tabela 3).

**Tabela 3:** Promene u rangovima alternativa prema scenarijima promene težinskih koeficijenata kriterijuma

Težinski koef.	Rang alternative (alternative koje su promenile rang, u odnosu na početni, označene su crvenim slovima)	
	SAW	TOPSIS
Početni	A2 > A1 > A7 > A8 > A4 > A5 > A3 > A6	A1 > A2 > A6 > A7 > A4 > A5 > A3 > A8
Scenario 1	A2 > A1 > A5 > A8 > A4 > A7 > A3 > A6	A2 > A1 > A5 > A6 > A4 > A8 > A3 > A7
Scenario 2	A2 > A1 > A6 > A8 > A4 > A7 > A3 > A5	A2 > A1 > A5 > A8 > A4 > A7 > A3 > A6
Scenario 3	A1 > A2 > A7 > A8 > A4 > A3 > A5 > A6	A1 > A2 > A7 > A8 > A4 > A3 > A5 > A6
Scenario 4	A3 > A2 > A7 > A8 > A5 > A1 > A6 > A4	A3 > A2 > A7 > A8 > A4 > A1 > A6 > A5
Scenario 5	A4 > A1 > A7 > A8 > A5 > A2 > A6 > A3	A5 > A1 > A7 > A8 > A4 > A2 > A6 > A3
Scenario 6	A4 > A2 > A7 > A8 > A5 > A1 > A6 > A3	A5 > A2 > A7 > A8 > A4 > A1 > A6 > A3
Scenario 7	A3 > A1 > A7 > A8 > A5 > A2 > A6 > A4	A3 > A2 > A7 > A8 > A4 > A1 > A6 > A5
	COPRAS	VIKOR
Početni	A1 > A2 > A7 > A8 > A3 > A4 > A5 > A6	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 1	A2 > A1 > A5 > A8 > A4 > A6 > A3 > A7	A2 > A1 > A4 > A7 > A5 > A8 > A3 > A6
Scenario 2	A2 > A1 > A6 > A8 > A4 > A7 > A3 > A5	A2 > A1 > A5 > A8 > A4 > A7 > A3 > A6
Scenario 3	A1 > A2 > A7 > A8 > A4 > A3 > A5 > A6	A2 > A1 > A5 > A8 > A4 > A7 > A3 > A6
Scenario 4	A3 > A2 > A7 > A8 > A5 > A1 > A6 > A4	A3 > A1 > A7 > A8 > A5 > A2 > A6 > A4
Scenario 5	A4 > A1 > A7 > A8 > A5 > A2 > A6 > A3	A4 > A1 > A7 > A8 > A5 > A3 > A6 > A2
Scenario 6	A4 > A2 > A7 > A8 > A5 > A1 > A6 > A3	A4 > A2 > A7 > A8 > A5 > A1 > A6 > A3
Scenario 7	A2 > A3 > A7 > A8 > A4 > A1 > A6 > A5	A2 > A1 > A7 > A8 > A5 > A3 > A6 > A4

Nakon analize promene u rezultatima rangiranja usled promene težinskih koeficijenata kriterijuma izvršena je analiza promena rezultata rangiranja izazvanih izmenom merne skale kojom izražavaju vrednosti alternativa. U tu svrhu, generisana su dva scenarija. U prvom scenariju, vrednosti alternativa po dva najznačajnija kriterijuma među kriterijumima lingvističkog karaktera: *posebni uslovi i nadoknade za skladištenje i manipulisanje opasnim teretom* (kriterijum  $K_8$ ,  $w = 0,08$ ) i *bezbednost tereta i pratećeg osoblja* (kriterijum  $K_9$ ,  $w = 0,10$ ), izražene su na mernoj skali: veoma dobro (VD = (8, 9, 9)), dobro (D = (6, 7, 8)), skromno (S = (4, 5, 6)), loše (L = (2, 3, 4)) i veoma loše (VL = (1, 1, 1)), koja je sa prethodno upotrebljenom fazifikovanom Likert-ovom skalom, povezana pozitivnom afinom transformacijom:  $y = 2x - 1$  [2]. U drugom scenariju vrednosti alternativa po kriterijumu *dužina ukupnog puta transporta* ( $K_2$ ) izražena je u nautičkim miljama – nm, umesto u km (1 km = 0,54 nm), a vrednost po kriterijumu *visina lučkih такси* ( $K_{10}$ ) izražena u USD (umesto u EUR, za odnos vrednosti EUR:USD, uzet je 1 EUR = 1,11669 USD), dok je način izražavanja vrednosti po ostalim kriterijumima ostao nepromenjen u odnosu na postupak utvrđivanja inicijalnih rangova.

Promene u mernim skalama, prema navedenim scenarijima, uglavnom nisu dovele do promenama u rangovima alternativa, osim u dva slučaja – u jednom scenariju kod metode SAW i jednom scenariju kod metode COPRAS, pri čemu je značajnija promena kod metode SAW jer se radi o promenama ranga kod prvorangiranih alternativa (tabela 4).

**Tabela 4:** Promene u rangovima alternativa prema scenarijima promene mernih skala kriterijuma

Merna skala	Rang alternativa	
	SAW	TOPSIS
Početna	A2 > A1 > A7 > A8 > A4 > A5 > A3 > A6	A1 > A2 > A6 > A7 > A4 > A5 > A3 > A8
Scenario 1	A1 > A2 > A7 > A8 > A4 > A5 > A3 > A6	A1 > A2 > A6 > A7 > A4 > A5 > A3 > A8
Scenario 2	A2 > A1 > A7 > A8 > A4 > A5 > A3 > A6	A1 > A2 > A6 > A7 > A4 > A5 > A3 > A8
	COPRAS	VIKOR
Početna	A1 > A2 > A7 > A8 > A3 > A4 > A5 > A6	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 1	A1 > A2 > A7 > A8 > A3 > A5 > A4 > A6	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 2	A1 > A2 > A7 > A8 > A3 > A4 > A5 > A6	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8

S obzirom da su u modelu odlučivanja prisutni i kriterijumi prihodnog i kriterijumi rashodnog karaktera, sprovedena je i analiza promene rezultata rangiranja na promenu načina formulacije kriterijuma. Na primer, kriterijum  $K_1$  (lokacijska pogodnost luke) može se prikazati i u formi rashodnog kriterijuma, kao lokacijska nepogodnost luke. Prihodna ( $X_j^+$ ) i rashodna ( $X_j^-$ ) formulacija ovog kriterijuma povezane su funkcijom oblika  $X_j^- = 1/X_j^+$ , pri čemu je poželjno da vrednost alternative po ovom kriterijumu bude što manja ukoliko je kriterijum rashodnog karaktera. Tako su izabrana tri kriterijuma prihodnog karaktera ( $K_1$ ,  $K_3$  i  $K_9$ ), čiji je način formulacije menjan kroz sedam scenarija (formirani na bazi različitih kombinacija načina formulacije ova tri kriterijuma), tabela 5.

**Tabela 5:** Scenariji sa različitim načinima formulacije kriterijuma

Scenario	Način formulacije kriterijuma		
	$K_1$ - lokacijska (ne)pogodnost luke	$K_3$ - operativna (ne)efikasnost luke	$K_9$ - (ne)bezbednost tereta i pratećeg osoblja
Početna formulacija	$X^+$	$X^+$	$X^+$
Scenario 1	$X^-$	$X^+$	$X^+$
Scenario 2	$X^+$	$X^-$	$X^+$
Scenario 3	$X^+$	$X^+$	$X^-$
Scenario 4	$X^-$	$X^-$	$X^+$
Scenario 5	$X^-$	$X^+$	$X^-$
Scenario 6	$X^+$	$X^-$	$X^-$
Scenario 7	$X^-$	$X^-$	$X^-$

Rezultati rangiranja alternativnih morskih luka, izabranim metodama VKO, prema ovim scenarijima, prikazani su u tabeli 6. Može se videti da su konzistentni na promene u načinu formulacije kriterijuma ostali jedino rangovi dobijeni primenom VIKOR metode, mada treba uočiti i da je TOPSIS metoda pokazala stabilnost kada se radi o odnosu dve prvorangirane alternative.

**Tabela 6:** Promene u rangovima alternativa prema scenarijima promene načina formulacije kriterijuma

Način formulacije	Rang alternativa	
	SAW	TOPSIS
Početni	A2 > A1 > A7 > A8 > A4 > A5 > A3 > A6	A1 > A2 > A6 > A7 > A4 > A5 > A3 > A8
Scenario 1	A2 > A1 > A7 > A8 > A4 > A5 > A3 > A6	A1 > A2 > A5 > A6 > A3 > A7 > A4 > A8
Scenario 2	A2 > A1 > A7 > A8 > A4 > A5 > A3 > A6	A1 > A2 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 3	A1 > A2 > A6 > A8 > A4 > A5 > A3 > A7	A1 > A2 > A6 > A7 > A4 > A5 > A3 > A8
Scenario 4	A2 > A1 > A7 > A8 > A4 > A5 > A3 > A6	A1 > A2 > A5 > A6 > A4 > A7 > A3 > A8
Scenario 5	A1 > A2 > A6 > A8 > A4 > A5 > A3 > A7	A1 > A2 > A5 > A6 > A3 > A7 > A4 > A8
Scenario 6	A1 > A2 > A6 > A8 > A4 > A5 > A3 > A7	A1 > A2 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 7	A1 > A2 > A6 > A8 > A4 > A5 > A3 > A7	A1 > A2 > A5 > A6 > A4 > A7 > A3 > A8
	COPRAS	VIKOR
Početni	A1 > A2 > A7 > A8 > A3 > A4 > A5 > A6	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 1	A2 > A1 > A6 > A8 > A3 > A5 > A4 > A7	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 2	A1 > A2 > A7 > A8 > A4 > A5 > A3 > A6	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 3	A1 > A2 > A7 > A8 > A3 > A5 > A4 > A6	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 4	A2 > A1 > A6 > A8 > A4 > A5 > A3 > A7	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 5	A1 > A2 > A6 > A8 > A3 > A5 > A4 > A7	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 6	A1 > A2 > A7 > A8 > A4 > A5 > A3 > A6	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8
Scenario 7	A1 > A2 > A6 > A8 > A4 > A5 > A3 > A7	A2 > A1 > A5 > A7 > A4 > A6 > A3 > A8

#### 4. ZAKLJUČAK

Evidentno je da su sve primenjene metode manje ili više osetljive na promenu ulaznih veličina. Međutim, niti veća stabilnost rezultata nedvosmisleno govori u prilog pouzdanosti metode, niti veće promene u rezultatima upućuju na nepreciznost metodološkog postupka. Usled toga je, prema [1], potrebno više pažnje treba posvetiti izboru relevantnih kriterijuma i alternativa, nego izboru same metode. Svaka metoda ima svoj skup ograničenja [12] i donosilac odluke mora biti svestan i prednosti i mana primenjene metode prilikom prihvatanja rezultata [4]. Neke smernice ([5, 10]) ukazuju da pri izboru adekvatne metode VKO treba imati u vidu: usklađenost

metode sa prirodom problema odlučivanja (mogućnost modeliranja strukture problema logičkim principima metode), konzistentnost matematičkog aparata metode i bliskost sa logikom odlučivanja čoveka, transparentnost uzročno-posledičnih veza između ulaznih i izlaznih podataka (mogućnost argumentovanja odluka), usaglašenost prirode i kvaliteta zahtevanih podataka sa značajem atributa odlučivanja i značajem problema u celini, mogućnost korekcije ulaznih parametara modela odlučivanja u toku procesuiranja podataka (u toku primene metode), osetljivost rezultata metode na promene ulaznih parametara, mogućnost upotrebe zavisnih i nelinearnih kriterijuma/atributa, način tretiranja neodređenosti podataka i neizvesnosti u problemima odlučivanja, mogućnost kombinovanja metode sa drugim metodama i tehnikama, jednostavnost primene, zahtevane vremenske i druge resursi za primenu metode, pogodnost metode za programiranje i dostupnost softverske podrške u primeni metode, pogodnost metode za grupne oblike odlučivanja i participaciju interesa stejkholdera i dr.

Izbor metode treba da bude rezultat kompromisa između opšteg cilja odlučivanja, prilagodljivosti matematičke podloge metode procesu odlučivanja, pretpostavkama i dostupnim podacima u konkretnom problemu odlučivanja, kao i eventualnim drugim operativnim ograničenjima u datoj situaciji (raspoloživo vreme, finansijska sredstva, kadrovski resursi), pri čemu se ne smeju zanemariti ni preferencije i bliskost donosioca odluke sa metodama koje se razmatraju kao pogodan alat u odlučivanju.

## LITERATURA

- [1] Athawale, V.M., Chakraborty, S. (2012). Material selection using multi-criteria decision-making methods: A comparative study, *Proceedings of Institution of Mechanical Engineers - Part L, Journal of Materials Design and Applications*, 226(4), 266–285.
- [2] Bach, E., Bridy, A. (2013). On the number of distinct functional graphs of affine-linear transformations over finite fields, *Linear Algebra and its Applications*, 439, 1312–1320.
- [3] Camparo, J. (2013). A geometrical approach to the ordinal data of Likert scaling and attitude measurements: The density matrix in psychology, *Journal of Mathematical Psychology*, 57, 29–42.
- [4] Chauhan, A., Vaish, R. (2014). A Comparative Study on Decision Making Methods with Interval Data, *Journal of Computational Engineering*, Vol 2014, article ID 793074.
- [5] De Montis, A., De Toro, P., Droste-Franke, B., Omann, I., Stagl, S. (2000). Criteria for quality assessment of MCDA methods, *3<sup>rd</sup> Biennial Conference of the European Society for Ecological Economics*, Vienna, Austria.
- [6] D'Este, G.M., Meyrick, S. (1992). Carrier selection in a RO/RO ferry trade, Part 1 - decision factors and attitudes, *Maritime Policy and Management*, 19(2), 115–126.
- [7] French, S. (1988). *Decision theory - an introduction to the mathematics of rationality*, Ellis Horwood Ltd., Chichester, UK.
- [8] Kahneman, D., Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47(2), 263–292.
- [9] Lubojević, S. (2016). Model odlučivanja organa saobraćajne službe u zadacima strategijskog transporta, doktorska disertacija, Univerzitet odbrane u Beogradu, Beograd.
- [10] *Multi-criteria analysis: A manual* (2009). Department for Communities and Local Government, London, UK.
- [11] Niavis, S., Tsekeris, T. (2012). Ranking and causes of inefficiency of container seaports in South-Eastern Europe, *European Transport Research Review*, 4(4), 235–244.
- [12] Shanian, A., Savadogo, O. (2009). A methodological concept for material selection of highly sensitive components based on multiple criteria decision analysis, *Expert Systems with Applications*, 36, 1362–1370.
- [13] Steven, A.B., Corsi, T.M. (2012). Choosing a port: An analysis of containerized imports into the US, *Transportation Research Part E: Logistics and Transportation Review*, 48(4), 881–895.
- [14] Triantaphyllou, E., Mann, S.H. (1989). An Examination of the Effectiveness of Multi-Dimensional Decision-Making Methods: A Decision-Making Paradox, *Decision Support Systems*, 5, 303–312.



# RUDARSTVO, GEOLOGIJA I ENERGETIKA

---

## MINING, GEOLOGY AND ENERGETICS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business







## RANKING OPEN-PIT MINES USING TYPE-1 CRITERION INTERVAL EXTENSION IN EDAS++ AS A COMPARATIVE MCDM METHOD

ŽELJKO PRAŠTALO<sup>1</sup>, ANĐELKA ŠTILIĆ<sup>2</sup>, STEVAN ĆORLUKA<sup>3</sup>, PAVLE STJEPANOVIĆ<sup>4</sup>

<sup>1</sup> Mining Institute Ltd. Belgrade, zeljko.prastalo@ribeograd.ac.rs

<sup>2</sup> Academy of Applied Studies Belgrade, andjelka.stilic@gmail.com

<sup>3</sup> Mining Institute Ltd. Belgrade, stevan.corluka@ribeograd.ac.rs

<sup>4</sup> Mining Institute Ltd. Belgrade, pavle.stjepanovic@ribeograd.ac.rs

**Abstract:** Due to its wide application, multi-criteria decision-making has easily found its way to the mining industry as well. Mining is particularly interesting from the multi-criteria analysis aspect, as mining projects require analyzing numerous factors affecting the exploitation process. However, multi-criteria problems are among the poorly structured decision-making models, given they take place in stochastic conditions. This primarily refers to the number of criteria and the complexity of their mutual relationships, which can be characterized by a complete contradiction, and the methodologically diverse space for determining preferences or weight factors that significantly affect decision-making results. In this paper, the use of EDAS++, normalized matrix data values, and interval criteria, presented through a mathematical model and example, allows ranking in conditions where some of the criteria have an optimal value interval of attributes that are equally relevant and therefore invariant with respect to decision. EDAS++ was used as a comparative method in analyzing open-pit limestone mines to obtain the broadest possible range of ranking methods while working on a doctoral dissertation entitled "Decision-making in case of different mining project solutions by comparative single and multi-criteria modeling."

**Keywords:** limestone, open-pit limestone mine, exploitable mineral raw material, supply chain, multi-criteria decision-making

### 1. INTRODUCTION

As in other industries, multi-criteria decision-making and analysis are not foreign to mining. Different multi-criteria methods and their different applications in works dealing with open-pit mines can be found in [1-12] [13], and others' work.

Ranking-based decision-making in the process of finding an acceptable solution by multi-criteria analysis was applied in the preparation of the doctoral dissertation entitled "Decision-making in case of different mining project solutions by comparative single and multi-criteria modeling." In addition to the existing, traditional multi-criteria analysis methods, the open-pit limestone mines' ability to participate in the process of flue gas desulphurization and thermal power plants has instigated the use of the EDAS++ method.

The need for evaluation and comparison aimed at determining the extent to which each alternative is trustworthy, preferable, and useful for the decision-maker requires observing past and present data, needs and availability through alternative options to be used in the future.

Numerous authors refer to the term "multi-criteria decision-making" as making decisions based on multiple criteria that are often contradictory. [14] state that each problem has four common characteristics: multiple criteria, appearance of conflicting criteria, immeasurable units, and problem solution [13].

### 2. MODEL FORMING INPUTS

The analyses of all limestone resources that have the characteristics necessary for this process have been taken from the above dissertation as continuous data in multi-year work on finding and verifying the necessary input data. The research directed the analysis towards the possibility of defining decision-making by comparing different multi-criteria analysis methods against the original desire to reach an optimal solution for supplying thermal power plants with limestone.

The thermal power plants' requirement for the necessary characteristics of the input raw material in the FGD (flue gas desulphurization) process has largely defined the decision-making criteria and their individual impact. According to the Study [15,16] on the possibility of supplying limestone for the needs of flue gas desulphurization in TPP Kostolac, TPP Nikola Tesla, and new thermal capacities, deposit classification parameters can be divided into three groups by the relevance of elements, as presented in Table 1.

**Table 1:** Deposit classification criteria

Type of criteria	Border of criteria
<i>Elimination criteria</i>	CaCO <sub>3</sub> content Distance > 150 km Lack of owner interest due to their own needs Undefined company status
<i>Economic criteria</i>	Transport conditions (road, rail, river transport) Deposit distance from the thermal power plant Deposit profitability
<i>Reliability criteria</i>	Capacity to supply contracted quantities Possibility for logistic support Quality, environmental protection, and occupational health and safety system

The reliability criterion was tested by considering the legal operations of existing mines, i.e., whether the existing mines have all the necessary permits and verified reserves obtained in a legally acceptable manner. This segment is defined by a simple request made to the relevant ministry and obtaining a list of exploitation permits. These certificates show the quantities of verified reserves to consider, which are an essential ranking criterion, along with mines' willingness to participate in this process. The open-pit mine capacity and mine lifespan are given in tons per year and projected years, respectively.

The annual capacities of open-pit mines and the selling price of mineral raw materials were obtained through direct contact with mine management and review of existing projects. The selling price and transport costs are presented in RSD per ton.

Real data for the price costs of individual modes of transport are presented through market analysis. Note that for obtaining the necessary multi-criteria ranking model, the data defined in 2014 were used as an initial example for continuous attribute values processing.

The influence of the above criteria on the decision selecting an open-pit limestone mine was obtained based on the analyses from the Study [15] the final result of which is given in Table 2.

**Table 2:** Influence of criteria on open-pit limestone mine selection

Type of criteria	Influence
<i>Open-pit mine capacity</i>	0.10
<i>Mine lifespan</i>	0.05
<i>Quantity of calcium carbonate</i>	0.25
<i>Selling price</i>	0.30
<i>Transport costs</i>	0.30

### 3. EDAS++ METHOD TEST EXPERIMENT

In this paper, one of the recent multi-criteria decision-making methods, EDAS, was chosen to select an open-pit mine that would serve the future Kolubara B thermal power plant. This method aggregated values obtained by adding up the positive distance from the average value and the values obtained by adding up the negative distance from the average value [17]. The original, EDAS method author, Mehdi Keshavarz-Ghorabae, suggests that the method can provide both optimistic and pessimistic solutions that give decision-makers flexibility in the final evaluation. In addition, this method and its extensions provide insight into the input data structure throughout the calculation process and monitor their transformations, including ranking [17-19].

As selecting the open-pit mine allowed for five criteria, some with smaller, some with a larger range of attribute values by criteria, and one of the criteria is invariant if the basic requirement is met, the EDAS++ extension of the EDAS method was applied. Type-1 criterion interval was put to use on the first criterion, wherein the emphasis is on the interval within which all criterion function values are invariant [19].

#### 3.1. Mathematical model

**Step 1.** Defining key criteria, weighting factors for criteria and alternatives to solving the problem of multi-criteria decision making.

**Step 2.** Forming a decision making matrix  $X = [X_{ij}]_{n \times m} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1m} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2m} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3m} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{nm} \end{bmatrix}$  (1)

**Step 3.** Mapping matrix X to matrix Q:

$$x_{i,j} \rightarrow q_{i,j} \text{ so that } q_{i,j} = x_{i,j} \tag{2}$$

$$\text{where for } x_{i,j} \in [a, b] \rightarrow q_{i,j} = b, \text{ for } x_{i,j} > b \rightarrow q_{i,j} = m, \text{ for } x_{i,j} < a \rightarrow q_{i,j} = x_{i,j}. \tag{3}$$

For the remaining classic benefit and non-benefit criteria, mapping is performed as follows:

$$\text{the value of the attribute } x_{i,j} \text{ is mapped to } q_{i,j} \text{ so that } q_{i,j} = x_{i,j} \tag{4}$$

**Step 4.** Normalization of the decision-making matrix, where a normalized decision matrix is obtained based on the mapped matrix [17-19]:

$$r_{ij} = \frac{q_{ij} - q_j^-}{q_j^+ - q_j^-} \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n \tag{5}$$

$$R = [R_{ij}]_{n \times m} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \dots & r_{1m} \\ r_{21} & r_{22} & r_{23} & \dots & r_{2m} \\ r_{31} & r_{32} & r_{33} & \dots & r_{3m} \\ \dots & \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & r_{n3} & \dots & r_{nm} \end{bmatrix} \tag{6}$$

**Step 5.** Determining the average solution for each criterion separately

$$AV = [AV_j]_{1 \times m} = \left[ \frac{\sum_{i=1}^n x_{i1}}{n}, \frac{\sum_{i=1}^n x_{i2}}{n}, \dots, \frac{\sum_{i=1}^n x_{im}}{n} \right] = [x_1^*, x_2^*, \dots, x_m^*] \tag{7}$$

**Step 6.** Forming a matrix  $[PDA_{ij}]_{n \times m}$  where the matrix elements represent a positive distance from  $[AV_j]_{1 \times m}$  and matrix  $[NDA_{ij}]_{n \times m}$  where the matrix elements represent a negative distance from  $[AV_j]_{1 \times m}$ . The elements of matrices  $[PDA_{ij}]_{n \times m}$  i  $[NDA_{ij}]_{n \times m}$  are calculated depending on whether the criteria are benefit- 'income' or 'expenditure', as follows:

$$d_{ij}^+ = \begin{cases} \frac{\max(0, (x_{ij} - x_j^*))}{x_j^*} & ; j \in \Omega_{max} \\ \frac{\max(0, (x_j^* - x_{ij}))}{x_j^*} & ; j \in \Omega_{min} \end{cases} \tag{8}$$

$$d_{ij}^- = \begin{cases} \frac{\max(0, (x_j^* - x_{ij}))}{x_j^*} & ; j \in \Omega_{max} \\ \frac{\max(0, (x_{ij} - x_j^*))}{x_j^*} & ; j \in \Omega_{min} \end{cases} \tag{9}$$

where  $\Omega_{max}$  is a group of 'income' criteria and  $\Omega_{min}$  is a group of 'expenditure' criteria.

**Step 7.** Forming one-dimensional matrices of weighted sums  $[SP_i]_{n \times 1}$  and  $[SN_i]_{n \times 1}$  in which elements are the result of multiplication of matrices  $[PDA_{ij}]_{n \times m}$  and  $[NDA_{ij}]_{n \times m}$  with one-dimensional matrix of weighting factors  $[W]_{m \times 1}$ .

**Step 8.** Normalization of one-dimensional matrices  $[SP_i]_{n \times 1}$  and  $[SN_i]_{n \times 1}$ :

$$[NSP_i]_{n \times 1} = \frac{[SP_i]_{n \times 1}}{(\max[SP_i]_{n \times 1})} = \frac{1}{(\max[SP_i]_{n \times 1})} * [SP_i]_{n \times 1} \tag{10}$$

$$[NSN_i]_{n \times 1} = 1 - \frac{[SN_i]_{n \times 1}}{(\max[SN_i]_{n \times 1})} = 1 - \frac{1}{(\max[SN_i]_{n \times 1})} * [SN_i]_{n \times 1} \tag{11}$$

**Step 9.** Forming a one-dimensional matrix of calculating assessment  $[AS_i]_{nx1}$  the elements:

$$as_i = \frac{1}{2}(nsp_i + nsn_i) \quad (12)$$

where:  $as_i$  the elements of  $[AS_i]_{nx1}$ ,  $nsp_i$  are the elements of  $[NSP_i]_{nx1}$ , and  $nsn_i$  are the elements of  $[NSN_i]_{nx1}$ .

**Step 10.** Assessment of the calculations where  $0 \leq AS_i \leq 1$ . The alternative with the highest  $AS_i$  is the best ranked alternative [17-19].

### 3.2. Testing and results

The application of the EDAS++ method began with defining key criteria and the direction of the criterion function, weighting factors for the criteria and selection alternatives, and continued with forming the decision matrix [17-19].

Step 1 was to define key criteria, weighting factors for the criteria and alternatives for solving the multi-criteria analysis problem presented in Table 3.

**Table 3:** Baseline table

	Open-pit mine capacity	Open-pit mine lifespan	CaCO <sub>3</sub>	Costs	
				Selling price	Transport costs
	(t/yr)	(yr)	(%)	(RSD/t)	
Impact	↑	↑	↑	↓	↓
Criterion weight	$w_1=0.1$	$w_2=0.05$	$w_3=0.25$	$w_4=0.3$	$w_5=0.3$
MINE 1	350000	80.86	97.04	520	970
MINE 2	200000	66.5	99.57	520	313.43
MINE 3	50000	45.87	98.38	520	771.56
MINE 4	800000	12.39	98.5	560	585.8
MINE 5	1000000	9.7	97.54	570	771.56
MINE 6	150000	62.19	94.39	650	585.8
MINE 7	162000	44.15	97.97	580	931.17
MINE 8	300000	21.99	97	600	931.17
MINE 9	280000	62.07	95.19	570	628.97
MINE 10	300000	57.51	97.04	630	466.66
MINE 11	333000	145.77	97.29	660	731.86
MINE 12	300000	60	95.34	540	499.46
MINE 13	900000	59.67	94.96	630	568.16
MINE 14	750000	105.98	97.02	630	568.16
MINE 15	400000	64.11	96.06	530	1362.89

which was followed by forming a decision matrix.

Mapping the data from decision matrix X to decision matrix Q within Step 3, according to the interval type criteria rules [19], the interval Type-1 criterion used in this case for Criterion 1 (open-pit mine capacity), for where the interval is  $[a, b]$ :  $a = 75000$ ,  $b = 1000000$ , was performed as follows:

$$\text{for } x_{i,j} \in [a, b] \rightarrow q_{i,j} = b, \quad (13)$$

$$\text{for } x_{i,j} > b \rightarrow q_{i,j} = m, \quad (14)$$

$$\text{for } x_{i,j} < a \rightarrow q_{i,j} = x_{i,j} \quad [19]. \quad (15)$$

For the remaining classic benefit and non-benefit criteria, as well as for special unspecified criterion interval conditions, mapping was performed as follows: the value of the attribute  $x_{i,j}$  is mapped to  $q_{i,j}$  so that  $q_{i,j} = x_{i,j}$ . (16)

Step 4 implied normalization:

$$Q \rightarrow R; \quad r_{ij} = \frac{q_{ij} - q_j^-}{q_j^+ - q_j^-} \quad i = 1, 2 \dots m; j = 1, 2 \dots n \quad (17)$$

after which the procedure was continued according to the EDAS and established extensions of methods' steps.

The ranking results shown in Table 4, ranked using the EDAS++ method of multi-criteria decision-making, can also be viewed as MINE2 < MINE4 < MINE3 < MINE12 < MINE1 < MINE5 < MINE9 < MINE7 < MINE10 < MINE14 < MINE8 < MINE11 < MINE13 < MINE15 < MINE6 with the smallest ranking number meaning the best ranked open-pit mine.

**Table 4:** EDAS++ method test result

	$nsp_i$	$nsn_i$	$as_i = \frac{1}{2}(nsp_i + nsn_i)$	Ranking
MINE 1	0.3891320073	0.6733696129	0.5312508101	5
MINE 2	1	1	1	1
MINE 3	0.5094907464	0.7327624763	0.6211266113	3
MINE 4	0.4149636385	0.920180363	0.6675720007	2
MINE 5	0.1420861848	0.8406915774	0.4913888811	6
MINE 6	0.1185493908	0	0.05927469539	15
MINE 7	0.1348550882	0.6961329695	0.4154940288	8
MINE 8	0.02103179046	0.4976559978	0.2593438941	11
MINE 9	0.1414044495	0.7134692617	0.4274368556	7
MINE 10	0.2349625811	0.5844648889	0.409713735	9
MINE 11	0.1508040943	0.3136521895	0.2322281419	12
MINE 12	0.4175016657	0.7388110133	0.5781563395	4
MINE 13	0.1310134647	0.2627044036	0.1968589342	13
MINE 14	0.1981896276	0.5885078777	0.3933487526	10
MINE 15	0.2967238177	0.03556247814	0.1661431479	14

#### 4. CONCLUSION

Defining the optimal model for solving the mining problem of this paper, open-pit limestone mines ranking, requires analyzing several possible scenarios. In the process of analyzing the presented problem, the EDAS++ with interval criteria Type-1 extension was applied as one of the options for decision-making. This method was initially developed as a method for analyzing social and personal doubts, but in this case, it was applied to analyze a technical problem. The inputs are defined based on the most important criteria given in the Study [15] on the possibility of supplying limestone for the needs of flue gas desulphurization in TPP Kostolac, TPP Nikola Tesla, and new thermal capacities.

The best-ranked open-pit limestone mine was MINE2. The analysis results are within limits close to those of other methods, primarily the PROMETHEE method, so the EDAS++ method can be equally used for analyzing technical and technological tasks.

#### REFERENCES

[1] Vujić, S., Hudej, M., & Miljanović, I. (2013). Results of The Promethee Method Application in Selecting The Technological System at The Majdan III Open Pit Mine / Wyniki Zastosowania Metody Promethee Do Wyboru Systemu Technologicznego W Kopalni Odkrywkowej Majdan III. Archives of Mining Sciences, 58(4), 1229–1240. <https://doi.org/10.2478/amsc-2013-0084>

- [2] Dimitrijević, B., Vujic, S., Matic, I., Majjanac, S., Praštalo, J., Radosavljević, M., & Čolaković, V. (2014). Multi-criterion analysis of land reclamation methods at Klenovnik open pit mine, Kostolac coal basin. *Journal of Mining Science*, 50(2), 319–325. <https://doi.org/10.1134/s106273911402015x>
- [3] Nolan, T. A., & Keçojević, V. (2014). Selection of overburden surface mining method in West Virginia by analytical hierarchy process. *International Journal of Coal Science & Technology*, 1(3), 306–314. <https://doi.org/10.1007/s40789-014-0019-0>
- [4] Kizil, M. S., Abdalla, S., & Canbulat, I. (2014). Underground coal mine layout selection using analytical hierarchy process. *Mining Technology*, 123(1), 20–29. <https://doi.org/10.1179/1743286313y.0000000053>
- [5] Wang, C., Tu, S., Zhang, L., Yang, Q., & Tu, H. (2015). Auxiliary transportation mode in a fully-mechanized face in a nearly horizontal thin coal seam. *International Journal of Mining Science and Technology*, 25(6), 963–968. <https://doi.org/10.1016/j.ijmst.2015.09.013>
- [6] Baek, J., Choi, Y., & Park, H. S. (2016). Uncertainty Representation Method for Open Pit Optimization Results Due to Variation in Mineral Prices. *Minerals*, 6(1), 17. <https://doi.org/10.3390/min6010017>
- [7] Bouhedja, A., Idres, A., Boutrid, A., Bounouala, M., Benselhoub, A., & Talhi, K. (2016). Application of promethee mathematical model for choosing a secondary breakage process of the oversized blocks in limestone quarries. *Mining Science*, 23.
- [8] Yari, M., Bagherpour, R., & Almasi, N. (2016). An Approach To The Evaluation And Classification Of Dimensional Stone Quarries With An Emphasis On Safety Parameters. *Rudarsko-Geološko-Naftni Zbornik*, 31(1), 15–26. <https://doi.org/10.17794/rgn.2016.3.2>
- [9] Yakovlev, V. L., Zyryanov, I. V., Akishev, A. N., & Sakantsev, G. G. (2016). Determination of open pit diamond mine limits with regard to stripping time difference. *Journal of Mining Science*, 52(6), 1143–1149. <https://doi.org/10.1134/s1062739116061674>
- [10] Šubaranović, T., Vujić, S., Radosavljević, M., Dimitrijević, B., Ilić, S., & Krunic, D. J. (2019). Multi-Attribute Scenario Analysis of Protection of Drmno Open Pit Mine against Groundwater. *Journal of Mining Science*, 55(2), 280–286. <https://doi.org/10.1134/s1062739119025564>
- [11] Aryafar, A., Rahimdel, M. J., & Tavakkoli, E. (2020). Selection Of The Most Proper Drilling And Blasting Pattern By Using Madm Methods (A Case Study: Sangan Iron Ore Mine, Iran). *Rudarsko-Geološko-Naftni Zbornik*, 35(3), 97–108. <https://doi.org/10.17794/rgn.2020.3.10>
- [12] Yari, M., Bagherpour, R., Khoshouei, M., & Pedram, H. (2020). Investigating A Comprehensive Model For Evaluating Occupational And Environmental Risks Of Dimensional Stone Mining. *Rudarsko-Geološko-Naftni Zbornik*, 35(1), 101–109. <https://doi.org/10.17794/rgn.2020.1.8>
- [13] Farkaš, B., & Hrastov, A. (2021). Multi-Criteria Analysis for the Selection of the Optimal Mining Design Solution—A Case Study on Quarry "Tambura." *Energies*, 14(11), 3200. <https://doi.org/10.3390/en14113200>
- [14] Hwang, C.-L.; Yoon, K. *Multiple Attribute Decision Making; Lecture Notes in Economics and Mathematical Systems; Springer: Berlin/Heidelberg, Germany, 1981; Volume 186, ISBN 978-3-540-10558-9.*
- [15] Studija o mogućnosti snabdevanja krečnjakom za potrebe odsumporavanja dimnih gasova TE Kostolac, TE Nikola Tesla i novih termo kapaciteta (2014). Rudarski Institut d.o.o
- [16] Praštalo, Ž., Simović, I., Milošević, M. D. (2015). Uloga rudarstva u zaštiti životne sredine. 40. Naučna konferencija - Održavanje mašina i opreme, Budva, Montenegro. Institut za istraživanja i projektovanja u privredi. p. 478–486.
- [17] Štilić, A., Nicić, M., Zimonjić, B., & Njeguš, A. (2019). Primena višekriterijumske metode EDAS u rangiranju kandidata za rad u turističkoj privredi i uvođenje korektivnog koraka. *Turističko poslovanje*, (23), 61-75. <https://doi.org/10.5937/turpos0-21644>
- [18] Štilić, A., & Štilić, I. (2022). Selection of Exterior Wall System and MCDM Derived Decision. *Journal of Facade Design and Engineering*, 10(1), 1–28. <https://doi.org/10.47982/jfde.2022.1.01>
- [19] Štilić, A. (2020). Novel EDAS++ method: Interval type criteria and extension to EDAS. *Turisticko Poslovanje*, 25–26, 39–52. <https://doi.org/10.5937/turpos0-29612>

## THE SERBIAN CLIMATE CHANGE ACT: THE PERSPECTIVE OF IMPLEMENTATION OF CARBON CAPTURE AND STORAGE PROJECTS IN THE OIL AND GAS INDUSTRY

SERGUEI FOMINYKH<sup>1</sup>

<sup>1</sup> NIS j.s.c. Novi Sad [serguei.fominykh@emba.imd.org](mailto:serguei.fominykh@emba.imd.org)

**Abstract:** Serbia has a great potential for the implementation of carbon capture and storage projects in the oil and gas industry. The recently adopted Climate Change Act provides the stimulation mechanism named Clean Development Mechanism (CDM) that proved its efficiency under the Kyoto Protocol. CDM would motivate the major industry players, especially Naftna Industrija Srbije j.s.c. (NIS), to focus on carbon capture and storage (CCS) projects aiming to benefit from the reduction of CO<sub>2</sub> emissions and contribute to the overall sustainable development of Serbia. Currently, the Serbian Government has not enabled CDM implementation under the Climate Change Act via by-laws due to uncertainties in the world carbon markets. The CDM should be enabled in the nearest future, as any delay in deployment of CCS would negatively affect the Republic of Serbia's target to reduce its GHG emissions by 9.8% below 1990 levels by 2030, and NIS's goal to foster CCS technologies.

**Keywords:** Oil & Gas Industry, Carbon Capture and Storage, CO<sub>2</sub> emissions, Climate Change Act

### 1. INTRODUCTION

The purpose of this article is to analyse whether the Climate Change Act would motivate the major players for the execution of carbon capture and storage (CCS) projects in the oil and gas industry in Serbia, or it will not affect their strategies and policies regarding the carbon capture and storage practices. Carbon capture and storage is the process of removing CO<sub>2</sub> from industrial processes such as the exploration of oil and gas fields that burn fossil fuels, and CO<sub>2</sub> can either be taken out before combustion occurs or after the combustion occurs. When the CO<sub>2</sub> is removed, it is transported and placed in long-term storage that prevents its remission to the atmosphere.

More than a year ago, in March 2021, the Republic of Serbia adopted the Law on climate change, i.e., the Climate Change Act, fulfilling its obligations under the Paris Agreement and harmonizing its domestic legislation with the applicable EU regulations. This law introduced a system for reducing the greenhouse gas (GHG) emissions such as carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), nitrous oxide (N<sub>2</sub>O), fluorocarbons (HFCs), perfluorocarbons (PFCs), sulphur hexafluoride (SF<sub>6</sub>), and nitrogen trifluoride (NF<sub>3</sub>), all caused by human activities and industries, which has led to exposure to climate change not only of the population of the Republic of Serbia but also the world's population [17]. The system for reducing greenhouse gas (GHG) emissions should be a possible solution for the better and more complete fulfillment of the Paris Agreement and other acts and instructions adopted under the United Nations Framework Convention on Climate Change (UNFCCC).

According to the Republic of Serbia's Nationally Determined Contribution (NDC) under the UNFCCC, the country has pledged to reduce its GHG emissions by 9.8% below 1990 levels by 2030. In line with the available data for 2020, Serbia has produced 43.14 million tonnes of CO<sub>2</sub>, including 4.98 million tonnes from oil fuels and 8.45 million tonnes from gas fuels [14].

The major player in the Serbian oil and gas industry is Naftna Industrija Srbije j.s.c. (NIS j.s.c. Novi Sad), one of the largest vertically integrated oil and gas companies in Southeast Europe [11]. The company and its key production capacities are based in the Republic of Serbia, and the main activities of NIS j.s.c. Novi Sad are exploration, production, and refining of crude oil and natural gas, as well as sales of a wide range of petroleum products including both motor and vehicle fuels. According to the company's Sustainable Development Reports, NIS j.s.c. Novi Sad produced 1.627 million tonnes of CO<sub>2</sub> in 2020, and 1.69 million tonnes of CO<sub>2</sub> in 2021, where the share of the oil and gas production and exploration business was 21 percent or 0.35 million tonnes [10]. The data show that NIS j.s.c. could have a major role in both production and reduction of CO<sub>2</sub> emissions, which reminds us of the importance of studying possibilities of implementation of the new technologies such as carbon capture and storage.

Keeping this in mind, in the first part of the article the author gives a review of available literature concerning the implementation of carbon capture and storage projects in the oil and gas industry. In the second

part of the article, a review of the motivation mechanism available under the Climate Change Act is given. In the third part of the article, an analysis of potential benefits related to the deployment of CCS projects is conducted, and the environmental, economic, and social are identified and well explained. In the last part of the article, the author studies the potential of carbon capture and storage in oil and gas fields in Serbia and provides insights into the future of CCS projects in the Serbian oil and gas industry.

## 2. LITERATURE REVIEW

The recent implementation of the clean development mechanism in Eastern Europe was mostly analyzed in the studies of Casino, Collado, & Nassar S. (2022), and in the studies of Kainou (2022), as well. They all confirmed that the clean development mechanism stimulates the deployment of projects aiming to decrease GHG emissions [1, 6]. Accordingly, innovative technologies and the following transformation of the oil and gas industry are the keys to successful decarbonization and achieving sustainability goals, which is confirmed by Hills (2020) and Kyriakopoulos (2021) [4, 9].

The risks and benefits related to the deployment of CCS projects were recently analyzed by Torvanger (2019) who revealed at least three groups of the effects, such as environmental, economic, and social. The risks and benefits related to the deployment of CCS projects have also found their place in the special reports published by the International Energy Agency – IEA (2020), and the Intergovernmental Panel on Climate Change – IPCC (2005), which have recognized the potential of CCS technology as a mechanism for the GHG emission reduction.

In addition to the mentioned IEA and IPCC special reports, the official data from both UNDP Serbia and NIS j.s.c. Novi Sad sustainability development reports were analyzed to confirm the necessity of reduction of CO<sub>2</sub> emissions in the oil and gas industry in Serbia. The author has also analyzed the publicly available results of research and case studies on the long-term planning of carbon capture and storage in Serbia performed by Nešić & Karas (2019) of NIS Scientific and Technological Centre based on their accumulated experience in Russia and Serbia, as well as scholars Tomić, Karavić, and Danilovic (2021), from the Faculty of mining and geology, University of Belgrade, Serbia.

## 3. STIMULATION MECHANISM UNDER THE CLIMATE CHANGE ACT

The Serbian Climate Change Act provides the measures aimed at motivating the reduction of GHG emissions that can be implemented through programs and projects under the framework named the Clean Development Mechanism (CDM) set out in Article 15 of this Act. Initially, this mechanism was introduced under the Kyoto Protocol of December 11, 1997, and came into force in Serbia on January 7, 2008, but it was also included in the Law on Air Protection, which is considered one of the first endeavors in Serbia in the domain of climate change and GHG emission reduction.

The CDM is one of the three mechanisms under the Kyoto Protocol, providing the world's first international carbon finance scheme which allows emission-reduction projects to earn certified emission reduction (CER) credits, each equivalent to one tonne of CO<sub>2</sub> [3]. Companies could acquire tradeable certified emission reduction credits by investing in energy conservation and new energy projects to meet a part of their emission targets during the first committed period from 2008 to 2012. A political decision was made to set no numerical target for the second commitment period from 2012 under the Kyoto Protocol [1].

Certified emission reduction credits issued under the CDM have been traded and priced in major commodities exchanges around the world. Thus, in the first year of the first commitment period (2008–2012), the credit price hit a record high of €25 per tonne of CO<sub>2</sub>, but in 2012 the credit price crashed, falling to €0.5 per tonne of CO<sub>2</sub> [6]. Eventually, the market for certified emission reduction credits tumbled into a 'carbon panic' and the pricing mechanism broke down [6]. Despite the recent revival of carbon markets, the credit prices under CDM carbon financing schemes remain exposed to high volatility and uncertainty, and there are no clear predictions as to whether this mechanism will achieve its primary objective.

Regarding the concrete efforts in Serbia, there are seven CDM projects related to the first commitment period from 2008 to 2012 registered under Kyoto Protocol in Serbia, all listed below [20]:

- Alibunar Biogas Plant Construction Project;
- LGF Recovery at the Bubanj Landfill Site;
- Reduction of Methane Leakages in the Gas Distribution Networks;
- Wind Farm Cibuk 1;
- Wind Farm Kladovo 1;
- Wind Farm Kosava I+II; and
- Wind Farm Plandiste 1.



At the same time, the scope of the CDM practical application in Serbia is unclear due to the following reasons:

- targets for the second commitment period under the Kyoto Protocol are not defined and did not come into force;
- CDM might be superseded by the mechanisms under the Paris Agreement; and
- CDM might be superseded by EU Emissions Trading System (EU ETS).

In the light of the above-mentioned uncertainties on the world carbon markets, the Serbian Government has not enabled CDM implementation under the Climate Change Act via by-laws yet as stipulated in section 2 Article 18 of the Climate Change Act. The future of GHG emission reduction in our country lies in both activities of the public administration and regulatory agencies and activities of the private sector and individuals. Thus, the initiative for GHG emission reduction should be strengthened as Serbia proceeds to align its regulations and politics in accordance with the European Union.

#### **4. BENEFITS OF IMPLEMENTATION OF CCS PROJECTS**

Based on the literature review and analysis of the acts and agreements on the climate change, we could divide the benefits of implementation of CCS in the oil and gas industry into three groups:

- environmental;
- economic; and
- social;

All three groups of the benefits listed above have a positive impact both local and global and are considered a crucial factor for the justification of CCS projects in the oil and gas industry. Without meeting, at least, the minimum environmental, economic, and social benefits and requirements, the CCS projects would not even be launched, and that is why it is important to identify most of them and understand their role in the deployment of CCS projects. Accordingly, both the social acceptance and law justification must be provided and aligned with the goals of the society [2, 16].

##### **4.1. Environmental benefits**

The key environmental effect of CCS deployment might be defined as additional to the primary climate change mitigation effect through reduced CO<sub>2</sub> emissions. Some of the CCS projects in the oil and gas industry might have an effect related to the removal of CO<sub>2</sub> from the atmosphere, which is meant necessary in many climate policy scenarios. In such cases, CCS deployment is essential for two removal technologies, namely Bioenergy with CCS (BECCS) and Direct Air Capture (DAC) [19]. Both the BECCS and DAC technologies are innovative and disruptive, and they challenge entities of the private and public sectors to commit to emissions reductions.

As the overall GHG emissions level is relatively high, the Republic of Serbia would benefit from the implementation of the CCS projects that would bring it closer to the target to reduce its GHG emissions by 9.8% below 1990 levels by 2030. Deployment of the CCS projects, along with fostering other technologies and solutions for the reduction of GHG emissions, could be one of the best approaches to achieving sustainable development goals and building a better environment for society.

##### **4.2. Economic benefits**

The most significant economic benefits of the deployment of CCS projects in the oil and gas industry are the trading of CER credits and enhanced oil recovery.

By enabling the CDM under the Climate Change Act, the Serbian Government will provide the companies implementing CCS projects with an opportunity to benefit from trading the CER credits. So, the future of GHG emissions reduction, especially CO<sub>2</sub>, could be relied upon the trading of the CER credits, which will ease the achievement of the sustainable development goals and obligations under the Paris Agreement.

While the CDM is not enabled under the Climate Change Act of the Republic of Serbia, the most effective established form of economic benefits from carbon capture and storage projects these days is enhanced oil recovery (EOR), whereby the CO<sub>2</sub> is injected into oil and gas reservoirs to increase their extraction. The first CO<sub>2</sub> EOR in Serbia was applied by NIS j.s.c. Novi Sad in Rusanda oil field in 2015 with the injection of 150 000 cubic meters of CO<sub>2</sub> per day into one well [7]. This project was initiated to store separated CO<sub>2</sub> from

natural gasses and to increase oil recovery, which has both economic and environmental effects, and it is considered a safe, secure, and cost-effective project.

With the aim to reduce CO<sub>2</sub> emissions and empower oil recovery projects, the NIS j.s.c. Novi Sad has invested more than 30 million euros in the construction of the high-pressure acid capture technology (HiPACT) unit which can separate 99% of CO<sub>2</sub> from gas [7]. The main reason for the development of this project was the high content of CO<sub>2</sub> in natural gas mixtures (80%) from a certain number of gas fields, and it has shown that there is a wide range of possibilities that can improve both GHG emissions reduction and extraction of oil.

By continuous CO<sub>2</sub> injection, reservoir pressure started to increase, and this effect has been observed and measured in all wells. As a result, NIS j.s.c. Novi Sad benefited from increased extraction of oil in the Rusanda oil field, but the exact quantities of enhanced extraction and received income are not publicly disclosed due to legal limitations. Accordingly, the economic effects were achieved, and the NIS j.s.c. Novi Sad was not only to meet the economic benefits, but it was also the Republic of Serbia that collected higher taxes and reduced level of the pollution.

There are more oil fields, such as Rusanda, that have a CO<sub>2</sub> EOR potential in Serbia, and their testing, assessment, and enhanced exploration could be expected to happen in the following years.

### 4.3. Social Benefits

The reduction of CO<sub>2</sub> emissions under deployed CCS projects have positive effects on the overall climate-related ecosystem, including human and animal health.

The deployment of CCS projects decreases the damage to human and animal health caused by the effects of ozone depletion, formation of a particular matter, ionizing radiation, and photochemical oxidation [15]. Additionally, the Republic of Serbia in general, and Serbian society in particular, would benefit from new jobs created by the deployment of new CCS projects in the oil and gas industry. Besides the jobs created and protecting human and animal health, the deployment of CCS projects could be the beginning of a new era that would foster the change of the population's perspective and habits when it comes to pollution and environmental care.

Nowadays, every player in the oil and gas industry, including NIS j.s.c. Novi Sad, as well, is implementing social responsibility policies and issues the company's sustainability development reports annually demonstrating the social impact of the implementation of projects aimed to decrease GHG emissions.

## 5. POTENTIAL OF CARBON CAPTURE AND CARBON STORAGE

The world oil and gas industry currently offers a prospective solution for a significant decrease in GHG emissions, and that solution envisages a capture of CO<sub>2</sub> emissions and their storage in suitable geological formations [12]. However, potential carbon storage areas in Serbia might be the depleted or partially depleted oil and gas fields which have many advantages compared with other geological formations including very low exploration costs, good traps and seals, well-described reservoir properties, and the presence of surface facilities and existing oil and gas pipelines that can be reused for CO<sub>2</sub> transportation and injection [5, 8].

The geographical position of Serbia is at crossroads between Central and South-Eastern Europe covering the southern part of the Pannonian and central Balkan. The northern part of Serbia is in the Pannonian basin where most of the oil and gas fields are located and fall within the boundaries of Hungary, Croatia, and Romania. The Serbian part of the Pannonian basin covers around 26 000 km<sup>2</sup> with over the 80 discovered oil and gas fields with more than 300 reservoirs. More than 400 exploration wells were drilled and around 1400 appraisal wells were also drilled. Currently, there are 53 oil and gas fields in production with around 1000 wells [13].

Most of the Serbian oil and gas fields are mature fields and potential candidates for enhanced oil recovery [18], and the total CO<sub>2</sub> storage potential in Serbia would be around 110 Mt [11]. As the oil and gas fields in the Serbian part of the Pannonian basin have a significant capacity for CO<sub>2</sub> storage, it is strongly believed that carbon capture and storage projects could have a great potential in the oil and gas industry in the Republic of Serbia. Operating the oil fields and having key expertise in the area, NIS j.s.c. Novi Sad has an advantage in the implementation of carbon capture and storage projects in Serbia, which could make this company a national and regional leader in deploying CCS projects.

## 6. CONCLUSION

The Serbian Climate Change Act provides a framework for the CDM mechanism that confirmed its efficiency under the Kyoto Protocol. So far, Serbia has deployed seven CDM projects under the Kyoto Protocol but none of them were in the area of CCS. Considering the environmental, economic, and social benefits, all with a

positive impact on the Republic of Serbia and its society, the deployment of CCS projects in the oil and gas industry has a significant perspective not only when it comes to NIS j.s.c. Novi Sad but for all major companies in the oil and gas industry of the Republic of Serbia. The conducted research has also revealed that CCS projects in the oil and gas industry have a great potential in Serbia as the Serbian oil and gas fields have a capacity for massive CO<sub>2</sub> storage and the total CO<sub>2</sub> storage capacities in Serbia might be around 110 Mt. Keeping that in mind, the Republic of Serbia could become one of the most important and most influential countries in the region, especially in the domain of storage of removed CO<sub>2</sub>.

Unfortunately, the Serbian Government has not enabled CDM implementation under the Climate Change Act via by-laws yet as stipulated in its section 2 Article 18 due to uncertainties in the world carbon markets and further application of CDM in the European Union. Based on previous findings, we come to the conclusion that the CDM should be enabled regardless of the mentioned uncertainties as the players in the oil and gas industry may delay the deployment of CCS which would negatively affect the Republic of Serbia's target to reduce its GHG emissions by 9.8% below 1990 levels by 2030. The public administration of the Republic of Serbia should put its efforts into expanding the regulatory framework and fostering the CDM and CCS technologies which could make a significant reduction in GHG emissions. Additionally, new or updated climate change acts should be adopted, ideally, the ones which cover most of the segments of CDM, CCS, and other mechanisms and technologies focused on reduction in GHG emissions and stopping climate change.

## REFERENCES

- [1] Casino, J.M., Collado, R.R., & Nassar S. (2022). The clean development mechanism in Eastern Europe: an in-depth exploration. *Environmental Science and Pollution Research*. Retrieved from <https://link.springer.com/article/10.1007/s11356-022-20988-3>
- [2] De Medeiros Costa, K., & Arlota, C. (2021). *Carbon Capture and Storage in International Energy Policy and Law*. The Netherlands: Elsevier.
- [3] Doeh, D., Shergold, S., & Fominykh, S. (March 6, 2009). Consideration of JI projects in Russia. Russia: Lexology. Retrieved from <https://www.lexology.com/library/detail.aspx?g=d11edf4f-c79f-421d-abde-e1153f780018>
- [4] Hiil, J. (2020). *Environmental Social and Governance (ESG) Investing: A Balanced Analysis of the Theory and Practice of a Sustainable Portfolio*. UK: Academic Press.
- [5] Holloway, S., & Van Der Straaten, R. (1995). The Joule II project the underground disposal of carbon dioxide. *Energy Conversion and Management* 36, no. 6-9, pp. 519-522.  
doi:10.1016/0196-8904(95)00057-k
- [6] Kainou, K. (2022). Collapse of the Clean Development Mechanism scheme under the Kyoto Protocol and its spillover: Consequences of 'carbon panic'. *VOX EU*. Retrieved from <https://voxeu.org/article/collapse-clean-development-mechanism>
- [7] Karas, D., & Nesic, S. (2019). CO<sub>2</sub> Sources, transportation, and storage possibilities in Serbian oil and gas fields. *Materials of 10<sup>th</sup> Trondheim Conference on Co2 Capture, Transport and Storage, June 17-19, 2019*.
- [8] Knoope, M. M. J., Guijt, W., Ramírez, A., & Faaij, A. P. C. (2014). Improved cost models for optimizing CO<sub>2</sub> pipeline configuration for point-to-point pipelines and simple networks. *International Journal of Greenhouse Gas Control*, 22, pp. 25–46.  
doi:10.1016/j.ijggc.2013.12.016
- [9] Kuckshinrichs, W. (2015). Carbon Capture and Utilization as an Option for Climate Change Mitigation: Integrated Technology Assessment. In W. Kuckshinrichs & Hake J.F. (Eds.), *Carbon Capture, Storage and Use: Technical, Economic, Environmental and Societal Perspectives* (pp. 1-9). Switzerland: Springer International Publishing.
- [10] Naftna industrija Srbije [NIS j.s.c. Novi Sad]. (2021). *NIS Sustainable development report 2021*. Retrieved from <https://www.nis.rs/wp-content/uploads/2022/04/odrzivi-low.pdf>
- [11] Naftna industrija Srbije [NIS j.s.c. Novi Sad]. (2021). *About the company*. Retrieved from <https://www.nis.rs/en/company-information/>
- [12] National Energy Technology Laboratory [NETL]. (2010). *Site Screening, Selection, and Initial Characterization for Storage of CO<sub>2</sub> in Deep Geologic Formations*. No. DOE/NETL-401/090808.

Retrieved from <https://www.netl.doe.gov/File%20Library/Research/Carbon-Storage/Project-Portfolio/BPM-SiteScreening.pdf>

- [13] Nesic, S., Mitrovic, V., & Filipovic, M. (2017). Serbia evaluates its CO<sub>2</sub> storage potential. *Oil and Gas Journal*, December.
- [14] Richie, H., & Roser M., (2020). Serbia: CO<sub>2</sub> Country Profile. Retrieved from <https://ourworldindata.org/co2/country/serbia>
- [15] Seigo S., L'orange S., Dohle S., & Siegrist M. (2014). Public perception of carbon capture and storage (CCS): A review. *Renewable and Sustainable Energy Reviews*, 38, pp. 848-863.  
doi:10.1016/j.rser.2014.07.017
- [16] Schumann, D. (2015). Public Acceptance. In W. Kuckshinrichs & Hake J.F. (Eds.), *Carbon Capture, Storage and Use: Technical, Economic, Environmental and Societal Perspectives* (pp. 221-251). Switzerland: Springer International Publishing.
- [17] The Climate Change Act of the Republic of Serbia, Službeni glasnik RS, 26 (2021). Retrieved from <https://www.propisi.net/zakon-o-klimatskim-promenama/>
- [18] Tomic, L., Karovic-Maricic, V., & Danilovic, D. (2021). The preliminary selection of oil reservoir in Serbia for carbon dioxide injection and storage by a multicriteria decision-making approach: a case study. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*. Retrieved from <https://www.tandfonline.com/journals/ueso20>
- [19] Torvanger, A. (2019). Ancillary Benefits of Carbon Capture and Storage. *Ancillary Benefits of Climate Policy*, pp. 213–225.
- [20] United Nations Development Programme in Serbia [UNDP Serbia]. (2022). Serbia and Climate Change. Retrieved from <https://www.klimatskepromene.rs/obaveze-prema-un/kjoto-protokol/cmd-projekti/>



# SAOBRAĆAJ, TRANSPORT I KOMUNIKACIJE

---

## TRAFFIC, TRANSPORTATION AND COMMUNICATION



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business





## A DECISION TREE MODEL FOR AIRPORT ACCESS MODE CHOICE: SERBIAN AIR TRANSPORT MARKET

KATARINA KUKIĆ, SLAVICA DOŽIĆ, DANICA BABIĆ, MILICA KALIĆ

University of Belgrade – Faculty of Transport and Traffic Engineering,  
k.mijailovic@sf.bg.ac.rs, s.dozic@sf.bg.ac.rs, d.babic@sf.bg.ac.rs, m.kalic@sf.bg.ac.rs

**Abstract:** *In order to achieve the seamless multimodal door-to-door journey, it is important to understand passengers' preferences and behaviour regarding the airport access mode choice. The objective of this paper is to propose a model which will help prioritising the factors that influence airport access mode choice. The proposed model is based on Decision Tree with Gini index for measuring data impurity. The data related to Serbian air transport market are obtained from the online survey. The results show that public transport in Serbia must offer services which are safe, fast, reliable and performed at high frequency in order to attract more users.*

**Keywords:** *Decision Tree model, Airport access, Air transport market, Survey*

### 1. INTRODUCTION

A door-to-door journey for an air passenger represents an intermodal chain consisting of the following stages: access to the airport via different surface transport modes or systems, the airport and non-airport activities at the terminal(s) before the flight, the airline flight, the activities at the airport terminal(s) after the flight and egress from the airport to the final destination via different surface transport modes or systems [1]. Total travel time consists of pre-flight time, flight time and post-flight time. At most large airports, the landside access modes and systems are based on the road and railway transport modes (cars, taxis, buses, subway/metro, conventional and high-speed rail). At the regional airports, as an airport ground access mode, most frequently used is the road-based transport (e.g. "Nikola Tesla" airport in Belgrade).

The factors influencing choice of the airport landside access modes and their systems or individual/private cars are following: availability, access time, access cost, transport service frequency, reliability, punctuality, and convenience of the arrival time at the airport, convenience of storing and retrieving luggage and whether access with transfer or without transfers. The access time and price have been approximately linearly correlated with the airport access distance at almost all landside access modes and their systems across many European and US airports [2].

Air travellers' access mode choice models are based on individual characteristics of passengers (gender, age, car ownership, income, etc.) and alternative-specific attributes. In order to understand airport accessibility, researchers take into account trip purpose (business or leisure air trip), solo/group journey, the size of passenger group, number of baggage, etc. Additionally, time of access (peak or non-peak period) and journey distance (from passenger's journey origin to the airport) are important factors, too. The most important issue is to understand passengers' preferences and behaviours with respect to the access mode choice, considering the possible way to achieve the seamless multimodal door-to-door journey, as a future passenger service. Airport ground accessibility has been widely researched such as access mode choice in the light of passengers' preferences and behaviours [3, 4], modal split for relocated airports or an assessment of the introduction of a new mode [5, 6].

Machine learning and data mining techniques are proved to be very suitable tool for analysing and understanding passengers' behaviour in transport, due to their higher performances in comparison to traditional techniques. The data mining techniques could be described as solutions used to analyse huge amounts of data and turn it into useful information and knowledge, [7]. Basically, machine learning is divided into three classes: unsupervised (model obtained from unlabelled data, includes classification and regression), supervised (model is learning from labelled data, includes clustering and association) and reinforcement learning (model is learning by being rewarded for each correct move, and penalized for the incorrect action). Classification and Regression Trees (CART) is an algorithm proposed by [8] that can be used for both regression and classification problems and it is usually referred as decision tree algorithm.

Related to implementation of machine learning techniques in transport problems, it is noted that researchers are proposing increasingly complex models, e.g., for analysing passenger satisfaction with public transport,

while at the same time there is a need from managers and transport service providers for simpler models that are easier to interpret [9]. A data mining technique which can meet these goals is a decision tree (DT) algorithm. It is relatively simple but powerful supervised learning method that uses tree-like model of decisions and their possible consequences. Different procedures for DT algorithms have been developed. These algorithms may use different split criteria for selecting the attribute to be placed in a node and branching (Gini index, Information Gain, Information Gain Ratio, etc.), different criteria to stop the branching of the tree and methods for assigning a class label. There are numerous studies in which DTs are used as a tool to analyse public transport systems, for example [10], road safety [11], [12], air transport [13], as well as recent papers considering attitudes of intercity train users' [14] and satisfaction of highly educated people with airlines' services [15].

The goal of this paper is to propose a model, based on DT and data related to Serbian air transport market, which will be acceptable and understandable for decision makers, and which will help prioritising the factor that influence airport access mode choice. Furthermore, it will help improving service offered to passengers.

## 2. METHODOLOGY

### 2.1. The decision tree

DT is a flowchart tree structure with rules in form of if-then rules and it is formed by a limited number of levels from root node until leaf (terminal) nodes. It is a non-parametric model that does not assume any predefined underlying relationship between dependent and independent variables. For building decision tree, the parent node is divided into two child nodes through the splitter variable. Starting from the root node which contains entire dataset, a DT first splits the nodes on all the available variables and then selects the split which results in the most homogeneous child nodes.

In this paper, we used CART algorithm with Gini index for impurity ( $GI$ ) as attribute selection measure for making splits of the dataset. Gini impurity represents the probability that new data will be misclassified if they are assigned a random according to the distribution of class labels from the dataset. It is calculated by subtracting the sum of squared probabilities of each class from one, which is in the case of classification:

$$GI(node) = 1 - \sum_{i=1}^N \left( \frac{\text{number of class } i \text{ cases}}{\text{number of all cases in the node}} \right)^2 \quad (1)$$

where  $N$  is the number of classes of the independent variables. As a next step, we need to sum weighted Gini indexes for left and right child nodes (the weight of a child subnode is the number of samples in that node divided by the total number of samples in its' parent node) in order to obtaining weighted GI for entire split:

$$\text{WeightedGI}(node) = \text{weight}(\text{left subnode}) \cdot GI(\text{left subnode}) + \text{weight}(\text{right subnode}) \cdot GI(\text{right subnode}) \quad (2)$$

The same procedure is performed for all attributes and attribute with the smallest *WeightedGI* is selected for splitting the node, in order to minimize impurity of child nodes by selecting the best data partition. As a prediction is taken the most common label in each child node, [16]. The splitting goes until all child nodes are pure or their impurity cannot further be decreased, which can result by overfitting of the DT. That is usually prevented by setting constrains on number of levels (depth) in the tree, as well as on choosing minimal number of cases for parent or child nodes and by implementing cross validation procedure to obtain a more reliable result.

The goal is to create a model that predicts the value of a target variable by learning simple decision rules. Advantage of DT compared to other modelling methods is in the using if-then rules which allow us to obtain logic conditional structure and to create simple visual presentation of obtained model. Another advantage of this model is that it provides opportunity to obtain the rank of the independent variables according to their importance.

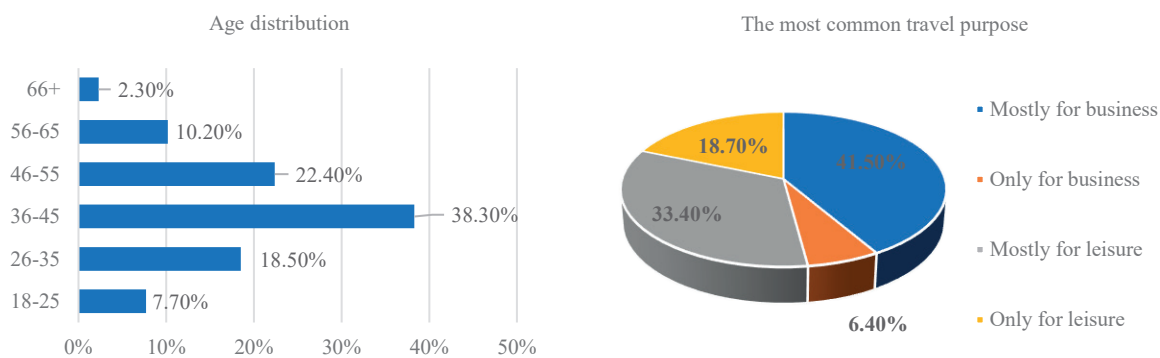
### 2.2. Data

Data considered in this study are obtained within the survey conducted in April and May 2021 for the purpose of the SYN+AIR project (894116) from the H2020-SESAR-2019-2 call. The survey was conducted online due to Covid-19 pandemic. Most of 2251 respondents were from Greece, Italy, Spain and Serbia (where project partners originated from), as well as from the other countries in Europe. The aim of the survey was to "quantify the tradeoffs that users consider when selecting travel alternatives and identify traveller characteristics that reflect their emotions, attitude, and travel behaviour", [17]. The questionnaire was available in five languages



(English, Serbian, Italian, Greek and Spanish) and it consisted of 25 questions divided in three parts: Mobility profile (frequency and most common purpose of travel by plane, habits related to planning a trip and, as a focus, group of questions related to mode choice to/from the airport), Travel preferences (questions related to three different hypothetical travel scenarios) and Socio-demographic profile. It was distributed online on EU survey platform, and disseminated via emails, social networks, project's website, etc. The complete questionnaire and descriptive statistics results from the survey can be seen in [17]. In this paper, we will focus only to respondents from Serbia.

Totally 562 respondents from Serbia filled the questionnaire, from which 66% were male, 33.4% female respondents, and the rest declare either as other or rather not say. Percentage of respondents who often travel by plane (in regular conditions, without Covid-19) is 43.6, while 30.4% of respondents travel rarely, 15.5% frequently and 10.5% of respondents almost never travel by plane. Sample distribution regarding age and most common purpose of travel by air are presented on Figure 1. Related to employment status 86.8% of respondents were employed, 4.6% students, 3.9% retired, 1.6% unemployed and 3% other. Approximate household income was assessed as average by 47% of respondents, 38.6% think their household income is high, 1.2% have low income, while 13.2% of respondents choose not to answer on this question.



**Figure 1:** Age and travel purpose distribution in sample from Serbia in SYN+AIR survey, 2021

The main part of this research is related to the questions about mode choice to/from the airport and factors which influence that choice. The first considered question relates to mode preference for travelling to/from the airport if all modes are available, and frequency distributions of the responses are presented in Table 1.

**Table 1:** Frequency distribution of responses related to airport access mode choice if all modes are available

Airport access mode choice if all modes are available	Frequency	Share
Bus	7	1.2%
Car (park at/near the airport)	37	6.6%
Car (someone drops me off/picks me up)	216	38.4%
Combination of modes (e.g. bus & train)	14	2.5%
Metro	119	21.2%
Other	2	0.4%
Taxi (or ridesharing services like Uber or Lyft)	124	22.1%
Train	43	7.7%
Total	562	100.0%

Second question used for DT model considers how much selected factors influence mode choice when travelling to and from the airport converted from five-point verbal scale (from not important to most important) to five-point numerical scale (respectively 1 to 5). Considered factors, as well as the means and the standard deviations (SD) are presented in Table 2. From results presented in Table 2, we see that the most important factor regarding mode choice to and from airport for respondents from Serbia is reliability, and it is followed by security, crowdedness, travel time and waiting time.

**Table 2:** How much do the following factors influence mode choice when travelling to/from the airport:

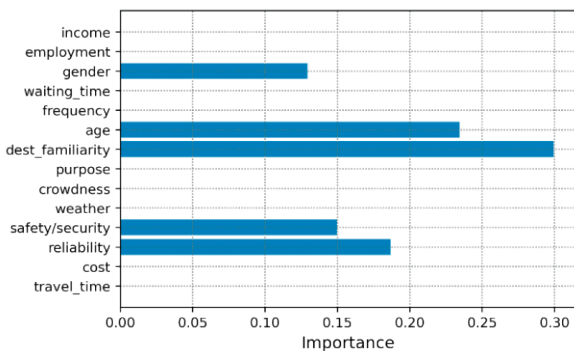
Factors	Mean	SD	Factors	Mean	SD
Reliability (e.g., whether your bus may be delayed or stuck in traffic)	3.89	0.86	Cost (e.g., total cost of a bus ticket)	2.95	0.98
Security (e.g., the possibility of getting mugged)	3.48	1.14	Familiarity of the city (e.g., whether it is your first time visiting a location, or if you are travelling within your own city)	2.86	1.03
Crowdedness (e.g., a crowded bus or crowded train platform)	3.28	0.93			
Travel Time (e.g., time spend in the bus)	3.18	0.95	Trip purpose (e.g., leisure or business)	2.59	1.10
Waiting Time (e.g., waiting for the train at the platform)	3.07	0.93	Weather (e.g., rainy or cold weather conditions)	2.51	1.02

### 3. DECISION TREE MODEL AND RESULTS

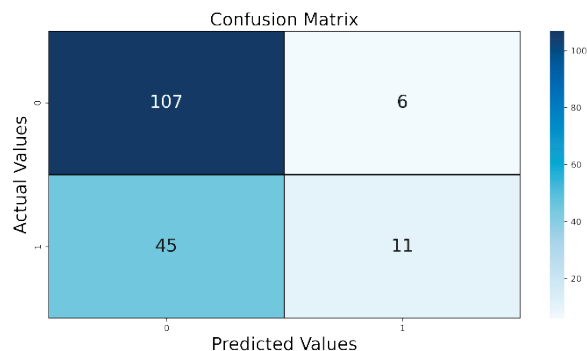
One of the goals of future new service examined within SYN+AIR project (single ticket for whole multimodal journey including air transport as the main leg) is shift of passengers from private (car or taxi) to public transport (PT). In order to better understand the needs and attitudes of respondents from Serbia related to factors that influence airport access mode choice, we grouped categories from Table 1 into two new categories: PT (joined Bus, Metro, Train, Other, Combination of modes) – coded as 1 and car or taxi (joined Car (park at/near the airport), Car (someone drops me off/picks me up) and Taxi (or ridesharing services like Uber or Lyft)) – coded as 0. Then, this new created variable was chosen as dependent variable. As independent variables we used factors from Table 2, as well as selected socio-demographic and mobility profile’s characteristics, totally 14 variables with their importance, Figure 2. From socio-demographic variables, we considered: gender (female – 1, male – 2, other – 3, rather not say – 4), age (scale 18-100), employment status (employed in public sector – 1, employed in private sector – 2, other – 3, student – 4, retired – 5, unemployed – 6) and approximate household income (average – 1, low – 2, high – 3, rather not say – 4). From mobility profile's characteristics we considered the most common purpose of travel by plane (mostly for business – 1, mostly for leisure – 2, only for business – 3, only for leisure – 4) and frequency of travel by plane (almost never – 1, frequently – 2, often – 3, rarely – 4).

A decision tree was built (in Python's decision tree classifier) to classify the respondents' airport access mode choice into PT as one category and Car or Taxi as second, as well as additionally to identify which independent variables play a key role in this classification. In order not to overfit the model, we used percentage split of 70% to build DT (70% of data are used for building the model, or totally 393 responses) and the remaining 30% are used for validating (totally 169 responses). The tree depth was specified as three, since from the obtained accuracy for tree depth is in range two to eight, we realized that results were similar for values two to five, but slightly better for tree depth equals three. Additionally, that depth was reasonable to keep the tree simple to understand.

The performance of the model is presented through the confusion matrix which shows the correctly and incorrectly classified instances for each class, Figure 3. As a measure of performance most widely used, here we present accuracy obtained directly from confusion matrix as a ratio of correctly classified instances (truly positive and truly negative) and all instances, which is in our model 118/169, or 70% and it is considered as acceptable value for DT algorithms precision compared to similar studies, [9] – [15].



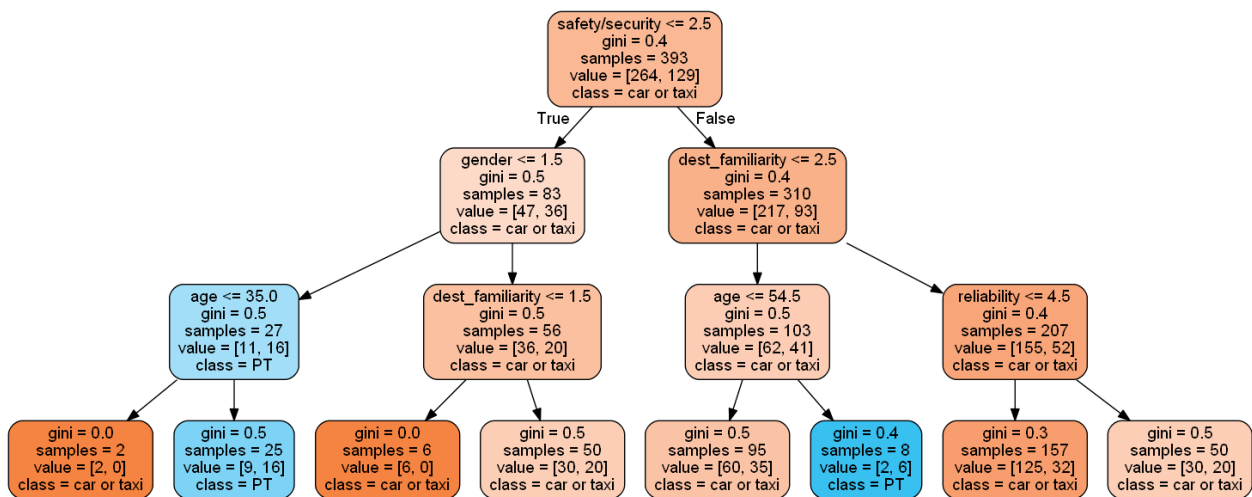
**Figure 2:** Factors importance ranking



**Figure 3.** Confusion matrix

The DT model with depth three is built with 15 nodes, of which eight are terminal nodes, Figure 4. As previously stated, the most common label in each child node is taken as a prediction. On Figure 4, nodes coloured in blue predict PT as a mode choice to/from the airport, while nodes in different tones of peach colour predict use of car or taxi (darker tone indicates higher data purity). Let us note here, that question related to choosing mode of transport to/from the airport that we considered was hypothetical question, stated as “If all of the following transport modes are available, which one would you choose to travel to/from the airport?”. Anyway, respondents gave answers based on their experience, and it is well known that people in Serbia use car or taxi as a main mode when travel to/from the airport in Serbia. But, with this question we wanted to examine mode choice to/from airport in both situations when the destination is familiar to respondents and also when it is not.

From the obtained DT model, we can conclude that the main primary splitters in the classification passengers from Serbia are: safety/security (the possibility of getting mugged for example in PT), gender, age, reliability (possibility that for example bus may be delayed or stuck in traffic) and familiarity with the destination where is the airport (e.g., whether it is the first time visiting a location, or travelling within hometown). Importance rankings were also produced and presented at Figure 2. It is interested to notice that familiarity with destination is identified as the most important factor for making choice whether to use PT or car/taxi as a mode choice to/from the airport. The same factor was rated as seventh of nine considered factors in question „How much do the following factors influence mode choice when travelling to/from the airport?“, see Table 2. Familiarity with destination was followed by age and gender, where female respondents who find safety/security important below 2.5 (on scale 1 to 5) and who are older than 35 are more likely to choose PT. Another group of respondents which is more likely to choose PT are respondents older than 55 who find safety/security important above 2.5, but familiarity of destination below 2.5.



**Figure 4:** Decision tree for classification airport access mode choice in two classes: PT and car or taxi

However, most respondents opt for a car or taxi because safety/security is most important for them (310 respondents), then the familiarity with destination (207 respondents), and last but not least, the reliability (157 respondents) that reflects the level of system performance in terms of punctuality, delays and capacity. People under the age of 55 tend to use less public transport and more private cars and taxi because they find it safer/secured. This result implies that PT in Serbia must offer services which are safe, fast, reliable and performed at high frequency in order to attract more users.

#### 4. CONCLUSION

The objective of this paper was to investigate what are the most important factors for air passengers from Serbia when choosing between use of public and private transport for airport access mode choice. Decision tree algorithm was chosen since the outcomes of this model are easily understandable and suitable to be visually represented. From obtained decision rules, useful information may be provided to transport service providers, public transport managers and operators in order to prioritize the measures that must precede the introduction of a new service providing a single ticket for door-to-door journey with air transport as the main leg. As not so positive side of DT algorithm, we must state its instability, in a sense that small changes in input training samples may cause dramatically large changes in output classification rules. Anyway, we think that

relatively simple way to identify the most important independent variables for classification which DT provides should be used in studies like this, which results need to be easily understandable to transport service providers, etc.

## Acknowledgement

Parts of this study has been supported by SESAR Joint Undertaking within the project “Synergies between Transport Modes and Air Transportation” (SYN + AIR) no. 894116 and the Ministry of Education, Science and Technological Development, Republic of Serbia, as part of projects TR36033 and TR36002 (2011–2022).

## REFERENCES

- [1] Babić, D., Kalić, M., Janic, M., Dožić, S., Kukić, K. 2022. Integrated Door-to-Door Transport Services for Air Passengers: From Intermodality to Multimodality. *Sustainability*, 14, 6503.
- [2] Janić M., 2019. *Landside Accessibility of Airports, Analysis, Modelling, Planning, and Design*, Springer International Publishing AG, part of Springer Nature 2019, ISBN 978-3-319-76149-7.
- [3] Tam, M.L., Lam, W., Lo, H.P. 2011. The Impact of Travel Time Reliability and Perceived Service Quality on Airport Ground Access Mode Choice. *Journal of Choice Modelling*, 4(2), 49-69.
- [4] Akar, G. 2013. Ground access to airports, case study: Port Columbus International Airport. *Journal of Air Transport Management*, 30, 25-31.
- [5] Jou, R., Hensher, D., Hsu, T. 2011. Airport ground access mode choice behaviour after the introduction of a new mode: A case study of Taoyuan International Airport in Taiwan. *Transportation Research Part E: Logistics and Transportation Review*, 47(3), 371-381.
- [6] Birolini, S., Malighetti, P., Redondi, R., Deforza, P. 2019. Access mode choice to low-cost airports: Evaluation of new direct rail services at Milan-Bergamo airport. *Transport Policy*, 73, 113-124.
- [7] Han, J., Kamber, M., Pei, J. *Data Mining: Concepts and Techniques*, 3<sup>rd</sup> Edition, Morgan Kaufmann, 2012, doi: 10.1016/C2009-0-61819-5. Boston.
- [8] Breiman, L., Friedman, J.H., Olshen, R. A., Stone, C. J. 1984. *Classification and Regression Trees*. Chapman and Hall/CRC.
- [9] de Oña, R., de Oña, J., 2015. Analysis of transit quality of service through segmentation and classification tree techniques. *Transp. A Transp. Sci.* 11, 365–387.
- [10] de Oña, J., de Oña, R., Calvo, F.R. 2012. A classification tree approach to identify key factors of transit service quality. *Expert Systems with Applications*, Vol. 39, Issue 12, 11164-11171.
- [11] Abellán, J., López, G., de Oña, J. 2013. Analysis of traffic accident severity using Decision Rules via Decision Trees. *Expert Systems with Applications*, Vol. 40, Issue 15, 6047-6054.
- [12] da Cruz Figueira, A., Pitombo, C.S., de Oliveira, P., Camargo Larocca, A. 2017. Identification of rules induced through decision tree algorithm for detection of traffic accidents with victims: A study case from Brazil. *Case Studies on Transport Policy*, Vol. 5, Issue 2, 200-207.
- [13] Wong, J.Y., Chung, P.H. 2007. Managing valuable Taiwanese airline passengers using knowledge discovery in database techniques. *Journal of Air Transport Management* 13(6), 362–370.
- [14] Farazi, N.P., Murshed, M.N., Hadiuzzaman, M. 2022. Application of machine learning to investigate heterogeneity in users’ perception of intercity train service quality in developing countries. *Case Studies on Transport Policy*, Vol. 10, Issue 1, 227-238.
- [15] Bellizzi, M.G., Eboli, L., Mazzulla, G., Postorino, M. N. 2022. Classification trees for analysing highly educated people satisfaction with airlines’ services. *Transport Policy*, Vol. 116, 199-211.
- [16] Ozdemir, S. *Principles of Data Science*, 2016. Packt Publishing Ltd, Birmingham.
- [17] Mavromatis, K., Kalić, M., Kukić, K., Ottomanelli, M., Colovic, A. SYN+AIR: D3.1 Report on Customer Journeys, H2020-SESAR-2019-2. (Available online: <https://syn-air.eu/>).

## AIRCRAFT ENCOUNTER CHARACTERIZATION: A SCENARIO-BASED SIMULATION APPROACH

LIDIA SERRANO-MIRA<sup>1</sup>, FEDJA NETJASOV<sup>2</sup>, LUIS PÉREZ SANZ<sup>1</sup>, JAVIER ALBERTO PÉREZ-CASTÁN<sup>1</sup>

<sup>1</sup> Universidad Politécnica de Madrid, Departamento de Sistemas Aeroespaciales, Transporte Aéreo y Aeropuertos  
[lidia.serrano@upm.es](mailto:lidia.serrano@upm.es), [l.perez@upm.es](mailto:l.perez@upm.es), [javier.perez.castan@upm.es](mailto:javier.perez.castan@upm.es)

<sup>2</sup> University of Belgrade – Faculty of Transport and Traffic Engineering, Division of Airports and Air Traffic Safety  
[f.netjasov@sf.bg.ac.rs](mailto:f.netjasov@sf.bg.ac.rs)

**Abstract:** Safety is paramount goal of air traffic control (ATC). One of the main concerns in everyday ATC operations is prevention of conflicts, i.e. losses of separation in the given airspace. The objective of the research is to find out main characteristics of different aircraft encounters leading to conflicts and collisions. For that purpose a set of metrics is proposed and a scenario-specific simulation model is built enabling experimenting with changes of different factors contributing to conflicts/collisions as well as separation minima.

**Keywords:** air traffic control, conflict, collision, simulation

### 1. INTRODUCTION

Safety is paramount goal of air traffic control (ATC). One of the main concerns in everyday ATC operations is prevention of conflicts, i.e. losses of separation in the given airspace. Separation minima ( $S_{min}$ ) is used by air traffic controllers (ATCo's) for distinguishing between safe operations and those in which safety is jeopardized – losses of separation. A common separation minima in horizontal plane used worldwide is 5 NM for en-route operations in surveillance environment. Conflicts are considered as incidents and if are not managed in appropriate way by ATCo's they could lead to collisions (accidents, events in which separation between aircraft is below double of horizontal aircraft size, 500 ft [1]).

International Civil Aviation Organization (ICAO) was requesting from ATC to manage flight operations in a way to satisfy a Target Level of Safety (TLS) [2]. TLS value of  $5 \cdot 10^{-9}$  accidents per flight hour is chosen as adequate for en-route flights by ICAO. There are different methodologies proposed for calculation of collision risks and their comparison with TLS [3, 4] which need to prove that Air Traffic Management (ATM) system is safe. In order for ATM system to be proactive as requested by ICAO [5] it is necessary to find a relationship between conflicts and collisions. For that purpose a characterisation of conflicts is needed.

The goal of the research presented in this paper is to find out main characteristics of different aircraft encounters leading to conflicts and collisions. Therefore, a set of metrics is proposed, and a scenario-specific simulation model is built enabling experimenting with changes of different factors contributing to conflicts/collisions as well as separation minima. Research is covering en-route flights and losses of separation in horizontal plane.

### 2. MODELING APPROACH AND METRICS

In order to model aircraft encounters a scenario-based deterministic simulation model of geometric/intersection type is developed in Matlab with specific inputs – initial positions of aircraft in horizontal plane and initial velocities (ground speeds) of both *host* and *intruder* aircraft (Figure 1).

A model is based on following assumptions: no ATCo interventions, aircraft trajectories crosses at certain angle, aircraft velocities are constant during encounter simulation, even in case of collision aircraft artificially continue to fly. Model is discrete in time and is looking into aircraft positions in constant time intervals which value could be varied due to the traffic surveillance system in use.

Metrics chosen to characterise encounters are the following (Figure 2):

- Distance at the closest point of approach (CPA) – DCPA [6],
- Time from the moment when separation minima is lost until CPA - TCPA,
- Conflict duration [6],
- Criticality of conflict =  $(S_{min} - DCPA) / S_{min}$  [6],
- Closing speed (derivative of the relative position vector),
- Acceleration (derivative of the closing speed vector).

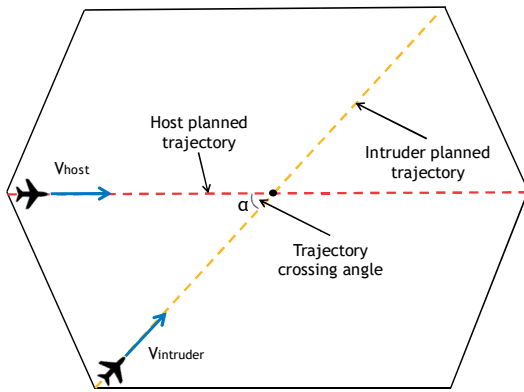


Figure 1: Scenario used in experiments

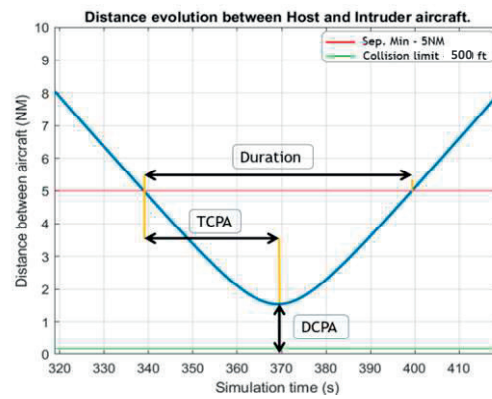


Figure 2: Change of distance between aircraft

### 3. EXPERIMENTS

#### 3.1 Change of intersection angle

In the first experiment an intersection angle was varied (Figure 3a) while velocities remain constant. In this experiment initial trajectory of host aircraft was not changed, but the trajectory of intruder aircraft was rotated around trajectory intersection point. Angles between two trajectories are taken from three categories of encounters [7]: overtaking (catching up – angles of  $20^{\circ}$  and  $40^{\circ}$ ), crossings ( $60^{\circ}$ ,  $90^{\circ}$  and  $120^{\circ}$ ) and head-to-head (H2H,  $145^{\circ}$  and  $175^{\circ}$ ).

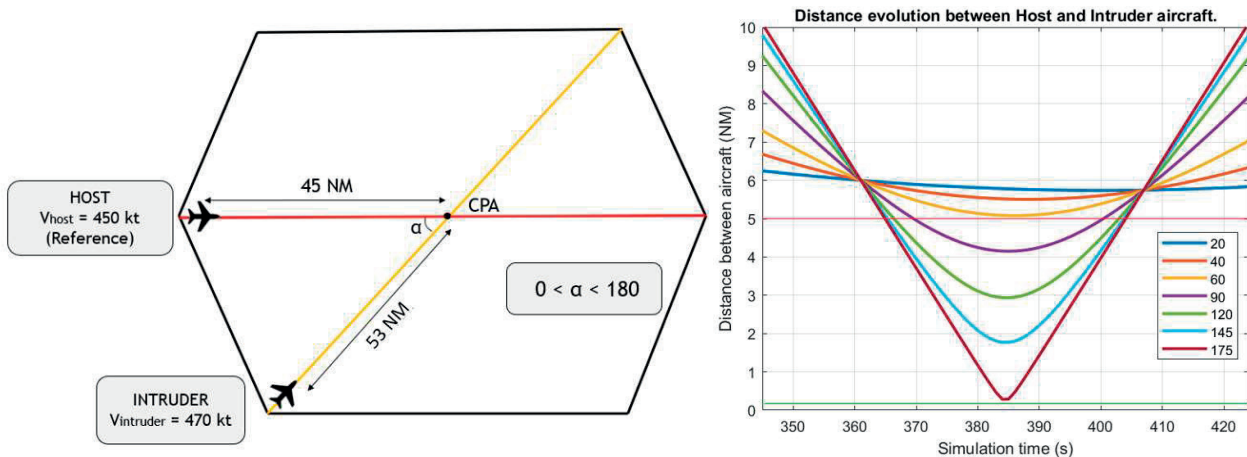


Figure 3: a) Scenario used in experiments with velocities, b) Change of distance between aircraft for different angles

Interesting finding was that when angle increases the curve closes – in case of conflict a curve is of U-shape (Figure 3b) but in case of collision is of V-shape. For near-collision encounters ( $175^{\circ}$  angle) the curve is approaching a V-shape. That means that with higher values of angle, conflict duration becomes longer.

Closing speed, on Figure 4a, corresponds with the gradient of the curve (Figure 3b). It shows the distance variation as a function of time. It is evident from this figure that closing speed curve is more vertical with increase of intersection angle.

Closing speed derivative (acceleration) shows how fast closing speed changes (Figure 4b). In Figure 5, the difference between the adjacent elements of the closing speed vector for different intersection angles has been plotted. As can be seen, the function defining these curves approximates the probability density function of the normal distribution. However, for near collision encounters this representation takes a sharper form (thorn like, Figure 5a). On Figure 5b curves corresponding to lower angles are presented. It is clear that the standard deviation is higher in those encounters where the minimum separation between aircraft is far from the collision, but below 5 NM. This means that the variation of the gradient in the curve of the relative distance as a function of time (Figure 3b), which is the closing speed, is slower than in the collision case.

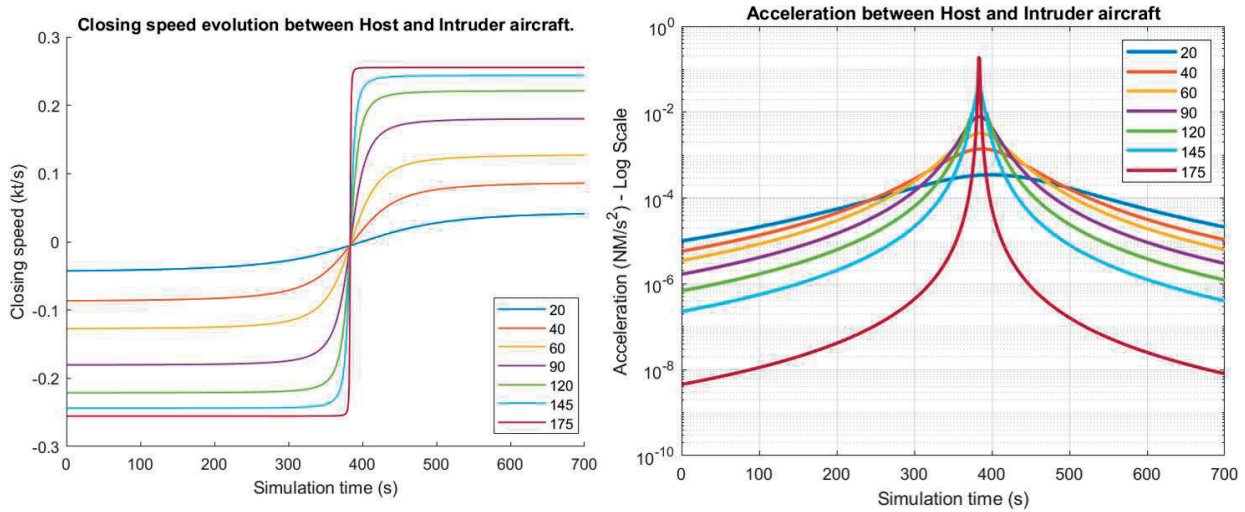


Figure 4: a) Closing speed for different angles, b) Acceleration for different angles

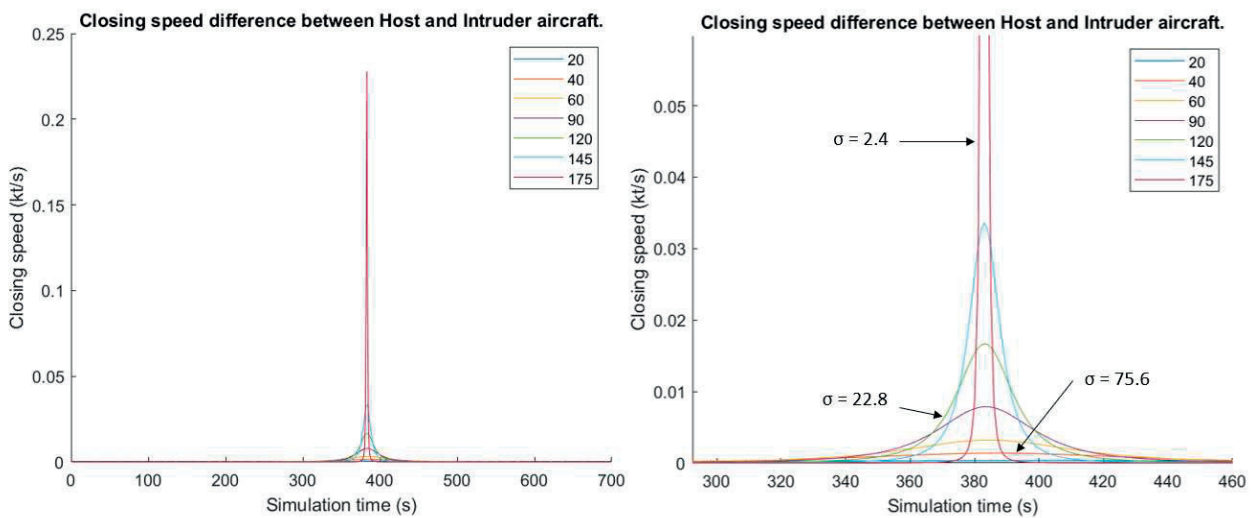


Figure 5: a) Difference between the adjacent elements of the closing speed vectors, b) Zoom on smaller angles

### 3.2 Change of intruder aircraft velocities

In the second experiment velocity of the intruder aircraft was varied (Figure 6) for a given values of intersection angles representing each encounter category. A range of velocity values (420 kt to 530 kt) was chosen to represent situations in which intruder is slower than host, have the same speed as host or is faster than host aircraft. From Figure 6 it can be seen that increasing the velocity, the CPA is reached earlier.

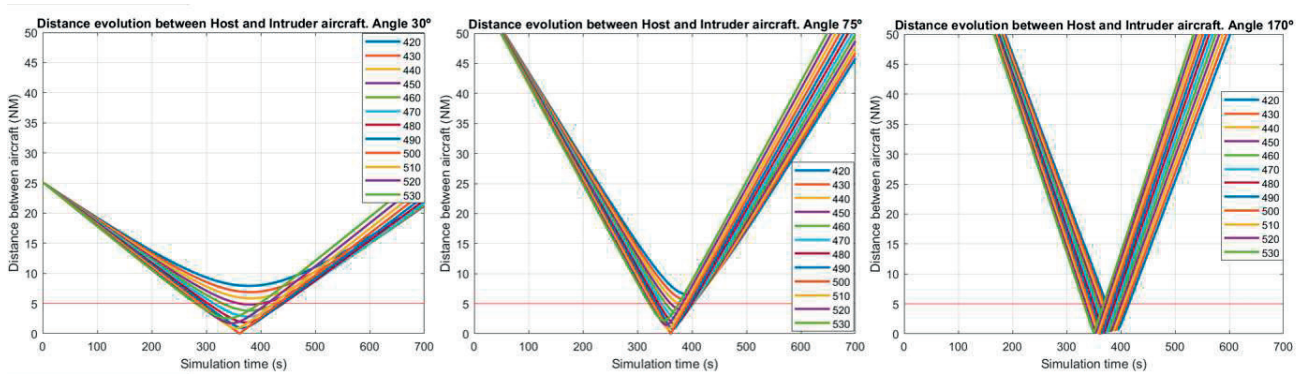
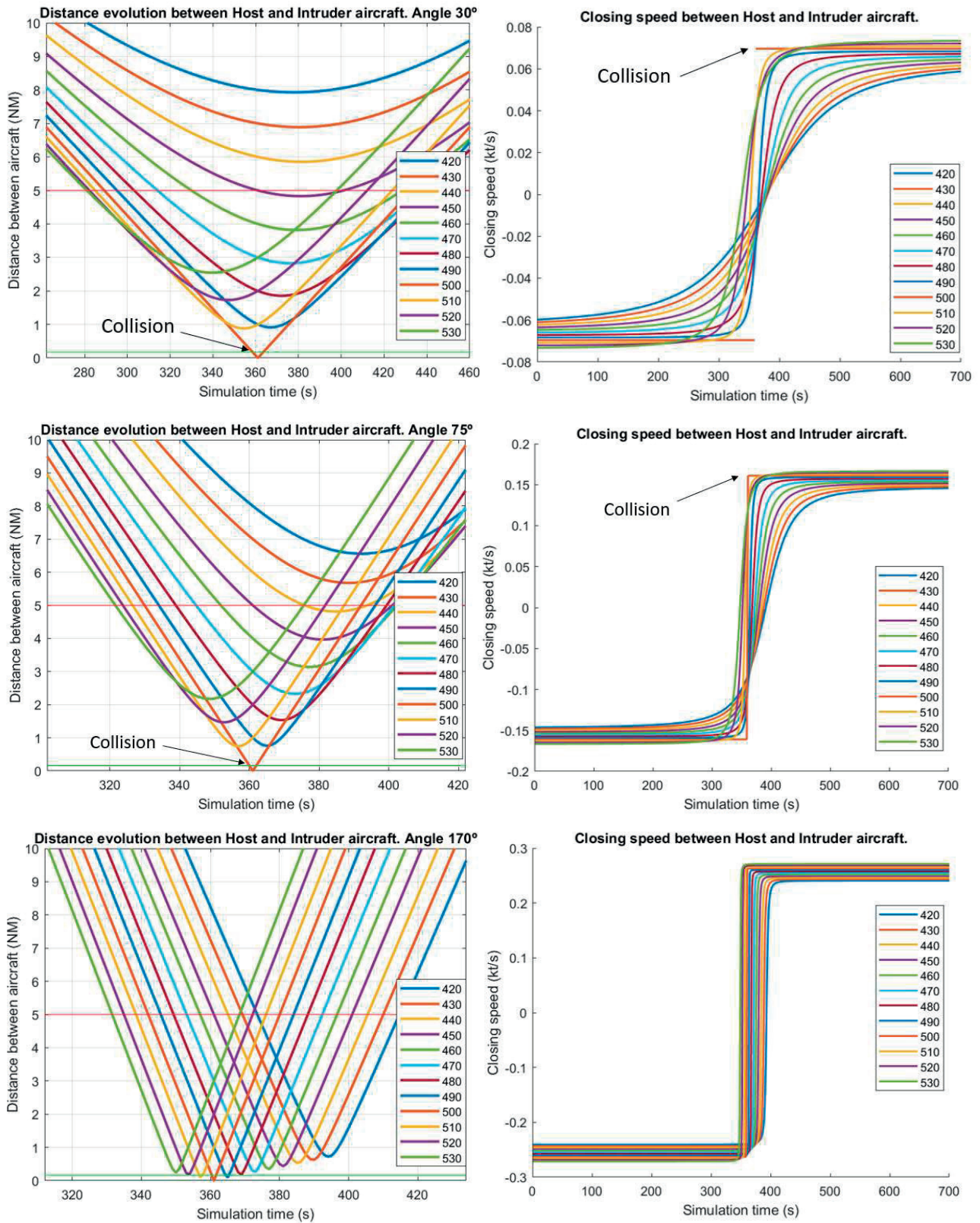


Figure 6: Influence on changing intruder velocity for chosen intersection angles 30<sup>0</sup>, 75<sup>0</sup> and 170<sup>0</sup>.

Also, from Figure 7 (left) it can be seen that duration of conflict situations is depending on speed but also, that for certain speeds conflict doesn't exist (cases with smaller intersection angles).

Figure 7 (right) presents closing speed – it can be seen that by increasing the speed a change of closing speed is becoming more vertical (especially in case of angle  $170^{\circ}$ ), while in case of potential collision is pure vertical (step change).



**Figure 7:** left) Distance variation between aircraft for different intruder velocities, right) Closing speed for different intruder velocities



Difference between the adjacent elements of the closing speed vector for an angle of 75° between airways has been plotted (Figure 8). It is shown the peaks of distributions at different times, but also huge increase in value in case of potential collision.

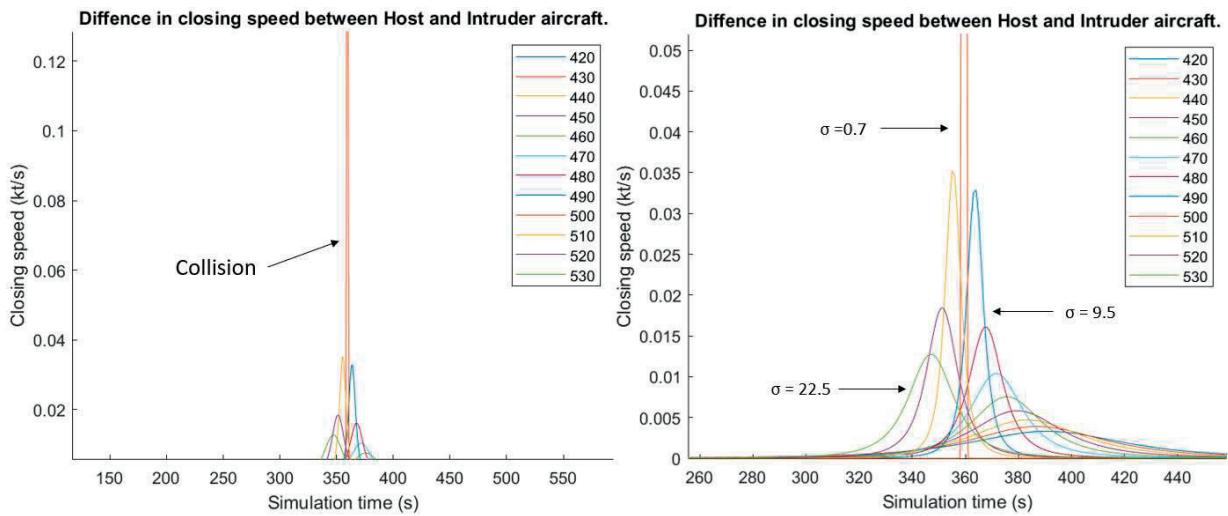


Figure 8: a) Difference between the adjacent elements of the closing speed vectors, b) Zoom on smaller angles

#### 4. RESULTS

Results of first scenario-based experiment combining changes of intersection angle with some changes of intruder velocities, relative to separation minima value of 5 NM are presented in Table 1. It can be observed that conflict could emerge in case of H2H encounter no matter what the velocities of both aircraft are, while in case of crossing encounters conflict emerge in cases when intruder velocity is equal or higher than host velocity. Finally, in case of overtaking encounter conflict emerge only in situation in which intruder is faster for more than 40 kt.

Table 1: Results of first experiment.  $V_h$  and  $V_i$  refer to the speed of Host and Intruder aircraft.

		$V_h (450) > V_i (430)$	$V_h = V_i (450)$	$V_h (450) < V_i (470)$	$V_h (450) < V_i (490)$
		SCENARIO1 MODIFYING V2	SCENARIO1 MODIFYING V2	SCENARIO 1 (PURELY)	SCENARIO1 MODIFYING V2
Overtaking (20° - 40°)	Conflict	NO	NO	NO	YES
	Criticality	NO	NO	NO	LOW (0.26 - 0.28)
	TimetoCPA (s)	NO	NO	NO	72 - 38
	Duration (s)	NO	NO	NO	145 - 76
	Gradient TCPA (NM/s)	NO	NO	NO	0.0183 - 0.0360
Crossing (60° - 90° - 120°)	Conflict	NO	YES (from 120°)	YES (from 90°)	YES
	Criticality	NO	LOW (0.2)	LOW - MEDIUM (0.17 - 0.41)	MEDIUM (0.33 - 0.46 - 0.61)
	TimetoCPA (s)	NO	13	15 - 18	28 - 22 - 20
	Duration (s)	NO	26	30 - 35	56 - 44 - 39
	Gradient TCPA (NM/s)	NO	0.0685	0.0537 - 0.1083	0.0588 - 0.0986 - 0.1448
H2H (145° - 175°)	Conflict	YES	YES	YES	YES
	Criticality	HIGH (0.38 - 0.91)	HIGH (0.51 - 0.93)	HIGH (0.64 - 0.94)	HIGH (0.76 - 0.96)
	TimetoCPA (s)	17 - 20	18 - 19	19 - 19	19 - 18
	Duration (s)	33 - 40	36 - 38	37 - 39	38 - 37
	Gradient TCPA (NM/s)	0.1122 - 0.2218	0.1397 - 0.2321	0.1621 - 0.2472	0.2004 - 0.2553

Results of second scenario-based experiment, relative to separation minima value of 5 NM are presented in Table 2. Similar behaviour is observed like in first experiment, related to emergence of conflict situations.

**Table 2:** Results of second experiment.  $V_h$  and  $V_i$  refer to the speed of Host and Intruder aircraft.

		$V_h=450$ kt $V_i=420$ kt	$V_h=450$ kt $V_i=440$ kt	$V_h=450$ kt $V_i=450$ kt	$V_h=450$ kt $V_i=470$ kt	$V_h=450$ kt $V_i=500$ kt	$V_h=450$ kt $V_i=520$ kt
Overtaking (30°) SCENARIO 2 (changing intersection angle)	Conflict	NO	NO	YES	YES	YES	YES
	Criticality	NO	NO	0.03	0.43	1	0.65
	TimetoCPA (s)	NO	NO	19	61	71	64
	Duration (s)	NO	NO	38	123	142	129
	Gradient TCPA (NM/s)	NO	NO	0.0081	0.0348	0.0696	0.0502
Crossing (75°) SCENARIO 2 (PURELY)	Conflict	NO	YES	YES	YES	YES	YES
	Criticality	NO	0.03	0.2	0.5	1	0.7
	TimetoCPA (s)	NO	9	20	28	31	29
	Duration (s)	NO	17	40	56	62	57
	Gradient TCPA (NM/s)	NO	0.0199	0.0517	0.0941	0.1610	0.1204
H2H (170°) SCENARIO 2 (changing intersection angle)	Conflict	YES	YES	YES	YES	YES	YES
	Criticality	0.85	0.89	0.91	0.94	1	0.96
	TimetoCPA (s)	20	19	19	19	19	19
	Duration (s)	40	39	39	38	38	37
	Gradient TCPA (NM/s)	0.2085	0.2232	0.2272	0.2376	0.2629	0.2521

## 5. CONCLUSION

Characterisation of aircraft encounters is used in order to learn how some important factors influence emergence of potential conflicts and collisions. Experimental results on a given scenario have shown that most critical situations which could easily lead from conflict to collision are with higher intersection angles, i.e. H2H encounters, but also crossing encounters with intruder velocity equal or higher than host aircraft velocity.

Findings could not be generalised but are presenting a solid foundation for further modelling of collision risk as a function of variable separation minima and different surveillance systems. Apart from that, findings will be used for definition of rules for creation of ATCo real-time decision supporting tool.

## Acknowledgement

The work of F. Netjasov was partially supported by the Project number 36033 commissioned by the Ministry of Education, Science and Technological development of the Republic of Serbia.

## REFERENCES

- [1] Netjasov F., Vidosavljevic A., Tomic V., Everdij M., Bloom H. (2013). Development, validation and application of stochastically and dynamically coloured Petri net model for ACAS operations for safety assessment purposes, *Transportation Research Part C - Emerging Technologies*, Vol. 33, pp. 167-195.
- [2] ICAO (2020). Annex 11 – Air Traffic Services (15th edition, Amendment 52), International Civil Aviation Organization, Montreal, Canada.
- [3] Serrano-Mira L., Pérez Sanz L., Pérez-Castán J. A. (2021). *Ad hoc minimum separation: a challenge for air traffic control*. Proceedings of 14th Conference on Transport Engineering (CIT 2021), Burgos, Spain, pp. 2803-2814. (<https://doi.org/10.36443/9788418465123>).
- [4] ICAO (2009). Cir. 319 - A Unified Framework for Collision Risk Modelling in Support of the Manual on Airspace Planning Methodology for the Determination of Separation Minima (Doc 9869). International Civil Aviation Organization, Montreal, Canada.
- [5] ICAO (2013). Annex 19 - Safety Management (1<sup>th</sup> edition), International Civil Aviation Organization, Montreal, Canada.
- [6] Netjasov F., Crnogorac D., Pavlović G. (2019). Potential safety occurrences as indicators of air traffic management safety performance: A network based simulation model, *Transportation Research Part C - Emerging Technologies*, Vol. 102, pp. 490-508.
- [7] ICAO (2020). Procedures for Air Navigation Services – Air Traffic Management (Doc. 4444, 16th edition, Amendment 9), Montreal, Canada.

## ANALIZA METEOROLOŠKIH USLOVA ZA POTENCIJALI RAZVOJ AERODROMA

### ANALYSIS OF METEOROLOGICAL CONDITIONS FOR POTENTIAL AIRPORT DEVELOPMENT

NIKOLA LUKAČEVIĆ<sup>1</sup>, BOJANA MIRKOVIĆ<sup>1</sup>, ANA UZELAC<sup>1</sup>, KATARINA KUKIĆ<sup>1</sup>

<sup>1</sup>Univerzitet u Beogradu – Saobraćajni fakultet, nikola.lukacevic@sf.bg.ac.rs

<sup>2</sup>Univerzitet u Beogradu – Saobraćajni fakultet, b.mirkovic@sf.bg.ac.rs

<sup>3</sup>Univerzitet u Beogradu – Saobraćajni fakultet, ana.uzelac@sf.bg.ac.rs

<sup>4</sup>Univerzitet u Beogradu – Saobraćajni fakultet, k.mijailovic@sf.bg.ac.rs

**Rezime:** *Vetar na lokaciji je jedan od ključnih faktora za utvrđivanje pravca pružanja poletno-sletne staze novog aerodroma. Cilj je da se omogući što veći koeficijent upotrebljivosti u odnosu na bočni vetar, ali tako da se postigne i kompromis sa ostalim uslovima od značaja. Za planiranje aerodroma se koriste podaci o vetru za duži niz godina, koji su snimani na najbližoj meteorološkoj stanici. Za potrebe analize potencijala za razvoj novog aerodroma na jednoj lokaciji u Opštini Bar, prikupljeni su podaci za period 2015-2021.godine, na meteorološkoj stanici Dobra Voda. Nakon detaljnog pročišćavanja podataka izdvojen je neprekinut niz od dve godine koji se može koristiti za dalju analizu. Na osnovu ruža vetrova dat je predlog orijentacije poletno-sletne staze, izračunat koeficijent upotebljivosti u odnosu na bočni vetar, i određen dominantan smer poletno-sletne staze.*

**Ključne reči:** *aerodrom, meteorološki uslovi, ruža vetrova, koeficijent upotrebljivosti*

**Abstract:** *Wind is one of the key factors influencing runway orientation at a new airport. The aim is to achieve high usability factor with respect to cross wind, having an appropriate trade-off with other significant location conditions. For airport planning purposes, wind data should be collected over longer time horizon at the nearest meteorological station. For the analysis of one location in Bar Municipality, for possible new airport development, wind data were collected for the period 2105-2021, at meteorological station Dobra Voda. After detailed cleaning of the database, usable unbroken two-year period is extracted for further analysis. Based on the wind roses runway orientation is proposed, utilization factor calculated and dominant runway direction is determined.*

**Keywords:** *airport, meteorological conditions, wind rose, usability factor*

## 1. UVOD

Meteorološki uslovi predstavljaju bitan faktor kod izbora lokacije novog aerodroma. Pojave vetra, oblačnosti i vidljivosti imaju uticaj na koeficijent upotrebljivosti aerodroma, a time i na veličinu i konfiguraciju aerodroma, kao i na vrstu navigacionih sredstava kojima treba opremiti aerodrom da bi se ovaj koeficijent povećao.

Dok se na koeficijent upotrebljivosti u odnosu na oblačnost i vidljivost može uticati unapređenjem navigacione opreme, koeficijent upotrebljivosti u odnosu na vetar isključivo zavisi od toga kako se u fazi planiranja orijentiše poletno-sletna staza (PSS). Utvrđivanje pravca pružanja PSS novog aerodroma ima za cilj postizanje što većeg koeficijenta upotrebljivosti u odnosu na bočni vetar, ali tako da se ostvari kompromis sa ostalim uslovima lokacije, i to topografskim uslovima koji direktno utiču na bezbednost odvijanja saobraćaja, ali i ekološkim, tj. izloženosti obližnjih naselja buci i štetnim gasovima.

Koficijent upotrebljivosti aerodroma u odnosu na vetar predstavlja odnos vremena u toku godine za koje je na aerodromu moguće letenje, u zavisnosti od bočnog vetra, prema ukupnom vremenu godine. Preporuka je da minimalni koeficijent upotrebljivosti na aerodromima za javni saobraćaj bude 95% [1, 2]. Teži se da se povoljan koeficijent postigne izgradnjom samo jedne adekvatno orijentisane PSS. Međutim, u slučaju da se traženi stepen otvorenosti ne može ostvariti na taj način, potrebno je planirati još jednu (vetrenu) PSS.

Kod izrade studije lokacije za novi aerodrom često se javlja problem nedostatka dovoljno reprezentativnih podataka za tu lokaciju. Meteorološki podaci treba da su izmereni u skladu sa propisima Svetske meteorološke organizacije za osmatranja na sinoptičkim stanicama (osmatranja na svaka tri časa ili češće). Na mnogim lokacijama koje se ispituju ne postoje meteorološke stanice, u kom slučaju se mogu koristiti podaci

sa najbliže meteorološke stanice, ali se pre toga mora utvrditi da nema bitnih razlika u meteorološkim pojavama na dve lokacije.

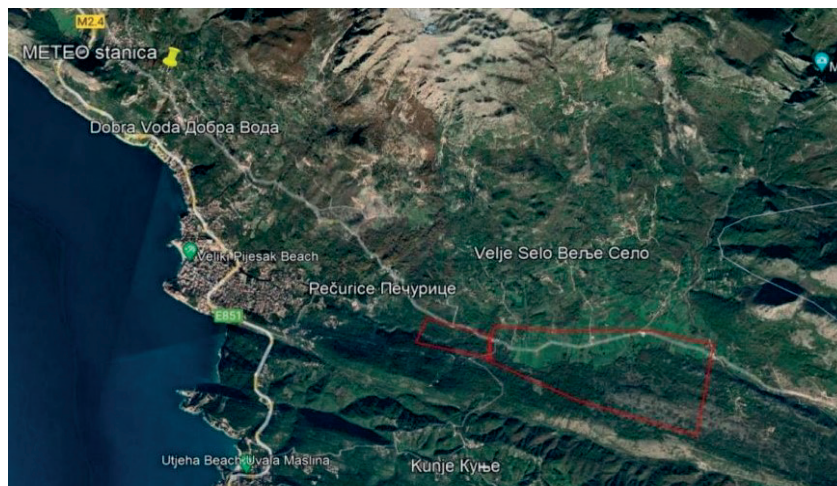
U ovom radu opisan je postupak pripreme podataka za analizu meteoroloških uslova jedne lokacije (Mrkovsko Polje) na teritoriji Opštine Bar, Crna Gora, i proračun koeficijenta upotrebljivosti za predloženu orijentaciju nove PSS. U poglavlju 2 opisan je postupak prečišćavanja sirovih podataka koje je meteorološka stanica beležila u periodu od 2015. do 2021. godine, da bi se došlo do upotrebljivih informacija o pojavama vetra na lokaciji koja se ispituje. Nakon što je izdvojen upotrebljiv niz podataka, u Poglavlju 3 je određen koeficijent iskorišćenja u odnosu na bočni vetar za predloženu orijentaciju PSS, i utvrđen dominantni smer PSS u upotrebi. U Poglavlju 4 su dati zaključci rada i dalji koraci.

## 2. PRIPREMA PODATAKA ZA ANALIZU

Za potrebe analize meteoroloških uslova lokacije, tj. za određivanje orijentacije i proračuna upotrebljivosti PSS potrebni su podaci o smeru, brzini i učestanosti vetra, poželjno za period od pet godina. Na osnovu ovih podataka formiraju su ruže vetrova po maksimalnoj brzini i po učestanosti. Kasnije, nakon utvrđivanja tačnog pravca pružanja PSS određuju se i koeficijenti upotrebljivosti aerodroma u odnosu na bočni vetar, kao i dominantni smer u upotrebi.

Za analizu meteoroloških uslova lokacije, korišćeni su podaci prikupljeni na najbližoj meteo stanici, koja se nalazi na lokaciji Dobra Voda, sa koordinatama 42°03'16,6" N, 19°08'27,5" E. Na slici 1 prikazana je lokacija meteo stanice (žuti pin) i okvirno područje novog aerodroma (crvene granice - "Mrkovsko polje"). Meteo stanica je locirana na oko 6km severo-zapadno od analizirane lokacije, zbog čega se vetar može delimično razlikovati od vetra na samoj lokaciji. U nekoj od narednih faza planiranja i projektovanja novog aerodroma, potrebno je postaviti mobilnu meteo stanicu na samu lokaciju, snimati podatke određeni period vremena, i uporediti ih sa podacima sa meteo stanice Dobra Voda, kako bi se utvrdila odstupanja i, ako je potrebno, izvršila korekcija proračuna.

Prikupljeni podaci mogu biti u neadekvatnom obliku, stoga je prvi korak proučiti podatke s kojima se namerava raditi i dovesti ih u oblik prikladan za rad. Podacima mogu da nedostaju vrednosti, mogu da sadrže suvišne vrednosti, vrednosti mogu da budu u neodgovarajućem obliku, i slično. Pročišćavanje podataka u ovom radu izvršeno je programom pisanim u Python-u.



**Slika 1:** Lokacija meteo stanice u odnosu na okvirnu lokaciju novog aerodroma [Google Earth]

Meteo stanica u Dobro Vodi snimala je podatke u periodu od 2015. do 2021. godine [3]. Inicijalno je zabeleženo 162.294.880 podataka, koji su raspoređeni u 1.341.280 vrsta i 121 kolonu. S obzirom da je meteo stanica podešena tako da beleži podatke na svakih 1 minut, primećeno je da je konačni broj snimljenih podataka značajno manji od očekivanog broja podataka (381.934.080). Ova razlika jasno je ukazala na to da je došlo do greške prilikom snimanja i da je značajan broj podataka izgubljen. Za preciznu i upotrebljivu analizu nije bitna količina podataka, koliko njihov kvalitet, stoga je bilo neophodno utvrditi u kojoj meri su snimljeni podaci odgovarajući, kao i ispitati mogućnost njihovog korišćenja u analizi. Za pouzdanost analize važno je da postoji neprekinut niz podataka, snimanih u istom koraku (vremenskom intervalu). Takođe, važno je da uzorak pokriva isti broj meseci u godini, na primer pet januara, pet avgusta itd.

Radi postizanja veće preglednosti i radi lakšeg manipulisanja sa podacima, prvo su odbačeni svi nepotrebni podaci. Svaka kolona snimljene baze podataka meteo stanice predstavljala je neku meteorološku karakteristiku kao npr. vlažnost vazduha, prosečna, minimalna i maksimalna temperatura, prosečna, minimalna i maksimalna brzina, smer vetra, padavine itd. Za dalju analizu potrebni su podaci o smeru i prosečnoj brzini vetra i iz toga razloga korišćenjem kôda u Pythonu zadržane su samo ove dve kolone (nova baza podataka sadrži 1.341.280 vrsta i 2 kolone).

Nakon izdvajanja podataka o vetru bilo je potrebno utvrditi da li su podaci snimani za svaki mesec i svaku godinu, kao i da li su podaci za svaki mesec potpuni. Uz pomoć kôda u Pythonu koji broji podatke za svaki mesec i godinu utvrđeno je da za određene periode podaci ne postoje ili su nepotpuni. To su april i maj 2015, period od oktobra 2017. do avgusta 2021. i decembar 2021. godine, i podaci za ove periode se ne mogu koristiti u analizi. Za dalju analizu korišćen je period od septembra 2015. do avgusta 2017. Ovaj period je najduži mogući neprekidni period za koji su podaci potpuni, a pritom se svaki mesec u godini pojavljuje podjednak broj puta.

Prethodnim brojanjem podataka za mesec/godinu (09/2015-08/2017) uočeno je da postoje značajne razlike u broju prebrojanih podataka. Tako, na primer, za septembar 2015. imamo 4.320 podataka, 86.400 podataka za septembar 2016, a za septembar 2017. 41.760 podataka. Analizom je utvrđeno da meteo stanica nije snimala podatke u jednakim vremenskim intervalima. Opažaji vetra beleženi su, u zavisnosti od meseca u godini, na svakih 1 minut ili na svakih 10 minuta. Takođe, utvrđeno je da određen broj podataka dupliran (za isti vremenski trenutak postoje dva ista ili dva različita podatka o vetru). Kako bi dalja analiza bila relevantna, napravljen je kôd koji briše sve duplirane podatke i koji uzima u obzir samo opažaje vetra beležene na svakih 10 minuta. Novoformirana baza podataka je konačna i sadrži 210.514 podataka (105.257 vrsta i 2 kolone). Ona sadrži najduži neprekinuti niz za koji su podaci potpuni, a da se pritom svaki mesec u godini pojavljuje podjednak broj puta.

Na osnovu konačne baze podataka formirana je tabela brzina-smer-učestanost, prikazana na slici 2, na osnovu koje je dalje rađena analiza meteoroloških uslova lokacije aerodroma. Broj opažaja u svakom polju je dobijen prebrojavanjem opažaja u određenom opsegu brzine i u određenom opsegu smera iz kojeg vetar duva. Svaki smer obuhvata opseg od  $22,5^{\circ}$ . Tako na primer, smer istok (E) obuhvata opseg  $78,75^{\circ}$ - $101,25^{\circ}$ .

	N	NNE	NE	ENE	E	ESE	SE	SSE	S	SSW	SW	WSW	W	WNW	NW	NNW
(0,1)	898	925	1026	833	697	551	454	453	375	388	420	513	647	928	1072	928
(1,2)	3531	5216	4149	3096	2782	1751	1426	2021	1707	1098	1005	1540	2009	3236	3289	2882
(2,3)	2504	5451	607	1029	1807	1521	1003	1710	1136	470	318	656	984	3518	2520	1593
(3,4)	1263	651	108	181	794	1527	743	574	283	63	32	83	216	1752	1788	1024
(4,5)	409	143	97	146	569	1605	675	214	78	16	9	14	37	369	926	514
(5,6)	230	71	99	148	386	1263	434	73	18	1	0	0	5	68	394	343
(6,7)	162	52	92	148	204	634	334	37	4	1	1	0	3	21	170	241
(7,8)	150	32	74	156	148	452	232	11	1	0	1	0	0	6	92	135
(8,9)	125	30	33	117	103	287	204	2	0	0	0	0	0	2	43	91
(9,10)	112	19	18	58	72	131	162	0	0	0	0	0	0	1	27	29
(10,11)	67	15	7	51	31	87	109	2	0	0	0	0	1	0	12	31
(11,12)	44	6	3	18	9	48	60	1	0	0	0	0	0	0	10	16
(12,13)	28	8	0	9	3	20	24	0	0	0	0	0	0	0	8	10
(13,14)	28	2	0	3	1	3	4	0	0	0	0	0	0	0	0	4
(14,15)	28	1	0	2	0	0	2	0	0	0	0	0	0	0	1	6
(15,16)	21	2	0	0	0	0	0	0	0	0	0	0	0	0	1	3
(16,17)	16	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(17,18)	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
(18,19)	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(19,20)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(20,21)	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(21,22)	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Slika 2: Brzina-smer-učestanost vetra, pročišćeni podaci, septembar 2015. - avgust 2017. godine

### 3. ANALIZA METEOROLOŠKIH USLOVA

Prilikom analize uticaja vetra na operacije vazduhoplova posmatraju se dve komponente vetra. Prva komponenta vetra je upravna na osu PSS i naziva se bočna komponenta vetra. Bočna komponenta vetra utiče negativno na vazduhoplove u poletanju i sletanju. U trenutku odlepljivanja točkova glavnog stajnog trapa pri poletanju, odnosno neposredno pri dodiru točkova glavnog stajnog trapa sa PSS pri sletanju, pri određenoj jačini bočnog vetra, može doći do nezgode, pa i udesa usled toga što vazduhoplov može biti smaknut sa PSS [4]. Bočni vetar utiče i na konfiguraciju aerodroma na taj način što određuje orijentaciju, a ponekad i broj PSS. Druga komponenta vetra je u pravcu PSS i može biti čeon ili repna. Operacije poletanja i sletanja se po pravilu odvijaju sa čeonim vetrom, što ima direktan uticaj na skraćenje dužine PSS za

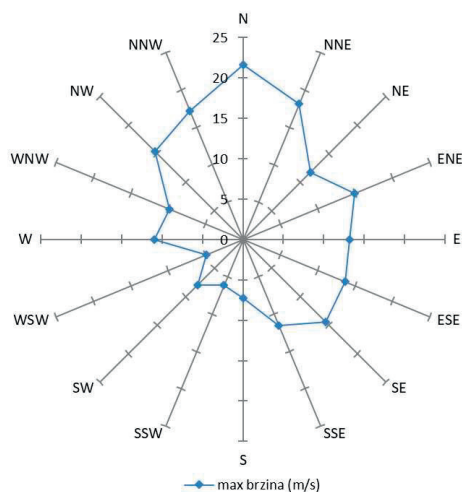
poletanje/sletanje. Dopušteno je da se operacije obavljaju i uz slab repni vetar. Analizom komponente vetra duž PSS može se utvrditi koji će biti dominantan smer u upotrebi nove PSS.

### 3.1. Ruže vetrova po maksimalnoj brzini i učestanosti

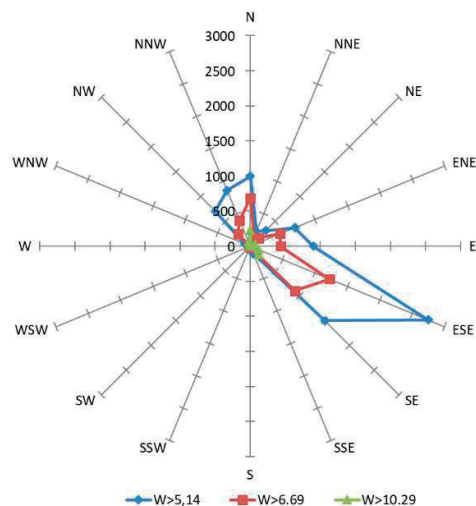
Na osnovu prethodno očišćenih podataka (slika 2) formiraju su ruže vetrova po maksimalnoj brzini i po učestanosti. Na osnovu ruža vetrova možemo odrediti poželjan pravac pružanja PSS u odnosu na vetar. PSS treba da se orijentiše tako da se izbegnu najjači i najučestaliji bočni vetrovi. Ruža vetrova po maksimalnoj brzini se dobija tako što se na svaki od 16 geografskih smerova nanosi maksimalna brzina vetra koja je osmotrena iz tog smera, slika 3.

Međunarodna regulativa (EASA [1] i ICAO [2]) propisuje dozvoljenu bočnu komponentu vetra u zavisnosti od referentne dužine PSS. Dozvoljne bočne komponente vetra su 20kts za PSS referentne dužine 1500m i veće, 13kts za PSS referentne dužine od 1200m do 1500m i 10 kts za PSS referentne dužine manje od 1200m. Konstruisane ruže vetrova po učestanosti za sve tri dozvoljene bočne komponente prikazane su na slici 4. Ruže vetrova po učestanosti konstruišu se za brzine veće od dozvoljenih bočnih komponenti, jer ove pojave utiču na upotrebljivost PSS.

Prema podacima sa meteo stanice Dobra Voda najintezivniji vetrovi su zabeleženi iz smera N, NNE, NNW, dok je najveći broj opažaja vetra jačine veće od graničnih komponenti bočnog vetra osmotren iz smera ESE. Na osnovu ruža vetrova po maksimalnoj brzini i po učestanosti, može se zaključiti da bi najpoželjniji pravac pružanja PSS u odnosu na vetar bio u pojasu ESE-WNW.



Slika 3: Ruža vetrova po maksimalnoj brzini



Slika 4: Ruža vetrova po učestanosti

### 3.2. Koeficijent upotrebljivosti aerodroma u odnosu na vetar

Pravac pružanja PSS se određuje osnovu ruže vetrova, ali i drugih ograničenja kao što su prirodne i veštačke prepreke, karakteristike terena, blizina naselja, itd. U praksi se najčešće ispituje više pravaca pružanja PSS. U primeru potencijalnog aerodroma na lokaciji u Opštini Bar, uzimajući u obzir oblik definisanog područja (slika 1), prirodne prepreke (teren) u okolini lokacije i ružu vetrova predložena je orijentacija PSS 110°/290°. Maksimalna dužina PSS sa pratećim zaštitnim površinama (osnovna staza i zaštitna površina kraja poletno-sletne staze) u okviru definisanog područja je 2250m. Uzimajući u obzir nadmorsku visinu lokacija (240m) i referentnu temperaturu (31°C), referentna (osnovna) dužina PSS je 1.812m. Zbog toga je proračunat koeficijent upotrebljivosti za sve tri granične bočne komponente vetra.

Koeficijent upotrebljivosti aerodroma u odnosu na vetar se računa kao:

$$U(\%) = \frac{n}{N} \cdot 100 = \frac{(N-n_1)}{N} \cdot 100 \quad (1)$$

gde je:

N – ukupan broj osmatranja,

$n$  – ukupan broj osmatranja u kojima je  $W_b \leq W_{bdoz}$ ,  
 $n_1$  – ukupan broj osmatranja u kojima je  $W_b \geq W_{bdoz}$ .

Za proračun koeficijenta upotrebljivosti potrebno je za svaki geografski smer utvrditi broj nepovoljnih opažaja vetra ( $n_1$ ). Prvo se za svaki smer određuje intenzitet vetra  $W$  iz tog smera koji daje  $W_{bdoz}$ , gde je  $\alpha$  ugao između ose PSS i odgovarajućeg geografskog smera

$$W = \frac{W_{bdoz}}{\sin \alpha} \quad (2)$$

Vrednosti ugla  $\alpha$  i intenziteta vetra  $W$  koji iz svakog smera proizvodi dozvoljene bočne komponente od 5,14; 6,69 i 10,29 m/s (tabela 1). Na osnovu podataka iz tabele 1 za svaki geografski smer dobija se broj nepoželjnih pojava vetra koje generišu  $W_b > W_{bdoz}$ , koje su prikazane u tabeli 2.

**Tabela 1:** Intenzitet vetra koji proizvodi graničnu bočnu komponentu na PSS sa orijentacijom 110°/290°

Geografski pravac	$\alpha$	$\sin \alpha$	5,14/ $\sin \alpha$	6,69/ $\sin \alpha$	10,29/ $\sin \alpha$
N-S	70	0,940	5,47	7,12	10,95
NNE-SSW	92,5	0,999	5,14	6,70	10,30
NE-SW	115	0,906	5,67	7,38	11,35
ENE-WSW	137,5	0,676	7,61	9,90	15,23
E-W	160	0,342	15,03	19,56	30,09
ESE-WNW	2,5	0,044	117,84	153,37	235,90
SE-NW	25	0,423	12,16	15,83	24,35
SSE-NNW	47,5	0,737	6,97	9,07	13,96

Na osnovu podataka iz tabele 2, za pravac pružanja PSS 110°/290°, koeficijenti upotrebljivosti PSS u odnosu na vetar u odnosu na sve tri bočne komponente su preko preporučenih 95% (tabela 3).

**Tabela 2:** Broj nepovoljnih opažaja iz svakog smera, celobrojne vrednosti

Smer vetra	N	NNE	NE	ENE	E	ESE	SE	SSE	S	SSW	SW	WSW	W	WNW	NW	NNW
$W_b > 5,14$	921	231	260	319	0	0	26	17	14	2	2	0	0	0	9	333
$W_b > 6,69$	619	134	107	89	0	0	0	3	1	0	1	0	0	0	0	98
$W_b > 10,29$	186	32	2	0	0	0	0	0	0	0	0	0	0	0	0	10

**Tabela 3:** Broj nepovoljnih opažaja i koeficijent upotrebljivosti za sve tri bočne komponente vetra

	$W_b > 5,14$ m/s	$W_b > 6,69$ m/s	$W_b > 10,29$ m/s
$n_1$	2133,40	1051,14	230,94
U%	97,97%	99,00%	99,78%

### 3.3. Utvrđivanje dominantnog smera u upotrebi

Na osnovu ruža vetrova po učestanosti, indikativno je da će dominantan smer korišćenja PSS biti sa zapada na istok, odnosno sa vode ka kopnu, jer će se u tom smeru operacije odvijati sa dominantno čeonim vetrom. Vetar koji proizvodi čeonu komponentu na PSS javlja se iz smera +/- 90 stepeni u odnosu na smer PSS koji se ispituje. Pored toga, operacije se mogu obavljati i sa repnim vetrom, ali prema preporukama ICAO do 5kts [5].

Procenat vremena tokom kojeg se dominantan smer može koristiti u odnosu na komponentu vetra duž PSS računa se kao:

$$D_{deg}(\%) = \frac{N - n_2}{N} \cdot 100 \quad (3)$$

gde je:

$N$  – ukupan broj osmatranja,  
 $n_2$  – ukupan broj osmatranja u kojima je  $W \geq 5$ kts.

Za proračun procenta vremena tokom kojeg se dominantan smer može koristiti u odnosu na komponentu vetra duž PSS, potrebno je za svaki smer utvrditi broj nepovoljnih opažaja vetra ( $n_2$ ). Za svaki geografski smer izračunat je intenzitet vetra  $W$  koji iz tog smera proizvodi repni vetar jači od 5kts (2,57 m/s):

$$W = \frac{W_{rdoz}}{\cos\alpha} \quad (4)$$

Vrednosti ugla  $\alpha$  i inteziteta vetra  $W$  koji proizvodi dozvoljenu komponentu repnog vetra od 2,57 m/s, date su u tabeli 4. Prikazani su samo smerovi iz kojih se može pojaviti repni vetar na smer upotrebe PSS od 110°.

**Tabela 4:** Intenzitet vetra koji proizvodi dozvoljenu repnu komponentu vetra u smeru 110°

Smer vetra	$\alpha$	$\cos\alpha$	$2,57/\cos\alpha$	$n_2$
N	110	-0,34	-7,5	559,00
SSW	92,5	-0,04	-58,96	0
SW	115	-0,42	-6,09	1,91
WSW	137,5	-0,74	-3,49	56,45
W	160	-0,94	-2,74	520,73
WNW	182,5	-0,99	-2,57	3716,08
NW	205	-0,91	-2,84	3880,52
NNW	227,5	-0,68	-3,81	1621,59

Ukupan broj opažaja vetra koji rezultiraju repnom komponentom većom od 5kts na smer 110° je 10.356,3. Na osnovu toga se dobija da se smer 110° može koristiti 90,16% vremena. U kontekstu položaja naselja Dobra Voda, ovo znači da bi se operacije sletanja odvijale preko naselja Dobra Voda, dok bi se poletanja odvijala ka Albaniji. To je povoljnija varijanta jer se operacije poletanja doživljavaju kao bučnije, a pored toga promenom procedura u prilazu je moguće dodatno smanjiti izloženost naselja buci aviona prilikom sletanja.

#### 4. ZAKLJUČNA RAZMATRANJA

U ovom radu prikazana je analiza meteoroloških uslova lokacije, u svrhu ispitivanja potencijala za izgradnju novog aerodroma na teritoriji Opštine Bar. Na osnovu podataka prikupljenih sa meteo stanice Dobra Voda, ustanovljeno je da su meteorološki uslovi lokacije Mrkovsko polje izuzetno dobri, sa koeficijentima upotrebljivosti preko 97% u odnosu na bočni vetar. Ruža vetrova je takva da se jasno izdvaja dominantan smer u upotrebi sa zapada na istok. Na osnovu prikazane analize, preliminarno odabrana lokacija u Opštini Bar se može oceniti kao povoljna za razvoj aerodroma sa aspekta meteoroloških uslova. Osim utvrđenog koeficijenta upotrebljivosti u odnosu na bočni vetar, i utvrđenog dominantnog smera u upotrebi, na osnovu raspoloživih podataka takođe je moguće uraditi analizu i po periodima dan/noć, letnja/zimska sezona, itd. za potrebe sledećih faza projektovanja aerodroma.

#### ZAHVALNICA

Ovo istraživanje podržalo je Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije.

#### LITERATURA

- [1] European Union Aviation Safety Agency (EASA), *Easy Access Rules for Aerodromes (Regulation (EU) No 139/2014)*, peto izdanje, jun 2021.
- [2] International Civil Aviation Organization (ICAO), *Annex 14, Volume I – Aerodrome Design and Operations*, osmo izdanje, jul 2018, ažuriran decembar 2020.
- [3] Meteorološki podaci sa meteorološke stanice Dobra Voda za period 2015. do 2021. godina
- [4] Mirković B., Tošić V., Babić O., *Vazduhoplovna pristaništa – Praktikum*, Saobraćajni Fakultet, Univerzitet u Beogradu, Beograd, 2010.
- [5] International Civil Aviation Organization (ICAO), *Doc 9157 - Aerodrome Design Manual, Part 1 – Runways*, četvrto izdanje, 2020.



## INFLUENCE OF AUTOMATION CHANGES ON GO-AROUND OPERATIONS: A DATA-DRIVEN APPROACH

DOROTEJA TIMOTIC<sup>1</sup>, BOJANA MIRKOVIC<sup>1</sup>, IRENE BUSELLI<sup>2</sup>, FEDJA NETJASOV<sup>1</sup>, CHEN XIA<sup>3</sup>, CHRISTIAN EDUARDO VERDONK GALLEG<sup>3</sup>

<sup>1</sup>University of Belgrade – Faculty of Transport and Traffic Engineering, Division of Airports and Air Traffic Safety  
[d.timotic@sf.bg.ac.rs](mailto:d.timotic@sf.bg.ac.rs), [b.mirkovic@sf.bg.ac.rs](mailto:b.mirkovic@sf.bg.ac.rs), [f.netjasov@sf.bg.ac.rs](mailto:f.netjasov@sf.bg.ac.rs)

<sup>2</sup>Zenabyte Srl, Genoa, Italy [irene.buselli@zenabyte.com](mailto:irene.buselli@zenabyte.com)

<sup>3</sup>Centro de Referencia de Investigacion, Desarrollo e Innovacion ATM (CRIDA) A.I.E., Madrid, Spain  
[cxia@e-crida.enaire.es](mailto:cxia@e-crida.enaire.es), [ceverdonk@e-crida.enaire.es](mailto:ceverdonk@e-crida.enaire.es)

**Abstract:** *A variety of causes led air traffic controllers (ATCOs) to instruct go-around procedures. The most frequent are non-stabilized approach, loss of separation or insufficient separation between aircraft, so that speed monitoring or separations in approaches are particularly significant for an improvement in operational safety. Nowadays, ATCOs use the speed parameter provided by their Controller Working Position (CWP) to apply relevant restrictions to aircraft within their responsibility. This parameter is obtained from the radar tracks and it can differ from the actual speed of the aircraft. The objective of this research was to select appropriate factors influencing go-around procedures and compare them before and after the introduction of the speed parameter on the CWP display at Barcelona airport.*

**Keywords:** *air traffic control, airport, safety, go-around*

### 1. INTRODUCTION

Some of the main causes affecting airport safety are the go-around procedures or the losses of separations (LoS) between consecutive aircraft in approach and landing. The statistics show that among worldwide fatal accidents between 2011 and 2020, 54% occurred during the final approach or in the landing phase [1]. According to accident analyses, the go-around procedures are often improperly executed because of their complexity and high tempo stress [2]. A go-around may be initiated by the pilot or the air traffic controller (ATCO) representing a decision not to continue an approach, or not to continue a landing, following procedures to conduct another approach or to divert to another airport [3]. Go-around can happen at any point from the final approach fix to the touchdown due to conditions such as wind shear, runway incursion or unstabilized approach [3]. However, go-around procedure usually occurs at low altitude and low speed, close to the ground, which makes these operations difficult to perform properly. In these situations, the ATCOs have a key role in ensuring the safe and efficient go-around operations. In fact, the ATCOs have to deal with continuously congested airspace with a high level of alertness in order to maintain the safety and efficiency of air traffic operations [4]. On the other hand, the close and timely precise cooperation between pilots and ATCOs is crucial to maintain safety of the approach/landing phase. To make that possible, better support of ATCOs decision making and instructions to aircrew relating to landing approaches is required.

The purpose of the research presented in this paper is to analyse the factors that contribute to go-around occurrences i.e. how they change with introduction of new information/label to the Controller Working Position (CWP). The idea is to analyse operational data in respect to the introduction of the specific information: indicated airspeed (IAS) or Mach number, in the flight label on the radar screen on CWP so that ATCO will have it at his/her disposal and will be able to apply new speed restrictions between pairs of approaching aircraft to maintain safe separation. The analysis is performed before and after the change on tower ATCOs CWP at Barcelona airport using the operational data in 2017 (before change introduction) and 2019 (after change introduction).

The aircraft speed in the final approach is a valuable well-known safety term. Currently, the label of a flight that appears on the CWP (Figure 1-left), shows the horizontal speed term, obtained from the radar track, but does not have to match with the current aircraft speed or IAS. However, the proposal is to add one more term, in the fourth row of the label, showing the IAS or Mach number. Being a known parameter on the aircraft, to be displayed in real time on CWP, it will be sent from the aircraft system as DAP (Downlinked Aircraft Parameters). A similar structure to the horizontal speed is proposed (Figure 1-middle), in dozens of knots with the indication "I" to understand that it is the IAS.

Depending on the flight level, the speed data presented in CWP will be indicated in Mach number or knots (IAS) (Figure 1-right picture). Therefore, if that speed data is not available, the tactical constraints will remain the same, so there will be no impact on the current Human Machine Interface. This change is expected to reduce the number of go-around by non-stabilized approaches or by LoS, among others, as the ATCO will be able to safely separate aircraft based on its IAS. Indirectly and due to the restrictions imposed by the ATCO, these speed changes will affect the pilots that they will have to modify according to the ATCO instructions.



**Figure 1:** Current (left), proposed (middle) flight label and flight label showing MACH on CWP (right)

The paper consists of four sections. The introduction and problem definition is presented in Section 1, as well as the purpose of this paper. In Section 2, five step methodology is presented, followed by the results obtained in each step in Section 3. Section 4 is the conclusion of the paper.

## 2. METHODOLOGY

In order to assess benefits due to introduction of the IAS label on CWP two scenarios are compared – before and after implementation of this change. Recorded operational data for two summer months (June and July) in year 2017 are used to represent before the change scenario. Same months in year 2019 are used to represent for after the change case. The first benefits of this change were expected in 2020, two years after its introduction. But due to COVID-19 pandemic impact on air traffic, year 2020 could not be used as a reference. Two sub-scenarios were planned, corresponding to nominal weather conditions and conditions with severe weather activity (with increased number of go-around).

The analysis is conducted through several steps as following:

- Step I – selection of the sample days for the analysis before and after the change;
- Step II – analysis of runway configuration schemes and available runway throughput;
- Step III – analysis of weather conditions;
- Step IV – extraction of the selected indicators and before and after-change scenario comparison;
- Step V – analysis of other possible contributors to go-around operation.

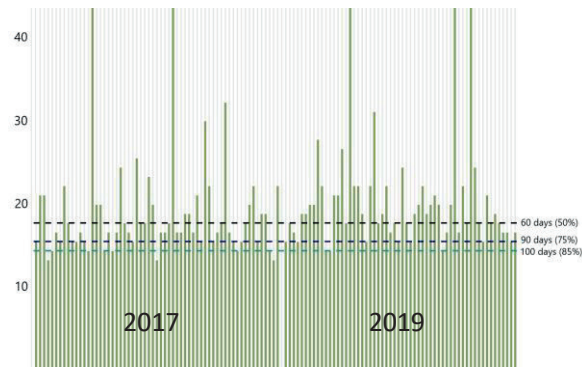
## 3. RESULTS

### 3.1. Step I –sample days

First step was to select the sample days with high traffic occupancy in summer period of 2017 and 2019. The daily number of arrivals for all sample days in both years was required in order to extract days with high occupancy.

The sample is composed by 119 complete days (61 in 2017 and 58 in 2019). To select the high-occupancy days the maximum Arrival Sequencing and Metering Area (ASMA) congestion was computed for each day, and then the days with the highest values were selected. This approach was adopted with the idea to focus more on highest congestion and how it is managed. In the Figure 2, the three dashed lines represent three possible thresholds for selection.

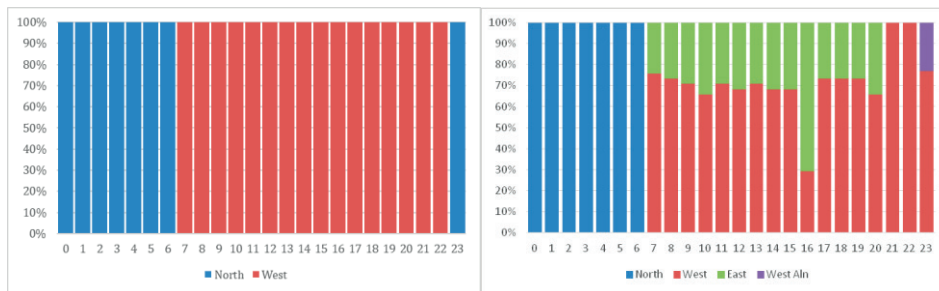
For further analysis the second threshold is adopted. The selected sample includes: 41 days in 2017 (June: 19 days and July: 22 days), and 48 days in 2019 (June: 24 days and July: 24 days).



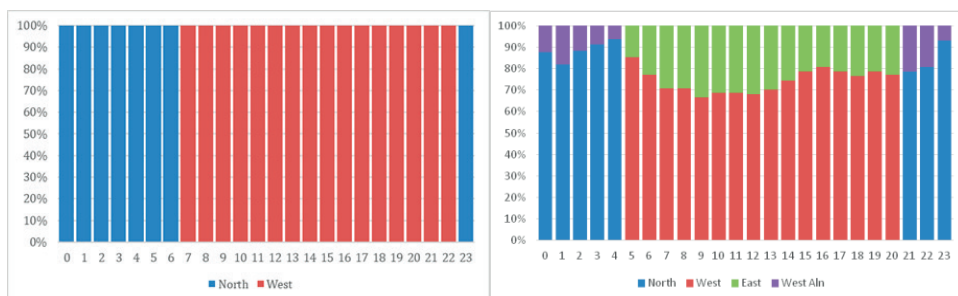
**Figure 2:** The sample days extraction (left) and number of weekdays in the sample (right) for 2017 and 2019

### 3.2. Step II – runway configuration and runway throughput

In case of Barcelona airport [5] there are three runway configurations in use during the day: West configuration, parallel runways (25L and 25R, default configuration: 07-23h), East configuration, parallel runways (07L and 07R), and North configuration, intersecting runways (07R and 02, mostly active during the night). In each configurations runways are used in segregated mode, one for arrivals and one for departures. In Figure 3 and Figure 4 comparison between planned runway configuration and actual runway configurations during the day is presented. Alternative configuration, East, is used up to 30% of time (with exception of one hour in 2017, when it was 70% of time in use). Apart from three basic configurations, West Aln configuration also appears as alternative configuration during the night time.



**Figure 3:** Daily distribution of Planned and Actual Runway configuration, selected days in summer 2017



**Figure 4:** Daily distribution of Planned and Actual Runway configuration, selected days in summer 2019

### 3.3. Step III – weather

Due to data observed from METAR reports, prevailing weather conditions are relatively mild, and due to that could not be directly connected to go-around operations. Cross-winds and tail-winds are analysed for each go-around operation separately (in step V). Wind direction and speed on altitudes/heights within final approach, as well as appearance of wind shear could be contributing factors, but those data were not available.

### 3.4. Step IV – key indicators

The number of go-around operations are proposed as a key indicator for tracking potential benefits of the change. The separation between aircraft pairs on final approach and wake turbulence category for each aircraft were extracted from the data as well as separation distribution before and after the change. Possible separation violations were also addressed.

The total number of arrivals for the selected days was 21.214 and 25.761 operations in 2017 and 2019, respectively. Within selected days in a traffic sample the percentage of go-around operations in relation to total number of arrivals was 0,27% (0,31% and 0,23% of number of arrivals in June and July, respectively) for 2017, and 0,35% (0,32% and 0,38% of number of arrivals in June and July, respectively) for 2019.

Share of go-around operations in a total number of arrivals is very low, as expected. The distribution of go-arounds per days of week is shown in Figure 5 (left). In 2017 the highest percentage of go-arounds occurred on Monday, and in 2019 on Thursday and Sunday. In 2017 the highest share of go-around operations occurred in evening (17th and 21st hours - 12,07% of go-arounds), while in 2019 the highest share of go-around operations was in the morning (8th hour - 15,38% of go-arounds) (Figure 5, right).

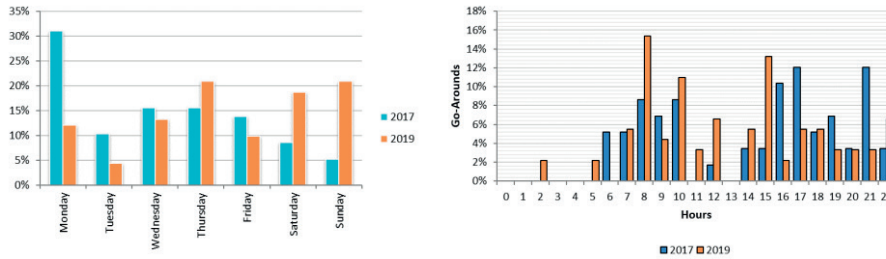


Figure 5: Go-around operations per days (left), and hourly distribution of go-around (right) in 2017 and 2019

In Figure 6, the hourly distribution of go-around is merged with the average hourly distribution of arrivals for 2017 and 2019 respectively. It seems (without entering into correlation analysis) that a higher number of arrivals are followed by higher number of go-around in both years.

According to wake turbulence categorization, in 2017, there were 86.21% go-around operations performed by medium aircraft lead by medium aircraft (M-M aircraft pair, the follower performing go-around). The number of go-around operations between other pairs of aircraft is much smaller. As in 2017, the vast majority of realized go-around operations in 2019 were associated to M-M aircraft type pair (75.82% of go-around operations).

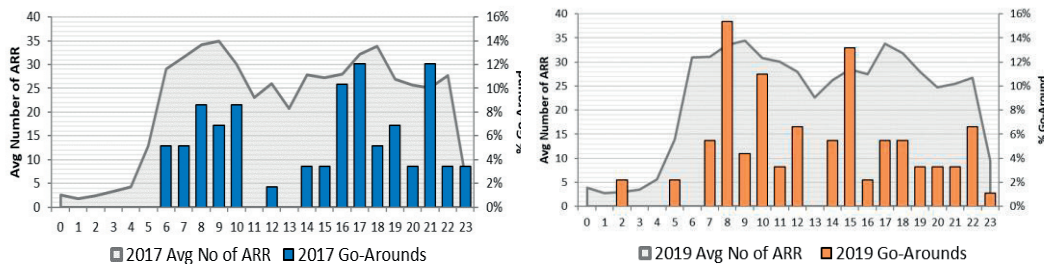


Figure 6: Average number of arrivals and go-around hourly distribution, in 2017 and 2019

Separation distribution for M-M aircraft pairs is presented in Figure 7 for sample days in 2017 and 2019. Comparing the two cases, it can be seen that the separation minima are met, but that the average separation for the case of M-M aircraft category is slightly lower in 2019 (5.83 NM in 2019 vs. 5.99 NM in 2017), which can be interpreted as slight improvement of the system performance due to IAS introduction in the CWPs flight label. For other aircraft pairs, though, the separation was higher in 2019 than in 2017.

The M-M aircraft pairs are observed separately, since that combination of aircraft is the most frequent when it comes to go-around cases. Theoretical distribution best fitting the empirical one is given in Figure 8. Mean separation is 5.82 NM in 2017 and 5.74 NM in 2019. Median also moves towards lower value for 5.0 NM in 2017, to 4.88 NM in 2019.

Ground separation violations were not assessed, since the data required to perfume this analysis were not available in the data-base (runway occupancy time, runway incursion events, etc.).

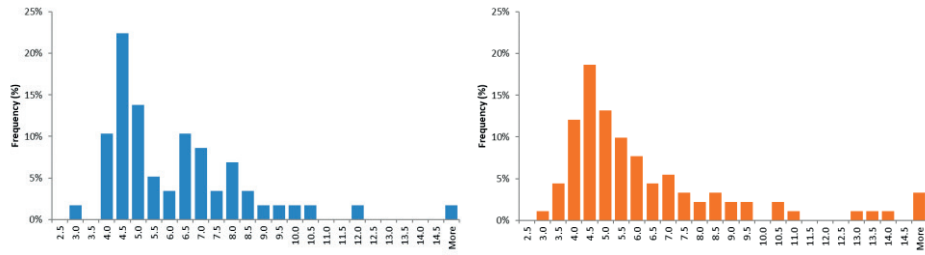


Figure 7: Separation distribution for all aircraft pairs (go-around cases) in 2017 and 2019

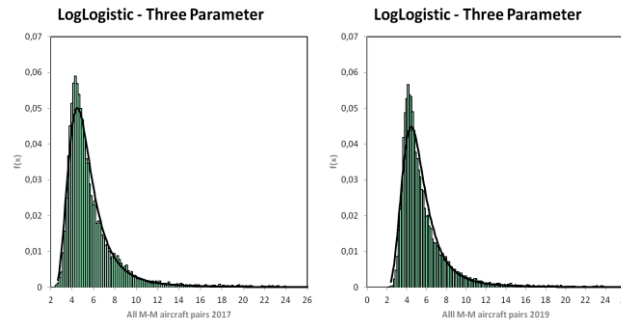


Figure 8: Theoretical distribution of M-M aircraft pair's separation, in 2017 and 2019

### 3.5. Step V – other indicators

In this step, selected factors possibly contributing to go-around were analysed, aiming to explain increased number of go-arounds in 2019 compared to 2017. This analysis includes possible connection of go-around operations to airlines or aircraft types, to runway configuration in use, or other detected anomalies. Go-around operations at Barcelona Airport are mostly performed by low cost airlines that may be a consequence of different descending profiles that include fuel conservation. Low cost airlines mostly fly uniform fleets that consists predominantly of A320 family or B737-700, -800 and MAX aircraft, and these aircraft types take the major shares in go-around operations. A320 aircraft performed 55% of go-arounds in 2017 and 38% in 2019, followed by B737-800 aircraft with 22% (2017) and 19% (2019) share in all go-arounds.

One possible reason for go-around may be postponed ATCO decision for the runway configuration change, allowing operations with stronger tailwind component than commonly accepted 5kt. From Figure 3 and Figure 4, it can be seen that alternative configuration (East) is in use approximately 30% of the time, with one exception in 2017, when it was in use 70% of time between 16h and 17h. This may indicate unfavorable wind conditions during this hour, or one hour before. However, the data does not confirm connection between go-around operations and runway configuration change decisions in these periods. During 16h to 17h hour period there were 10.3% go-arounds (7% when West was active + 3.3% when East was active) in 2017, and only 2.2% in 2019 (all when West was active). All go-arounds in 2017 were performed with headwind or tailwind below 1kt. During previous hour, 15h to 16h, there were only 3.4% go-arounds in 2017 (1.7% when West was active + 1.7% when East was active), but even 13.2% in 2019 (11% when West was active + 2.2% when East was active). This could indicate that in 2019 ATCOs allowed somewhat stronger tailwind to keep West configuration in use, which could have resulted in increase of go-around operations between 15h and 16h. The data do not confirm this, as in 2019 all go around operations in this period were performed with headwind.

Only 8.6% go-around cases possibly occurred due to tailwind (greater than 5kt) in 2017, and only 2.2% cases in 2019 (Figure 9), while there were only 1.7% cases in 2017 and 3.3% cases in 2019 with crosswind above the most restrictive 10kt crosswind limit (for aircraft with reference field length up to 1200m (Figure 10)).

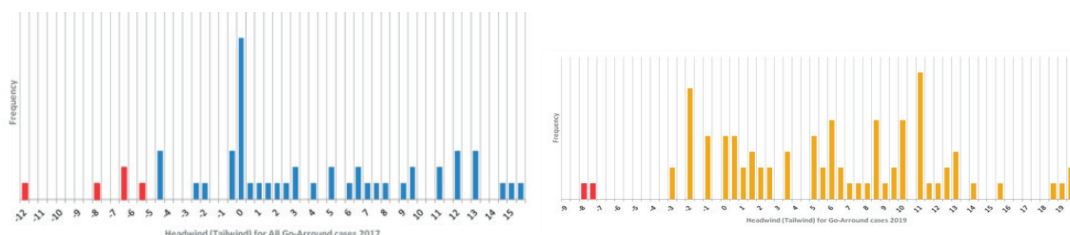
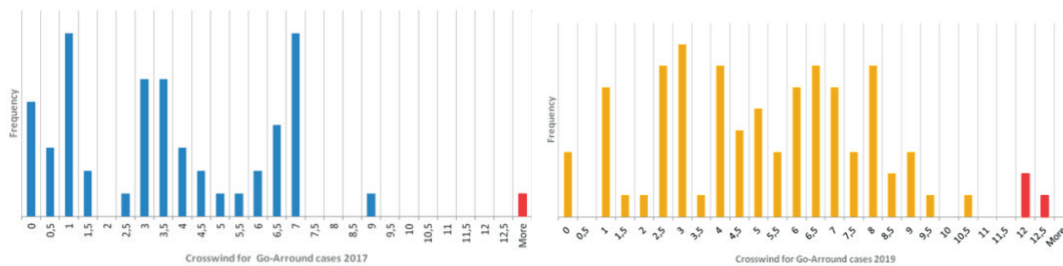


Figure 9: Head/tailwind component for go around cases in 2017 and 2019



**Figure 10:** Crosswind component for go around cases in 2017 and 2019

#### 4. CONCLUSION

The go-arounds are the most safety-critical operations, so it is necessary to perform them in the safest possible manner. Therefore, the speed of aircraft in the final approach is a valuable well-known safety term. IAS incorporation in the flight label for the CWP is expected to reduce the number of go-arounds by non-stabilized approaches or by LoS, among others.

Based on the performed analysis, benefits caused by introduction of IAS speed label on the screen of CWP, were not confirmed using the number of go-arounds. Number of go-arounds increased in 2019 (after change), compared to same period in 2017 (before change). However, the relative number of go-around operations does not increase from 2017 to 2019, meaning that the current safety level is kept despite an increased demand pressure. Other benefits of the change could be associated to indicators that cannot be extracted directly from the traffic data, like decreased ATCO workload, leading to increased throughput in the future.

The main benefits of IAS introduction were expected to be observed in Summer 2020 (after two years since introduction), but due to the COVID-19 impact, only 1.5 years data after the change (including Summer 2019 and half of the Summer 2018) were available. So, the analysis should be repeated when the air traffic stabilizes. Accordingly, the further research should take into account some indicators and data related to them that were not analyzed in the current research due to the lack of available information (i.e. runway occupancy time and runway incursions by another aircraft).

#### Acknowledgement

This paper is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 892542 (FARO - saFety And Resilience guidelines for aviatiOn, <https://faro-h2020.eu/>) under European Union's Horizon 2020 research and innovation programme. The opinions expressed herein reflect the author's view only. Under no circumstances shall the SESAR Joint Undertaking be responsible for any use that may be made of the information contained herein. The work of D. Timotic, B. Mirkovic and F. Netjasov was partially supported by the Project number 36033 commissioned by the Ministry of Education, Science and Technological development of the Republic of Serbia.

#### REFERENCES

- [1] Boeing. (2021). Statistical Summary of Commercial Jet Airplane Accidents: Worldwide Operations 1959-202. Retained from the Internet under: [https://www.boeing.com/resources/boeingdotcom/company/about\\_bca/pdf/statsum.pdf](https://www.boeing.com/resources/boeingdotcom/company/about_bca/pdf/statsum.pdf)
- [2] Dehais, F., Behrend, J., Peysakhovich, V., Causse, M., & Wickens, C. D. (2017). Pilot flying and pilot monitoring's aircraft state awareness during go-around execution in aviation: A behavioral and eye tracking study. *The International Journal of Aerospace Psychology*, 27(1-2), 15-28.
- [3] Dai, L., Liu, Y., & Hansen, M. (2021). Modeling go-around occurrence using principal component logistic regression. *Transportation Research Part C: Emerging Technologies*, 129, 103262.
- [4] Dhief, I., Alam, S., Lilith, N., & Mean, C. C. (2022). A machine learned go-around prediction model using pilot-in-the-loop simulations. *Transportation Research Part C: Emerging Technologies*, 140, 103704.
- [5] Barcelona Airport site. Retained from the Internet under: <https://www.barcelona-airport.com/>

## ODREĐIVANJE RASPOREDA RADA DIZALICA U LUČKIM KONTEJNERSKIM TERMINALIMA

### THE YARD CRANE SCHEDULING PROBLEM AT PORT CONTAINER TERMINALS

MILOŠ NIKOLIĆ<sup>1</sup>, IVANA VUKIĆEVIĆ BIŠEVAC<sup>2</sup>, JOVKO JAĆIMOVIĆ<sup>3</sup>, IVANA JOVANOVIĆ<sup>4</sup>, KATARINA VUKADINOVIĆ<sup>5</sup>, BRANKA DIMITRIJEVIĆ<sup>6</sup>

<sup>1,2,3,4,5,6</sup> Univerzitet u Beogradu – Saobraćajni fakultet, <sup>1</sup>m.nikolic@sf.bg.ac.rs, <sup>2</sup>i.vukicevic@sf.bg.ac.rs, <sup>3</sup>jjacimovic@gmail.com, <sup>4</sup>ivana.jovanovic@sf.bg.ac.rs, <sup>5</sup>k.vukadinovic@sf.bg.ac.rs, <sup>6</sup>b.dimitrijevic@sf.bg.ac.rs

**Rezime:** U ovom radu razmatran je problem određivanja rasporeda rada skladišnih dizalica u lučkim kontejnerskim terminalima. Za ovaj problem u literaturi je predložena matematička formulacija mešovitog celobrojnog linearnog programiranja. Autori ovog rada su uočili određene nedostatke u predloženoj formulaciji. U radu su date potrebne ispravke predložene matematičke formulacije i izvršena su njena testiranja na dva skupa hipotetičkih primera. Dobijeni rezultati su upoređeni sa rezultatima koji su za te primere dobijeni proždrljivim heurističkim algoritmom. Dobijeni rezultati ukazuju da se optimizacionim pristupom mogu postići značajne uštede, međutim vreme rada računara može da predstavlja veliki problem s obzirom na to da se ono značajno povećava sa većim brojem zahteva koje dizalica treba da obavi.

**Ključne reči:** Skladišne dizalice, Lučki kontejnerski terminali, Mešovito celobrojno linearno programiranje

**Abstract:** In this paper, we studied and solved the yard crane scheduling problem at port container terminals. To solve this problem, in the literature, a Mixed Integer Linear Programming formulation is proposed. We have seen a few disadvantages within the proposed formulation. We proposed necessary changes in the mathematical formulation, and performed tests on two sets of hypothetical examples. For these examples, the obtained results are compared to the results obtained by the greedy heuristics. Obtained results indicate that significant savings can be achieved by optimization. However, the CPU time increases significantly when the number of tasks that yard crane needs to perform increases.

**Keywords:** Yard cranes, Port container terminals, Mixed Integer Linear Programming

## 1. UVOD

Zbog globalne pandemije koja je imala negativni uticaj na transport tokom 2020. godine, pomorski transport kontejnera zabeležio je pad od 1,1% [8]. Iako je sa smanjenjem pomorskog kontejnerskog transporta smanjen i promet kontejnera u lukama, u najnovijim izveštajima o pomorskom transportu eksperti Konferencije Ujedinjenih nacija za trgovinu i razvoj (UNCTAD) navode da je povećano prosečno zadržavanje kontejnerskih brodova u luci za 2,3%.

Kao osnovni pokazatelji kvaliteta lučke usluge u literaturi se najčešće navode ukupni troškovi brodarka prouzrokovani pristajanjima i opslugama brodova u lukama, kao i ukupna vremena zadržavanja brodova u lukama. Svaki brod prema planu plovidbe ima unapred određeno očekivano vreme završetka opsluge. Imajući u vidu da je mera (ne)zadovoljstva korisnika razlika između očekivanog i opaženog kvaliteta usluge [7] operatori lučkih kontejnerskih terminala (LKT) teže da izvrše uslugu na vreme, odnosno da opsluže brodove u okviru očekivanog vremena.

Poslednjih godina u LKT dolazi do pomeranja zagušenja/nagomilavanja sa operativne obale u druge delove terminala [4]. U ovom radu pažnja će biti posvećena efikasnijem korišćenju skladišnih dizalica u lučkim kontejnerskim terminalima.

U sledećem poglavlju dat je kratak pregled literature. Zatim je u poglavlju 3 opisan problem koji se razmatra i prikazana je matematička formulacija problema koja je data u literaturi. Predložene izmene matematičke formulacije date su u poglavlju 4. Rezultati numeričkih testiranja prikazani su u poglavlju 5. U poslednjem poglavlju data su zaključna razmatranja i pravci budućih istraživanja.

## 2. PREGLED LUTERATURE

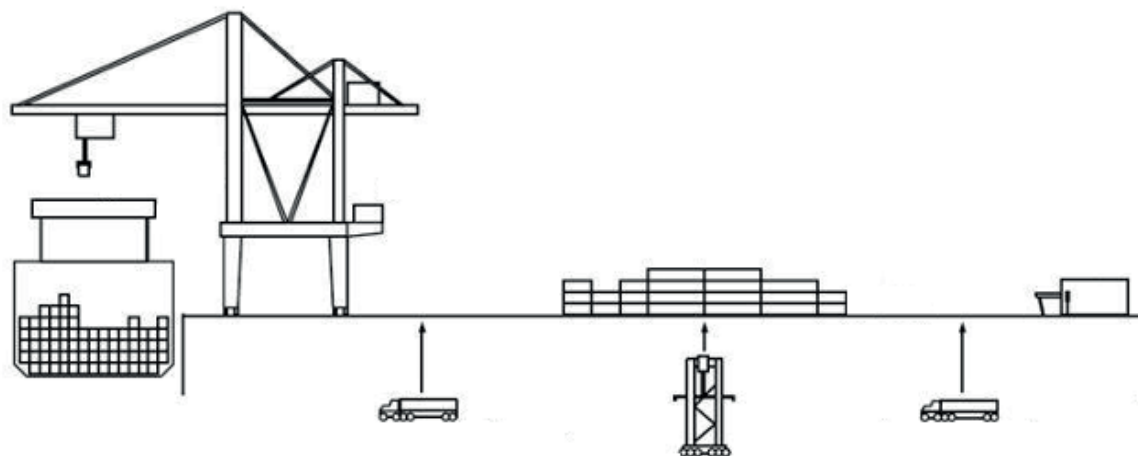
U detaljnom pregledu literature [4] navodi se da većina autora problem formiranja rasporeda rada skladišnih dizalica matematički formuliše kao model mešovitog celobrojnog programiranja, a zatim za rešavanje koriste heurističke i metaheurističke algoritme. Za rešavanje problema formiranja rasporeda rada skladišnih dizalica u radu [1] razvijena je hibridna metodologija zasnovana na heurističkom algoritmu i simulaciji. Problem formiranja rasporeda rada više dizalica formulisani je u radu [6] kao model celobrojnog programiranja za čije su optimalno rešavanje koristili algoritam granjanja i ograničavanja. Većina numeričkih primera, na kojima je testiran algoritam (broj zadataka  $\leq 25$ ), rešena je u okviru prihvatljivog vremena rada računara. Za sličan problem u radu [5] dat je novi model celobrojnog programiranja kojim se minimizira ukupno vreme završetka zadataka. Kako Cplex za 4h nije uspeo da pronađe optimalno rešenje za 15 zadataka, za rešavanje problema realnih dimenzija razvijena je heuristika zasnovana na dinamičkom programiranju.

Najčešći kriterijumi koji se u literaturi koriste za formiranje rasporeda rada skladišnih dizalica su: ukupno vreme potrebno za izvršavanje svih zadataka, ukupno odlaganje završetka zadataka, ukupno vreme čekanja zadataka i sl. Poslednjih godina, uvođenjem koncepta zelenih terminala, kriterijumi vezani za uštedu energije postaju sve značajniji i pri rešavanju zadataka na operativnom planu. Iako su ove dve grupe kriterijuma najčešće u pozitivnoj korelaciji, autori rada [2] pokazali su da postoje situacije u kojima mogu da budu i u negativnoj korelaciji. U cilju pronalaženja kompromisnog rešenja problem formiranja rasporeda rada skladišnih dizalica formulisali su kao problem rutiranja vozila sa mekim vremenskim prozorom (engl. Vehicle Routing Problem with Soft Time Windows, VRPSTW). Formulirani problem su rešavali koristeći dve kriterijumske funkcije. Prva kriterijumska funkcija je predstavljala ukupno vreme odlaganja završetka zadataka, a druga ukupnu potrošnju energije. Kako u LKT postoji neizvesnost u pogledu vremena pristajanja brodova autori rada [2] su u radu [3] formulisali problem formiranja rasporeda rada skladišnih dizalica uzimajući u obzir i neizvesnost povezanu sa vremenom nailaska broda, kao i sa vremenima nailaska i brojem sredstava kopnenih vidova saobraćaja.

## 3. OPIS PROBLEMA

Svi kontejneri, uvozni i izvozni, se određeno vreme zadržavaju u kontejnerskom skladištu dok čekaju da budu utovareni na brod ili otpremljeni sredstvima kopnenih vidova saobraćaja krajnjim korisnicima. Kontejneri se skladište u blokovima grupisanim u zone. Skladišnim dizalicama vrši se uskladištenje kontejnera dok se transport kontejnera unutar terminala najčešće vrši tegljačima sa poluprikolicama (Slika 1).

Primenom najnovijih tehnoloških inovacija pri opsluzi brodova u LKT sve češće dolazi do zagušenja prouzrokovanih tegljačima sa poluprikolicama koji čekaju na opslugu skladišnih dizalica.



Slika 1: Tehnologija opsluge kontejnera u LKT

U radu je razmatran i rešavan problem formiranja rasporeda rada skladišne dizalice tj. problem određivanja redosleda obavljanja grupe zadataka. Svaki zadatak podrazumeva da se izvrši uskladištenje ili iskladištenje jednog kontejnera. U trenutku kada se donosi odluka o rasporedu rada dizalice, svi zadaci koji treba da se realizuju su poznati. Cilj optimizacije u posmatranom problemu je minimizacija ukupnog vremena čekanja zadataka na realizaciju.

U radu [6] data je sledeća matematička formulacija razmatranog problema:



Minimizirati

$$F = \sum_{i=1}^n (t_i - h_i - r_i) \quad (1)$$

pri ograničenjima

$$t_i \geq r_i + h_i, \quad i = 1, 2, \dots, n \quad (2)$$

$$t_j - t_i \geq d_{ij} + h_j - (1 - x_{ij})M, \quad i, j = 1, 2, \dots, n; i \neq j \quad (3)$$

$$x_{ij} + x_{ji} = 1, \quad i, j = 1, 2, \dots, n; i \neq j \quad (4)$$

$$x_{ij} = \{0, 1\}, \quad i, j = 1, 2, \dots, n; i \neq j \quad (5)$$

gde su:

$n$  – broj zadataka koje dizalica treba da realizuje,

$r_i$  – vreme kada će zadatak  $i$  da bude spreman za realizaciju,

$h_i$  – vreme koje je potrebno dizalici da realizuje zadatak  $i$ ,

$d_{ij}$  – vreme koje je potrebno dizalici da dođe od pozicije zadatka  $i$  do pozicije zadatka  $j$ ,

$M$  – veliki pozitivan broj.

Promenljive u posmatranom problemu su:

$t_i$  ( $i = 1, 2, \dots, n$ ) – vreme kada će dizalica da završi realizaciju zadatka  $i$ ,

$x_{ij} = \begin{cases} 1 & \text{ako se zadatak } i \text{ realizuje pre zadatka } j \\ 0 & \text{u suprotnom} \end{cases}$ .

Funkciju cilja (1) koja predstavlja ukupno vreme čekanja zadataka na realizaciju potrebno je minimizirati. Ograničenjem (2) obezbeđuje se da vreme završetka realizacije zadatka ne može da bude pre zbira vremena spremnosti zadatka za realizaciju i vremena potrebnog dizalici da taj zadatak realizuje. Ograničenjem (3) dobija se relacija između vremena realizacije prethodnog i narednog zadatka. Ograničenjem (4) obezbeđuje se preciznost u određivanju redosleda realizacije zadataka. Promenljive  $x_{ij}$  definisane su kao binarne ograničenjem (5).

Pored već prikazane notacije u radu [6] uvedena je i oznaka  $d_{0j}$  koja predstavlja vreme kretanja dizalice od pozicije na kojoj se ona nalazi na početku procesa realizacije zadataka pa do lokacije zadatka  $j$ . Međutim, ukoliko se detaljnije pogleda matematička formulacija (1)-(5) može se uočiti da se u njoj nigde ne uzima u obzir gde se dizalica nalazi na početku procesa. Zbog toga, rešenje koje bi se dobilo ne bi odgovaralo realnosti pošto se ne uzima u obzir vreme koje je dizalici potrebno da dođe do mesta zadatka koji se prvi realizuje. Na taj način, početak realizacije prvog zadatka bi uvek bio u trenutku kada je on spreman za realizaciju, odnosno ako pretpostavimo da se  $k$  – ti zadatak prvi realizuje tada bi važno da je vreme  $t_k$  jednako  $r_i + h_i$ .

U narednom poglavlju date su izmene matematičke formulacije kako bi se prevazišao uočeni problem.

#### 4. MODIFIKOVANA MATEMATIČKA FORMULACIJA

Da bi se prevazišao problem neuzimanja u obzir početne lokacije na kojoj se nalazi dizalica potrebno je izvršiti modifikaciju ograničenja (3) i dodati jedno novo ograničenje. Modifikovana matematička formulacija se može prikazati na sledeći način:

Minimizirati

$$F = \sum_{i=1}^n (t_i - h_i - r_i) \quad (6)$$

pri ograničenjima:

$$t_i \geq r_i + h_i, \quad i = 1, 2, \dots, n \quad (7)$$

$$t_j - t_i \geq d_{ij} + h_j - (1 - x_{ij})M, \quad i = 0, 1, \dots, n; j = 1, 2, \dots, n; i \neq j \quad (8)$$

$$x_{ij} + x_{ji} = 1, \quad i, j = 1, 2, \dots, n; i \neq j \quad (9)$$

$$x_{0j} = 1, \quad j = 1, 2, \dots, n \quad (10)$$

$$t_i \geq 0, \quad i = 0, 1, 2, \dots, n \quad (11)$$

$$x_{ij} = \{0, 1\}, \quad i = 0, 1, \dots, n; j = 1, 2, \dots, n; i \neq j \quad (12)$$

Funkcija cilja (6) je identična funkciji cilja (1). Ograničenje (7) je indentično ograničenju (2). Ograničenje (3) je modifikovano na način da brojač  $i$  uzima vrednosti od 0 do  $n$ . Na taj način se uzima u obzir i vreme potrebno da dizalica dođe od početne pozicije (kada je  $i = 0$ ) do zadatka  $j$ . Ovo ograničenje je u izmenjenoj formulaciji prikazano pod brojem (8). Ograničenje (9) je identično ograničenju (4). Ograničenje (10) je novo ograničenje koje je dodato i ono definiše da promenljive  $x_{0j}$  moraju da uzmu vrednost 1, što znači da zadaci moraju da budu realizovani nakon što se dizalica nađe na početnoj poziciji, što i jeste uvek slučaj. Ograničenje (11) definiše promenljive  $t_i$  kao kontinualne, a ograničenje (12) definiše promenljive  $x_{ij}$  kao binarne.

## 5. REZULTATI TESTIRANJA

U radu su izvršena testiranja na dva skupa od po 5 hipotetičkih primera. U svakom primeru iz prvog skupa postoji 10 zadataka koje dizalica treba da realizuje, a u primerima iz drugog skupa ima po 15 zadataka. Za rešavanje zadataka mešovito celobrojnog linearnog programiranja korišćen je program CPLEX 20.1. Sva testiranja su izvršena na DELL laptop računaru sledećih karakteristika: procesor 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz 2.30 GHz, i 16 GB instalirane RAM memorije.

Rezultati dobijeni rešavanjem problema mešovito celobrojnog programiranja prikazani su u trećoj i četvrtoj koloni Tabele 1. U trećoj koloni su date vrednosti funkcija cilja, odnosno ukupna vremena čekanja zadataka na realizaciju, a u četvrtoj vremena rada računara koja su bila potrebna za njihovo dobijanje. Na osnovu prikazanih rezultata može se uočiti da veći broj zadataka koje dizalica treba da realizuje dovodi do značajnog povećanja vremena kašnjenja. Tako se za 10 zadataka vreme čekanja kretalo od 36,19 do 103,6 minuta, dok je u primerima sa 15 zadataka vreme čekanja bilo u opsegu od 229,43 do 306,78 minuta.

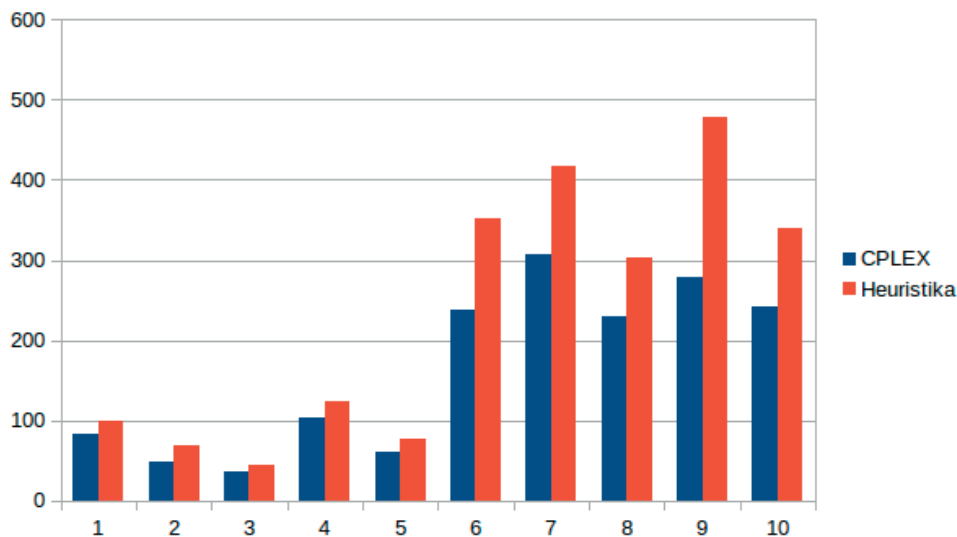
Veoma je važno uočiti da se vreme rada računara drastično povećava sa većim brojem zahteva koje je potrebno realizovati dizalicom. Tako je za rešavanje primera sa 10 zadataka bilo potrebno od 0,09 do 0,19 sekundi, dok je za rešavanje primera sa 15 zadataka bilo potrebno od 73,39 do 1693,58 sekundi. Dalja povećanja broja zadataka koje bi dizalica trebalo da realizuje dovela bi do još većih vremena rada računara. Autori ovog rada su pokušali da reše nekoliko primera sa 20 zadataka i nisu uspeli da dobiju rešenje ni posle nekoliko sati rada računara.

Zbog problema rešavanja primera sa većim brojem zahteva, u ovom radu autori su hteli da provere kvalitet rešenja koja se dobijaju jednim jednostavnim heurističkim pristupom. Heuristički algoritam koji je razvijen i primenjen u ovom radu zasniva se na ideji da dizalica realizuje zadatke redosledom kojim se oni pojavljuju. Ovaj jednostavni algoritam implementiran je u programskom jeziku Java, u Apatch NetBeans 13 razvojnom okruženju. Vremena čekanja zadataka na realizaciju dobijena primenom ovog pristupa za prethodno rešavane primere data su u Tabeli 1, u koloni 5. Potrebna vremena rada računara su uvek iznosila nula sekundi.

Poređenja optimalnih rezultata dobijenih CPLEX programom i rezultata koji su dobijeni heurističkim algoritmom mogu se izvršiti, kako na osnovu informacija o njihovom relativnom odstupanju (Tabela 1, kolona 6), tako i na osnovu grafičkog prikaza (Slika 2) Na osnovu vrednosti u koloni 6 može se uočiti da su relativna odstupanja za primere sa 10 zadataka u intervalu od 18,88 do 44,77 %, dok su za primere sa 15 zadataka ova odstupanja bila u opsegu od 32,31 do 71,73 %.

**Tabela 1:** Rezultati testiranja

Redni broj primera	Broj zadataka	CPLEX		Heuristički algoritam	Relativno odstupanje $\frac{F_h - F}{F} \cdot 100[\%]$
		Funkcija cilja (F) [min]	Vreme rada računara [s]	Funkcija cilja ( $F_h$ ) [min]	
1	10	82,47	0,14	98,8	19,80
2		47,75	0,16	69,13	44,77
3		36,19	0,13	45,4	25,45
4		103,6	0,19	123,16	18,88
5		61,77	0,09	77,29	25,13
6	15	238,8	724,30	352,40	47,57
7		306,78	1693,58	418,01	36,26
8		229,43	119,41	303,57	32,31
9		279,07	1409,30	479,25	71,73
10		243,06	73,39	338,85	39,41



Slika 2: Uporedni prikaz dobijenih vrednosti funkcija cilja  $F$  i  $F_h$

## 6. ZAKLJUČAK

Povećenje obima prevoza kontejnera može značajno da utiče na smanjenje efikasnosti rada pojedinih lučkih podsistema. Zbog toga se istraživači poslednjih 20 godina intenzivno bave razvojem različitih pristupa kojima bi ti problemi prevazišli.

U ovom radu razmatran je problem određivanja rasporeda rada skladišne dizalice u lučkom kontejnerskom terminalu. Za razmatrani problem predložene su izmene matematičke formulacije koja je data u literaturi. Izvršena su testiranja modifikovane matematičke formulacije na hipotetičkim primerima. Na osnovu dobijenih rezultata uočeno je da sa povećanjem broja zadataka koje dizalica treba da realizuje dolazi do značajnog povećanja kako ukupnog vremena čekanja zadataka na realizaciju, tako i vremena rada računara potrebnog za rešavanje.

U radu je testiran i jednostavan heuristički algoritam. Dobijeni rezultati ukazuju da kvalitet rešanja generisanih na taj način može značajno da odstupa od optimalnih rešenja. Zbog toga u okviru pravaca budućih istraživanja treba razmotriti razvoj boljih heurističkih algoritama ili raditi na poboljšanju generisanih rešenja primenom metaheurističkih algoritama.

## LITERATURA

- [1] Chang, D., Jiang, Z., Yan, W., & He, J. (2011). Developing a dynamic rolling-horizon decision strategy for yard crane scheduling. *Advanced Engineering Informatics*, 25(3), 485–494.
- [2] He, J., Huang, Y., & Yan, W. (2015). Yard crane scheduling in a container terminal for the trade-off between efficiency and energy consumption. *Advanced Engineering Informatics*, 29(1), 59–75.
- [3] He, J., Tan, C., & Zhang, Y. (2019). Yard crane scheduling problem in a container terminal considering risk caused by uncertainty. *Advanced Engineering Informatics*, 39, 14–24.
- [4] Kizilay, D., & Eliiyi, D. T. (2020). A comprehensive review of quay crane scheduling, yard operations and integrations thereof in container terminals. *Flexible Services and Manufacturing Journal*, 33(1), 1–42.
- [5] Ng, W. C. (2005). Crane scheduling in container yards with inter-crane interference. *European Journal of Operational Research*, 164(1), 64–78.
- [6] Ng, W. C., & Mak, K. L. (2005). Yard crane scheduling in port container terminals. *Applied Mathematical Modelling*, 29(3), 263–276.
- [7] Radivojević, G., Miljuš, M., & Vidović, M. (2007). *Logistički kontroling i performanse*. Univerzitet u Beogradu - Saobraćajni fakultet, Beograd.
- [8] United Nations. (2021). *Review of Maritime Transport 2021*, United Nations publication issued by the United Nations Conference on Trade and Development.



## PREDIKCIJA POSLEDICA SAOBRAĆAJNIH NEZGODA POMOĆU MODELA BAZIRANOG NA ALGORITMU LOGISTIČKE REGRESIJE

### PREDICTION OF CONSEQUENCES OF TRAFFIC ACCIDENTS MODEL BASED ON LOGISTIC REGRESSION ALGORITHM

PREDRAG GROZDANOVIĆ, KATARINA KUKIĆ, ANA UZELAC, SLAĐANA JANKOVIĆ

Univerzitet u Beogradu – Saobraćajni fakultet, p.grozdanovic@sf.bg.ac.rs, k.mijailovic@sf.bg.ac.rs, ana.uzelac@sf.bg.ac.rs, s.jankovic@sf.bg.ac.rs

**Rezime:** U ovom radu obrađeni su i analizirani podaci o saobraćajnim nezgodama na teritoriji Beograda u periodu od 2018. do 2020. godine. Cilj rada bio je da se kreira prediktivni model baziran na algoritmu logističke regresije. Model je kreiran u programskom jeziku Python i koristi se za predikciju posledica saobraćajnih nezgoda (SN). Na osnovu dobijenih rezultata analizirano je koji od razmatranih faktora najviše utiču na to kakve će biti posledice SN.

**Ključne reči:** Saobraćajne nezgode, Logistička regresija, Predikcija, Python

**Abstract:** This paper deals with and analyzes data on traffic accidents on the territory of Belgrade in the period from 2018 to 2020. The aim of this paper was to create a predictive model based on the logistic regression algorithm. The model was created in the Python programming language and is used to predict the consequences of a traffic accident (SN). Based on the obtained results, it was analyzed which of the considered factors most influence the consequences of SN.

**Keywords:** Traffic Accidents, Logistic Regression, Prediction, Python

#### 1. UVOD

Mašinsko učenje je jedna od trenutno najpopularnijih oblasti računarskih nauka, a istovremeno je i tehnika koja se može upotrebiti u najrazličitijim sferama privrede i nauke. Može se posmatrati kao generalizacija znanja na osnovu prethodnih iskustava i sve više se koristi kao jedna od tehnika prediktivne analitike. Pomoću mašinskog učenja, funkcioniše filteriranje neželjene pošte, softveri za prepoznavanje teksta i glasa, veb pretraživači, itd.

U ovom radu obrađen je algoritam logističke regresije koji spada u grupu klasifikacionih algoritama. Model baziran na algoritmu logističke regresije je testiran na podacima o saobraćajnim nezgodama (SN) na teritoriji Beograda u periodu od 2018. do 2020. godine. Podaci su preuzeti iz Integrisane baze podataka o obeležjima bezbednosti saobraćaja [1] Agencije za bezbednost saobraćaja, korišćenjem dostupne Web GIS aplikacije [2]. Korišćeni dataset u ovom radu poseduje 127 atributa saobraćajne nezgode i 29887 instanci. Modeli mašinskog učenja su razvijani i primenjivani za predikciju korišćenjem programskog jezika Python.

Zbog velikog porasta broja vozila drumskog saobraćaja u svetu, dolazi do velikih zagušenja saobraćaja, zagađenja životne sredine, ali i povećanja broja saobraćajnih nezgoda. U literaturi postoje brojni radovi koji analiziraju saobraćajne nezgode i njihove posledice. Predviđanje težine saobraćajne nezgode je od suštinskog značaja za upravljanje bezbednošću saobraćaja, pa iz tog razloga ne čudi činjenica da se najveći broj radova bavi upravo predviđanjem težine posledica saobraćajne nezgode.

U radu [8], predložen je hibridni model koji integriše algoritam Random forest (RF) i Bayesian optimization (BO), u cilju dobijanja što veće tačnosti prilikom predikcije ozbiljnosti posledica SN. U predloženom modelu, BO-RF, RF je usvojen kao osnovni prediktivni model, dok se BO koristi za podešavanje parametara RF modela. Predviđanjem iste ciljne varijable bavili su se autori rada [9] 2020. godine. Oni su za predikciju koristili hibridni model koji se zasniva na kombinaciji K-means i Random forest algoritma. K-means izdvaja skrivene informacije iz podataka o saobraćajnim nezgodama i stvara novu funkciju u setu za obuku, dok se RF algoritam koristi za predikciju ciljne varijable. U radu [4] iz 2002. godine, za predikciju posledica SN korišćen je model baziran na logističkoj regresiji. Cilj ovog rada bio je da se kreira prediktivni model koji će predvideti posledicu SN, pri čemu su sve posledice SN svrstane u dva moguća ishoda: SN sa fatalnim i SN sa nefatalnim ishodom. Radom [5] iz 2015. godine, pokazano je da se model baziran na algoritmu logističke regresije može koristiti i za predikciju žarišta saobraćajne nezgode. Na

osnovu definisanih ulaznih promenljivih, model iz pomenutog rada je sposoban da predvidi da li je posmatrano mesto žarište ili ne i to sa tačnošću od čak 86.67%.

## 2. PREDIKCIJA POSLEDICA SAOBRAĆAJNIH NEZGODA PRIMENOM PYTHON-A

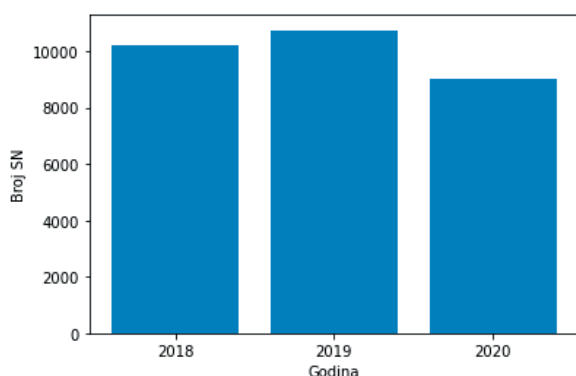
Mašinsko učenje deli se u tri osnovne grupacije: nadgledano, nenadgledano i učenje uslovljavanjem [6]. Python je jedan od najpopularnijih programskih jezika za istraživanje podataka. Zahvaljujući veoma aktivnim programerima i zajednici otvorenog kôda, razvijen je veliki broj korisnih biblioteka za naučna izračunavanja i mašinsko učenje. U ovom radu upravo je korišćen Python kao jezik za obradu i predikciju nad podacima o saobraćajnim nezgodama, primenom nadgledanog tipa mašinskog učenja. Testiranje je izvršeno pomoću pet prediktivnih modela baziranih na sledećim algoritmima: Random forest, K-nearest neighbors, Decision tree, Multinomial Bayes i Logistic regression. Od ovih algoritama koji su korišćeni za predikciju posledica SN, u ovom radu detaljno je analiziran algoritam logističke regresije.

Logistička regresija je proces modeliranja diskretnog ishoda u zavisnosti od ulaznih promenljivih. Najčešće logistička regresija modeluje binarni ishod, odnosno najčešće zavisna promenljiva može imati dve vrednosti kao što su tačno/netačno, da/ne, itd. Da bi se dobili potrebni podaci za analizu pomoću *logit()* algoritma, u Python-u je potrebno uvesti modul *statsmodels.api*.

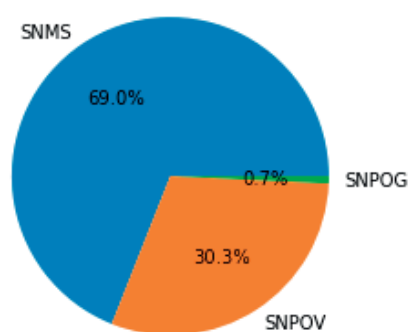
### 2.1. Izbor ulaznih promenljivih za predikciju

Cilj predikcije posledice SN bio je da se predvidi da li nezgoda za posledicu ima materijalnu štetu ili nastradala lica. Za posledicu sa nastradalim licima smatra se ona saobraćajna nezgoda u kojoj bar jedno lice ima povredu ili je poginulo. Pre formiranja prediktivnog modela za opisani problem, treba napomenuti da ovaj primer predstavlja problem klasifikacije, jer ciljna varijabla - saobraćajna nezgoda sa materijalnom štetom (SNMS) - ima dve moguće vrednosti 1 ili 0. Vrednost 1 označava da je posledica SN materijalna šteta, dok vrednost 0 ukazuje da je posledica SN nastradalo/a lice/a.

U okviru ovog koraka prvo je potrebno izvršiti uvoz podataka. Uvoz podataka se vrši pomoću neke od metoda iz paketa *pandas*, a izbor metode zavisi od formata u kom su sačuvani podaci. Za ciljnu varijablu u ovom primeru izabrana je kolona objekta tipa DataFrame koja ima oznaku *SNMS*, a koja je već opisana. Iz dostupnih podataka, analizirane su brojne karakteristike SN na teritoriji Beograda: broj SN po godinama za period 2018-2020, slika 1, broj SN grupisan po godinama i posledicama, po opštinama, po satima u toku dana, da li su se dogodile u naselju ili van naselja itd. Od karakteristika važnih za ovaj rad, izdvojena je raspodela SN po posledicama: sa materijalnom štetom, pri čemu se SN sa povređenim i poginulim licima (SNPOV i SNPOG) u modelu posmatraju kao jedna kategorija - SN sa nastradalim licima, slika 2.



Slika 1: Broj saobraćajnih nezgoda po godinama



Slika 2: Raspodela saobraćajnih nezgoda prema posledicama

Nakon toga izvršena je selekcija atributa i od 126 atributa, u uži izbor ušlo je 13. Preostalih 113 atributa izbačeno je iz daljeg razmatranja zbog nepostojanja ili izrazito male korelacije sa ciljnom varijablom. Izabranih 13 atributa smešteno je u listu *Ulaz*. Elementi ove liste su nazivi kolona koje čine ulazne promenljive za predikciju. Ulazne promenljive su: x1-'BrojVozaca', x2-'Cas', x3-'BrojTeretnihVozila', x4-'Vozac14', x5-'Vozac25', x6-'Vozac40', x7-'Vozac65', x8-'Vozac100', x9-'BrojAutobusa', x10-'BrojPesaka', x11-'BrojBicikala', x12-'BrojPrikljucnihVozila', x13-'BrojTraktora'. Atributi korišćeni za predikciju ciljne varijable su (za detaljniji opis atributa pogledati [3] i [7]):

- Vozac14, 25, 40, 65 i 100 – atribut binarnog tipa koji ukazuje na starosnu grupu kojoj pripadaju vozači koji su učestvovali u SN. Kategorija Vozac14 se odnosi na vozače do 14 godine (vozač bicikla), Vozac25 se odnosi na vozače od 15 do 25, itd.

- Cas – označava čas u kom se SN dogodila.
- BrojVozaca i BrojPesaka – označavaju broj vozača, odnosno pešaka koji su učestvovali u SN.
- BrojTeretnihVozila, BrojAutobusa, BrojBicikala, BrojPrikljucnihVozila i BrojTraktora – ukazuju na broj vozila određene kategorije koja su učestvovala u SN.

Za izbor ulaznih promenljivih korišćena je heat mapa koja prikazuje korelaciju između svake promenljive na grafiku. Za crtanje heat mape pored ulaznih promenljivih kao promenljiva prosleđena je i izlazna promenljiva  $y$  – SNMS kako bi se videla njena korelacija sa ulaznim promenljivim. Heat mapa se crta pomoću metode `heatmap()` paketa `seaborn` tako što se pomenutoj metodi kao parametar prosledi novokreirani objekat tipa `DataFrame` (`df`) koji sadrži ulazne i izlazne promenljive, nad kojim je pozvana metoda `corr()`. Metoda `corr()` računa korelaciju između svake kolone `df`-a nad kojima je pozvana.

Na grafiku prikazanom na slici 3, predstavljen je vizuelni prikaz korelacije između svih parova kolona prosleđenog `df`-a. Može se primetiti da najveću korelaciju sa izlaznom promenljivom  $y$  – SNMS ima ulazna promenljiva  $x_{10}$  – BrojPesaka i da je ta korelacija negativna. Negativna korelacija znači da se sa povećanjem broja pešaka u SN smanjuje šansa da SN ima materijalnu štetu.

Na osnovu prikazanog grafika može se videti da pojedine ulazne promenljive treba izbaciti iz daljeg razmatranja zbog postojanja korelacije sa drugim ulaznim promenljivama. Promenljive  $x_4$  – Vozac14,  $x_6$  – Vozac40,  $x_7$  – Vozac65 i  $x_{12}$  – BrojPrikljucnihVozila su izuzete iz daljeg razmatranja zbog postojanja korelacije sa drugim ulaznim promenljivama.



Slika 3: Heat mapa za SNMS

## 2.2. Kreiranje prediktivnog modela baziranog na algoritmu logističke regresije

Na osnovu izabranog ulaza i izlaza definisane su promenljive  $X$  i  $y$  sledećim kôdom:

```
X=df[Ulaz], y=df.SNMS
```

Promenljiva  $X$  sadrži samo one kolone `df`-a koje su definisane kao ulazne (nezavisne) promenljive. Promenljiva  $y$  sadrži jednu kolonu `df`-a SNMS koja čini izlaznu (zavisnu ili ciljnu) promenljivu. Nakon definisanja skupa ulaznih promenljivih i ciljne promenljive potrebno je da se skup podataka podeli na skup za trening i skup za testiranje. To je urađeno pomoću metode `train_test_split()`. Ovoj metodi je potrebno proslediti promenljive  $X$  i  $y$ , kao i definisati veličinu test skupa (u ovom slučaju to je 20%, što je 5976 instanci), pri čemu je veličina trening skupa automatski postavljena na 80%, odnosno trening skup obuhvata 23911 instanci. Promenljiva  $X_{\text{trening}}$  čuva ulazne trening podatke, a  $X_{\text{test}}$  ulazne test podatke, dok promenljiva  $y_{\text{trening}}$  čuva vrednosti ciljne varijable za trening podatke,  $y_{\text{test}}$  za test podatke. Na osnovu poređenja vrednosti  $y_{\text{test}}$  promenljive i promenljive koja će sadržati podatke predikcije ulaznih test podataka procenjuje se tačnost postavljenog modela.

Algoritam koji je korišćen za predikciju u ovom radu je *LogisticRegression*. Pre same predikcije, kod ovog modela, potrebno je izvršiti skaliranje podataka. Skaliranje podataka podrazumeva transformisanje vrednosti svih ulaza na interval od -1 do 1. Skaliranje je izvršeno pomoću metode `StandardScaler()`.

Potrebno je napomenuti da se u Python-u skaliranje radi tako što se prvo inicijalizuje metoda `StandardScaler()`, a onda se na nju nadoveže metoda `fit_transform()` za ulazni trening set, odnosno `transform()` za ulazni test set. Metoda `fit_transform()` istovremeno transformiše i uklapa podatke, dok metoda `transform()` samo vrši transformaciju prosleđenih podataka. Uvek se uklapanje radi samo za ulazne trening podatke, a ne za test podatke.

Nakon skaliranja podataka, potrebno je izvršiti inicijalizaciju modela pozivanjem metode `Logit()` iz `statsmodels.api` modula. Zatim je potrebno izvršiti uklapanje podataka u model pomoću metode `fit()`, kojoj se prosleđuju ulazni i izlazni trening setovi. Na ovaj način je kreiran prediktivni model i može se koristiti za predikciju na novim podacima.

### 3. REZULTATI

Kako bi se mogli analizirati rezultati predikcije potrebno je izračunati različite mere tačnosti koje ukazuju na kvalitet prediktivnog modela. U ovom radu korišćena je mera tačnosti `Accuracy_score` koja se dobija se na sledeći način [6]:

$$accuracy = \frac{\text{Broj tačnih predikcija}}{\text{Ukupan broj predikcija}} \quad (1)$$

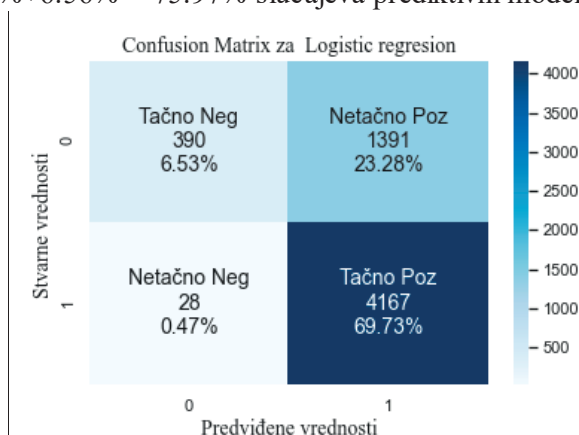
U slučaju binarne predikcije, najčešće se koristi `confusion matrix` za matricni prikaz broja tačnih i netačnih predikcija i jedne i druge klase. Na osnovu `confusion matrix` računa se i prethodno opisana metrika `accuracy`.

#### 3.1. Rezultati prvog testiranja algoritma logističke regresije – LOGR

Nakon kreiranja prediktivnog modela baziranog na logističkoj regresiji moguće je izvršiti predikciju na test skupu podataka. Predikcija se vrši tako što se pozove metoda `predict()` kojoj se prosleđuje ulazni test skup. Kao rezultat dobijaju se predviđene vrednosti za taj set koje se čuvaju u promenljivoj `yhat`, a koja ustvari predstavlja vrednosti logističke funkcije. S obzirom da se radi problem klasifikacije te vrednosti je potrebno konvertovati u jedinice i nule. Konvertovanje je izvršeno pomoću metode `map()`, pri čemu je vrednost koja označava granicu konvertovanja postavljena na 0.5. Na taj način one vrednosti koje su bliže jedinici, veće od 0.5, konvertuju se u 1, a one vrednosti koje su bliže nuli konvertuje se u 0, a dobijeni rezultat se čuva u listi `ypred`. Kreirani prediktivni model prvo je testiran za 9 ulaznih promenljivih za koje se smatra da su najznačajnije na osnovu korelacije prikazane heat mapom, na početku poglavlja.

```
log_reg.fit(Xttraining, yttraining)
yhat=log_reg.predict(Xtest)
ypred = list(map(round, yhat))
```

Prikazani grafik predstavlja `confusion matrix` u grafičkoj formi. Sa grafika prikazanog na slici 4, može se videti da je u ukupno  $69.41\% + 6.56\% = 75.97\%$  slučajeva prediktivni model bio tačan.



Slika 4: Confusion matrix za LOGR algoritam stilova

Pored `confusion matrix`, određena je `accuracy_score` mera tačnosti koja ukazuje da je tačnost ovog algoritma 75.97%. Od ukupno 5976 izvršenih predikcija model je bio tačan u 4540 slučajeva. Međutim, provera smislenosti modela je njegovo poređenje sa `Null modelom`. `Null model` u klasifikacionim problemima za ishod predikcije uvek uzima najčešću vrednost, a u ovom slučaju to je da posledica SN bude materijalna šteta. S obzirom da u posmatranom datasetu 69% SN za posledicu ima materijalnu štetu, onda se može reći da bi tačnost `Null modela` bila upravo 69%. Poređenjem tačnosti `accuracy_score` kreiranog modela sa tačnošću `Null modela` može se zaključiti da je kreirani model bolji, što je i prvi korak u predikciji.



### 3.2. Poboljšanje performansi LOGR algoritma i drugo testiranje

U okviru ovog poglavlja će se na osnovu analize vrednosti koeficijenata,  $Z$  vrednosti,  $p$  vrednosti i intervala poverenja, odrediti koje su promenljive značajne i poboljšati performanse kreiranog modela. Takođe, analizom pomenutih parametara izvešće se zaključci kako rast pojedinačnih varijabli utiče na šansu za pojavu SN sa materijalnom štetom. Python-ov modul *statsmodels.api* poseduje metodu *summary()* koja se poziva nad modelom logističke regresije *Logit()*, a koja daje prikaz vrednosti pomenutih parametara i to sve za svaku ulaznu promenljivu prediktivnog modela.

Da bi se lakše razumela interpretacija *LOGR modela*, treba prvo reći da *LOGR model* podrazumeva izračunavanje logaritma šanse za realizacijom nekog ishoda i to po sledećoj formuli (2):

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \quad (2)$$

gde su:  $\frac{p}{1-p}$  – šansa ili kvota da SN za posledicu ima materijalnu štetu,  $\beta_0, \beta_1, \dots$  - koeficijenti logističke regresije,  $x_1, x_2$  – ulazne promenljive. Međutim da bi se omogućila lakša interpretacija *LOGR modela*, potrebno je iz formule (2) izraziti kvotu, odnosno:

$$\frac{p}{1-p} = e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots)} \quad (3)$$

Da bi se koeficijent ulazne promenljive kod logističke regresije tumačio na isti način kao i kod linearne, potrebno je izračunati vrednost  $e^{coef}$ , gde je *coef* koeficijent logističke regresije ( $\beta_0, \beta_1, \dots$ ), dobijen metodom *summary()*. Na taj način izračunata vrednost ukazuje na to za koliko će se povećati ili smanjiti šansa za SN sa materijalnom štetom ukoliko se posmatrana ulazna promenljiva poveća ili smanji za jednu jedinicu. Ukoliko se šansa za nastanak SN sa materijalnom štetom (MS) označi kao  $OR = e^{coef}$  (*odds ratio*), onda se dobijena vrednost za  $OR$ , za posmatranu ulaznu promenljivu, može tumačiti na sledeći način:  $OR > 1$  utiče na ishod i to povećanjem vrednosti promenljive povećava se šansa za SN sa MS, dok vrednost  $OR < 1$  utiče na ishod i to povećanjem vrednosti promenljive smanjuje se šansa za SN sa MS.

Kolona izveštaja za  $Z$  vrednost ukazuje na značaj ulazne promenljive. Što je  $Z$  vrednost dalje od nule, to je promenljiva značajnija za predikciju.  $Z$  vrednost se računa kao količnik očekivane vrednosti (*coef*) i njene standardne greške (*S.E.*). Interval poverenja se odnosi na koeficijent ulazne promenljive i on u zavisnosti od njegove širine ukazuje u kojim granicama može varirati vrednost koeficijenta. Npr. interval poverenja od 95% ukazuje u kojim granicama varira vrednost koeficijenta za 95% slučajeva korišćenih podatka.

Koje promenljive će biti izbačene iz daljeg razmatranja određuje se na osnovu  $p$  vrednosti promenljivih dobijene metodom *summary()*. Međutim, najpre je potrebno zadati vrednost za  $\alpha$ , koja predstavlja prag prihvatanja neke promenljive, a u ovom slučaju iznosi 0.05. U konkretnom primeru, za promenljivu  $x_2$  treba prihvatiti hipotezu  $H_0$ , da ta promenljiva nije značajna i izbaciti je iz daljeg razmatranja, jer je njena  $p$  vrednost veća od zadate vrednosti za  $\alpha$ ,  $p_2=0.067$ . Za sve ostale ulazne promenljive treba odbaciti nultu hipotezu i usvojiti hipotezu  $H_1$ , a to je da su one značajne za predikciju ciljne promenljive.

Nakon odbacivanja promenljive  $x_2$  – *Cas*, model je ponovo testiran. Na osnovu dobijenih rezultata za *accuracy\_score*=0.7625, može se zaključiti da je povećana tačnost modela sa 75.97% na 76.25%, nakon uklanjanja pomenute ulazne promenljive. Izbacivanjem promenljive  $x_2$  ostvareno je smanjenje složenosti modela, a istovremeno i povećanje njegove tačnosti.

Izveštaj metode *summary()* sa izračunatim eksponencijalnim vrednostima koeficijenata i granica intervala poverenja *LOGR modela* prikazan je u tabeli 1. Na osnovu rezultata u koloni  $Z$ , može se reći da je najznačajnija ulazna promenljiva za predikciju  $x_{10}$  – *BrojPesaka*, jer je najudaljenija od nule, a najmanje značajna je  $x_5$  – *Vozac25*, jer je najbliža nuli. U koloni  $OR$ , prikazane su eksponencijalne vrednosti koeficijenata, a u kolonama *LowerCI* i *UpperCI*, eksponencijalne vrednosti granica intervala poverenja.

Iz table 1 se može videti da je vrednost za  $OR$  za promenljivu  $x_1$  jednaka 0.77 i s obzirom da je manja od 1, onda se može reći da se sa povećanjem vrednosti promenljive  $x_1$  za jednu jedinicu smanjuje šansa da posledica SN bude materijalna šteta, 0.77 puta, pri čemu vrednost  $OR$  može da varira u intervalu od 0.745 do 0.796, sa pouzdanošću 95%. Na isti način može se interpretirati zavisnost ciljne varijable sa ostalim promenljivama, recimo povećanje promenljive  $x_{10}$  – 'broj pešaka koji su učestvovali u SN', smanjuje šansu da posledica SN bude materijalna šteta, dok sa povećanjem promenljive  $x_3$  - 'BrojTeretnihVozila' raste šansa da posledica SN bude materijalna šteta. Na ovaj način je pokazano da se analizom rezultata iz table 1 može videti kako će se šansa za SN sa materijalnom štetom menjati u zavisnosti od promene vrednosti neke od promenljivih.

**Tabela 1:** Izveštaj metode summary() za model baziran na algoritmu logističke regresije

Promenljive	coef	OR= $e^{\text{coef}}$	S.E.	Z	P> Z	Lower CI	Upper CI
x1	-0.262	0.770	0.017	-15.295	0.000	0.745	0.796
x3	0.216	1.241	0.016	13.129	0.000	1.202	1.283
x5	0.047	1.048	0.016	3.034	0.002	1.017	1.081
x8	0.087	1.090	0.016	5.438	0.000	1.057	1.125
x9	0.131	1.140	0.016	8.140	0.000	1.104	1.177
x10	-3.709	0.025	0.061	-61.259	0.000	0.022	0.028
x11	-0.315	0.730	0.021	-14.958	0.000	0.700	0.761
x13	-0.069	0.934	0.013	-5.141	0.000	0.909	0.959

#### 4. ZAKLJUČAK

U ovom radu izvršena je predikcija posledica SN na osnovu podataka o saobraćajnim nezgodama na teritoriji Beograda u periodu od 2018. do 2020. godine. Podaci su obrađeni i analizirani pomoću programskog jezika Python. U radu su prikazane i opisane izabrane metode za rad sa podacima. Krajnji cilj rada bio je da se kreira model koji će moći da predviđa posledicu SN na osnovu određenih ulaznih podataka.

Kreirani model predviđa posledicu SN sa tačnošću od 76.25% čime je ovaj model bolji od *Null modela*. Takođe, zaključeno je da najveći uticaj na ciljnu varijablu ima promenljiva *x10*, a najmanji uticaj ima promenljiva *x5*. Sa povećanjem promenljive *x10* za jednu jedinicu, odnosno broja pešaka koji su učestvovali u SN, smanjuje se šansa da posledica SN bude materijalna šteta 0.025 puta. Druga najznačajnija promenljiva za predikciju je *x1 – BrojVozaca*. Na osnovu toga može se zaključiti da se sa povećanjem broja vozača, odnosno vozila koja su učestvovala u SN smanjuje šansa da posledica SN bude materijalna šteta 0.77 puta.

Povećanje tačnosti ovog modela potencijalno se može ostvariti većim skupom ulaznih podataka iz kojih bi model mogao da nauči. U budućnosti rezultati modela baziranog na algoritmu logističke regresije mogu se porediti sa rezultatima drugih prediktivnih algoritama koji se mogu primeniti za problem klasifikacije.

#### ZAHVALNICA

Ovaj rad delimično je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije, u okviru projekata pod brojem 036012 i TR36002.

#### LITERATURA

- [1] Agencija za bezbednost saobraćaja. „Baza podataka o obeležjima bezbednosti saobraćaja u Republici Srbiji“, <https://www.abs.gov.rs/sr/analize-i-istrazivanja/baza-podataka>. Pristupano [15.12.2021.]
- [2] Agencija za bezbednost saobraćaja. „Web Gis Aplikacija“. <http://195.222.99.60/ibbsPublic/>. Pristupano [15.12.2021.]
- [3] Agencija za bezbednost saobraćaja. „Uputstvo za razumevanje podataka ABS.“ [https://www.abs.gov.rs/admin/upload/documents/20171219103647-web\\_gis\\_razumevanje20podataka.pdf](https://www.abs.gov.rs/admin/upload/documents/20171219103647-web_gis_razumevanje20podataka.pdf). Pristupano [15.12.2021.]
- [4] Al-Ghamdi, A. (2002). Using logistic regression to estimate the influence of accident factors on accident severity, *Accident Analysis and Prevention*, 34(6), 729-741.
- [5] Lu, T., Donyao, Z., Lixin, Y., Pan, Z. (2015). The traffic accident hotspot prediction: Based on the logistic regression method. *2015 International Conference on Transportation Information and Safety (ICTIS)*, Wuhan, China, 107 – 110.
- [6] Ozdemir, S (2016). Principles of Data Science, Packt Publishing, Birmingham.
- [7] Pešić, D., Antić, B., Lipovac, K. (2019). Bezbednost saobraćaja – metode i analize, Univerzitet u Beogradu Saobraćajni fakultet, Beograd.
- [8] Yan, M. Shen, Y. (2022). Traffic Accident Severity Prediction Based on Random Forest. *Sustainability*, 14(3), 1729.
- [9] Yassin, S., S., Pooja (2020). Road accident prediction and model interpretation using a hybrid K-means and random forest algorithm approach. *SN Applied Sciences*, 2, 1576.

## EASY CHAIRPREDVIĐANJE KAŠNJENJA LETOVA U DOLASKU KORIŠĆENJEM MAŠINSKOG UČENJA

### PREDICTION OF FLIGHT DELAYS ON ARRIVAL USING MACHINE LEARNING

NENAD GOLUBOVIĆ<sup>1</sup>, SLAĐANA JANKOVIĆ<sup>2</sup>, ANA UZELAC<sup>3</sup>, SNEŽANA MLADENOVIĆ<sup>4</sup>

<sup>1</sup> Univerzitet u Beogradu – Saobraćajni fakultet, nenadgolubovic999@gmail.com

<sup>2</sup> Univerzitet u Beogradu – Saobraćajni fakultet, s.jankovic@sf.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Saobraćajni fakultet, ana.uzelac@sf.bg.ac.rs

<sup>4</sup> Univerzitet u Beogradu – Saobraćajni fakultet, snezanam@sf.bg.ac.rs

**Rezime:** Mašinsko učenje se može predstaviti kao generalizacija znanja na osnovu prethodnog iskustva. Danas, u Big Data eri, mašinsko učenje koristi se kao jedna od vodećih tehnika prediktivne analitike. U ovom radu predstavljene su osnovne faze u projektovanju modela mašinskog učenja namenjenog predviđanju kašnjenja letova. U radu je opisana metoda predviđanja kašnjenja letova bazirana na primenama algoritama mašinskog učenja, kao i implementacija svih faza procesa mašinskog učenja u programskom jeziku Python 3.0, uz pomoć biblioteka Pandas, Numpy, Matplotlib, Seaborn, Scikit-Learn. Python je predstavnik jezika za laku obradu podataka. Prateće biblioteke su razvijane za analitiku, vizualizaciju i manipulaciju podataka. Scikit-Learn biblioteka predstavlja biblioteku u programskom jeziku Python koja sadrži kolekciju algoritama mašinskog učenja. Implementacija procesa mašinskog učenja demonstrirana je na studiji slučaja predikcije kašnjenja letova na aerodromu „Denver International Airport“.

**Cljučne reči:** Veštačka inteligencija, Vazdušni saobraćaj, Mašinsko učenje, Python

**Abstract:** Machine learning can be presented as a generalization of knowledge based on previous experience. Today, in the Big Data era, machine learning is used as one of the leading techniques of predictive analytics. In this paper are presented the basic stages in designing of machine learning models intended for predicting flight delays. The paper describes a method of predicting delayed flights based on the application of machine learning algorithms as well as the implementation of all phases of machine learning process in the Python 3.0 programming language, with the help of Pandas, Numpy, Matplotlib, Seaborn, Scikit-Learn libraries. Python is a representative language for easy data processing. Accompanying libraries have been developed for data analysis, visualization, and manipulation. The Scikit-Learn library is a library in the Python programming language that contains a collection of machine learning algorithms. The implementation of the machine learning process was demonstrated in a case study of the prediction of delayed flights at Denver International Airport.

**Keywords:** Artificial Intelligence, Air Traffic, Machine Learning, Python

## 1. UVOD

Vazdušni saobraćaj predstavlja savremeni vid saobraćaja koji omogućava veliku povezanost i laku dostupnost udaljenih geografskih destinacija. Pored toga karakteriše ga velika brzina prevoza, visok kvalitet usluge, visoka cena opreme i visoka cena vazduhoplov. U prošlosti avio saobraćaj je predstavljao luksuzan način putovanja koji je za posledicu imao visoke cene putnih karata i malu potražnju od strane putnika. Stalnim usavršavanjem celokupnog sistema avio prevoza poboljšavaju se bezbednosni aspekti, kvalitet usluge, cene prevoza se smanjuju i posledično tome potražnja za avio saobraćajem ima konstantan rast. Usavršavanje vazduhoplov, proizvodnja bespilotnih vazduhoplov, usavršavanje sistema na vazduhoplovu i usavršavanje operativnih procesa sistema vazdušnog saobraćaja za posledicu ima rast količina generisanih podataka. Jedan komercijalni vazduhoplov može da proizvede 1 TB podataka tokom samo jednog leta. Tokom prosečnog letnjeg dana, u svetu se obavi 200.000 letova. Dakle, za samo jedan dan vazdušni saobraćaj može da proizvede 200.000 TB podataka samo od vazduhoplova. Količina podataka bi se znatno povećala kada bi svemu tome dodali količinu podataka koju proizvode ostali činiooci vazdušnog saobraćaja (kontrola letenja, aerodromi, kompanije i vazduhoplovne i međunarodne vlasti, putnici, kompanije za prihvatanje i otpremu vazduhoplova, itd.). Stoga podaci u vazdušnom saobraćaju predstavljaju dobru podlogu za Big Data analitiku. Važnu klasu tehnika

*Big Data* analitike čini prediktivna analitika bazirana na nadgledanom mašinskom učenju [2]. Mašinsko učenje je oblast veštačke inteligencije koja omogućava računarima da razmišljaju i uče sami. Nadgledano mašinsko učenje predstavlja podoblast mašinskog učenja, karakteriše ga podela podataka na skup ulaznih i izlaznih podataka. Ulazni podaci predstavljaju skup nezavisnih varijabli koje se nazivaju atributima. Izlazni podaci predstavljaju skup zavisnih varijabli koje se nazivaju i ciljnim promenljivim. Uz pomoć tehnike nadgledanog mašinskog učenja, moćnih softverskih alata i velikih količina podataka, moguće je kreiranje uspešnih prediktivnih modela u raznim sferama vazdušnog saobraćaja.

Konstantnim porastom putničke potražnje, porastom potražnje za brzu isporuku pošte i robe, razvojem sportske i generalne avijacije kao i ubrzanim razvojem bespilotnih vazduhoplova, danas sve veći problem u sektoru vazdušnog saobraćaja predstavlja zagušenje. Zagušenje možemo podeliti u četiri vrste: zagušenje u vazdušnom prostoru, zagušenje na aerodromima, zagušenje u okolini aerodroma, zagušenje na putničkim terminalima i u okolini terminala.

Zagušenje u vazdušnom prostoru možemo definisati kao pojavu povećanog obima vazdušnog saobraćaja koji je veći ili jednak od unapred definisanog kapaciteta određenog vazdušnog prostora. Za opslugu i regulisanje zagušenja u vazdušnom prostoru obično je zadužena oblasna kontrola leta. Zagušenje na aerodromima možemo definisati kao pojavu povećanog obima vazdušnog saobraćaja koji je veći ili jednak od kapaciteta opsluge aerodroma, kao i povećanog nivoa opsluge na aerodromima. Za zagušenje na aerodromima obično je zadužena aerodromska kontrola leta. Zagušenje u okolini aerodroma predstavlja zagušenje u vazdušnom prostoru koji se nalazi u blizini aerodroma. Za regulaciju zagušenja u okolini aerodroma obično je zadužena prilazna i oblasna kontrola leta. Zagušenje na putničkim terminalima i u okolini terminala predstavlja povećan obim saobraćaja, povećan broj putnika na putničkim terminalima i povećan obim drugih vidova saobraćaja, koji opslužuju aerodrom (taksi vozila, železnica, javni prevoz, itd.) i koji je veći ili jednak od kapaciteta opsluge. Glavnu karakteristiku navedenih vrsta zagušenja predstavlja njihova međusobna zavisnost. Smanjenjem pojave jedne grupe moguće je smanjiti pojavu zagušenja druge grupe. Pravilnim upravljanjem omogućava se da se vazdušni saobraćaj obavlja prema unapred definisanim planovima, bez pojave odstupanja od reda letenja, smanjuju se kašnjenja, smanjuje se opterećenje zaposlenih u vazdušnom saobraćaju, smanjuje se kompleksnost saobraćaja i povećava se kvalitet usluge.

Cilj ovog istraživanja je analiziranje raspoloživih podataka o vazdušnom saobraćaju na aerodromima i ispitivanje mogućnosti predikcije kašnjenja letova u dolasku na aerodromu Denver uz pomoć nadgledanog mašinskog učenja. Modeli mašinskog učenja su razvijani i primenjavani u programskom jeziku Python 3.0, u radnom okruženju *Spyder* uz pomoć različitih biblioteka. Korištene su biblioteke za obradu podataka *Pandas* i *Numpy*, za vizualizaciju *Matplotlib* i *Seaborn* i za modelovanje *Scikit-Learn*.

Rad je strukturiran na sledeći način: u drugoj sekciji opisana je metodologija korišćena u istraživanju, u trećoj sekciji predstavljeni su najznačajniji rezultati istraživanja, dok su u poslednjoj sekciji izdvojeni najvažniji zaključci.

## 2. METODOLOGIJA

Proces mašinskog učenja korišćen u radu sastoji se od sledeće četiri faze. Prva faza predstavlja pripremu podataka, koja uključuje prikupljanje podataka, analizu, čišćenje podataka i predprocesiranje. Druga faza obuhvata izgradnju i definisanje modela mašinskog učenja. Treća faza obuhvata testiranje modela. Četvrta faza obuhvata primenu modela.

„Izgradnja svakog od modela mašinskog učenja sastojala se od sledećih faza:

1. Definisanje cilja modela, u skladu sa ciljevima prediktivne analitike,
2. Izbor ciljne promenljive, tj. atributa iz skupa podataka čiju vrednost želimo da predvidimo primenom modela mašinskog učenja,
3. Izbor algoritma nadgledanog mašinskog učenja, u skladu sa prirodom ciljne promenljive i atributâ,
4. Izbor relevantnih atributa skupa podataka,
5. Priprema skupova podataka za učenje i za testiranje modela, prema zahtevima izabranog algoritma,
6. Podešavanje modela, tj. vrednosti hiperparametara specifičnih za svaku vrstu algoritma mašinskog učenja,
7. Učenje (treniranje) modela – primena izabranog algoritma mašinskog učenja na skup podataka za učenje, u cilju dobijanja hiperparametara modela“ [2].

Cilj prediktivnih modela je da predvide potencijalna kašnjenja na aerodromu „Denver International Airport“ (DEN). U istraživanju su razvijene dve grupe prediktivnih modela. Jedna grupa modela kao ciljnu varijablu ima atribut „Da li let kasni“, dok druga grupa modela ima ciljnu varijablu „Da li let kasni duže od 15

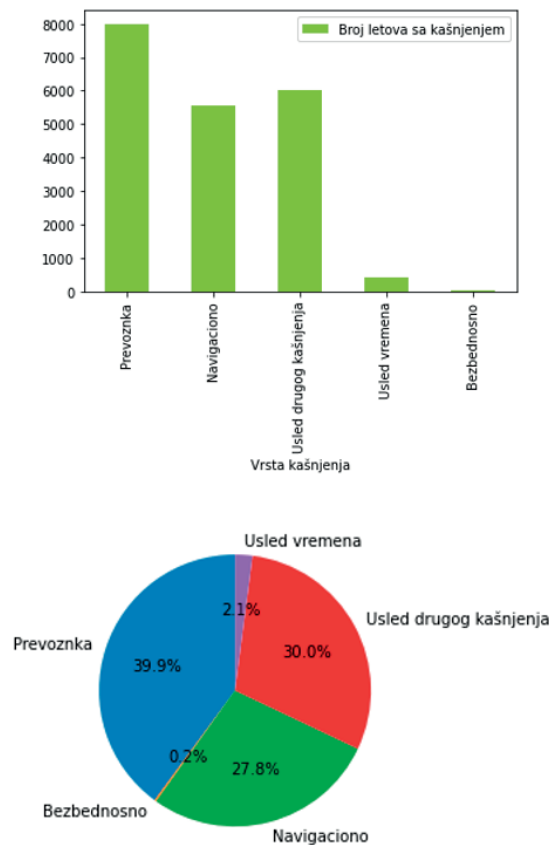
minuta“. Ciljne varijable su postavljene u formi pitanja, pa su zbog toga binarne, gde „1“ označava potvrđnu vrednost, a „0“ označava odričnu vrednost. Varijabla „Da li let kasni“ uzima vrednost „1“ ukoliko je vazduhoplov sleteo nakon planiranog vremena leta, u suprotnom uzima vrednost „0“. Varijabla „Da li let kasni duže od 15 minuta“ uzima vrednost „1“ ukoliko je vazduhoplov sleteo 15 minuta nakon predviđenog vremena sletanja, u suprotnom uzima vrednost „0“. Ciljana varijabla „Da li let kasni duže od 15 minuta“ je uzeta u razmatranje iz razloga što se u vazduhoplovstvu kašnjenje leta definiše kao sletanje vazduhoplova 15 minuta nakon planiranog vremena. Ukoliko vazduhoplov sleti u intervalu od 15 minuta, smatra se da je let na vreme (eng. *On time*). Kašnjenja manja od 15 minuta smatraju se slučajnim kašnjenjima.

Za algoritme mašinskog učenja izabrani su algoritmi klasifikacije: drvo odlučivanja (eng. *Decision Tree*), *k* najbližih suseda (eng. *K-Nearest Neighbours*), naivni Bajes (eng. *Naive Bayes*), nasumična šuma (eng. *Random Forest*).

Izbor relevantnih atributa izvršen je nakon detaljnije analize podataka. Na osnovu korelacije atributa i ciljne promenljive određena je relativnost atributa. Analiziranjem podataka o saobraćaju na aerodromu DEN, došlo se do zaključka da se kašnjenja mogu podeliti u pet grupa:

- kašnjenje prevoznika (eng. *Carrier Delay*),
- kašnjenje usled lošeg vremena (eng. *Weather Delay*),
- navigaciono kašnjenje (eng. *Navigation Delay*),
- kašnjenje usled drugog kašnjenja (eng. *Late Aircraft Delay*),
- bezbednosno kašnjenje (eng. *Security Delay*).

Na slici 1 prikazan je broj letova koji kasne, prema vrstama kašnjenja, i njihov udeo u ukupnom kašnjenju u periodu od 1.1.2022. do 1.4.2022. Na slici 2 prikazano je prosečno kašnjenje u minutima, po vrstama kašnjenja, i njihov udeo u ukupnom kašnjenju u periodu od 1.1.2022. do 1.4.2022.



Slika 1. Broj letova koji kasne, prema vrstama kašnjenja, i njihov udeo u ukupnom kašnjenju u periodu od 1.1.2022. do 1.4.2022.

Slika 2. Prosečno kašnjenje u minutima, prema vrstama kašnjenja, i njihov udeo u ukupnom kašnjenju u periodu od 1.1.2022. do 1.4.2022.

Na osnovu gornje analize, u modelu su korišćeni sledeći atributi:

- „Prevoznik“-prevoznik koji vrši usluge leta a na posredan način opisuje kašnjenje prevoznika;
- Temperatura“, „Vlažnost vazduha“, „Brzina vetra“, „Udari vetra“, „Pravac vetra“, „Vidljivost“ i „Vremensko stanje“ – podaci o vremenu u svakom satu u okolini aerodroma;
- „Polazni aerodrom“ - atribut koji na posredan način opisuje navigaciono kašnjenje;
- „Broj vazduhoplova“ - predstavlja obim saobraćaja, tj. broj vazduhoplova koji u istom satu sleću na DEN aerodrom. Ovaj atribut na posredan način opisuje kašnjenje leta usled drugog kašnjenja;
- Atributi koji karakterišu saobraćaj: „Sat dolaska vazduhoplova na aerodrom“, „Dan dolaska na aerodrom“, „Mesec dolaska na aerodrom“, „Distanca“ i „Kašnjenje pri poletanju“.

Priprema podataka za obučavanje i testiranje počinje prikupljanjem podataka. Korišćeni podaci preuzeti su sa zvaničnog sajta Ministarstva saobraćaja Sjedinjenih Američkih Država, Zavod za statistiku saobraćaja (eng. *United States Department of Transportation, Bureau of Transportation Statistic*), dok su podaci o vremenu preuzeti sa sajta organizacije *Visual Crossing*. Podaci obuhvataju podatke o saobraćaju i podatke o vremenu za period od 1. januara 2022. godine, do 1. aprila 2022. godine.

Da bi model bio efikasan potrebno je nakon prikupljanja podataka izvršiti prilagođavanje podataka algoritmima mašinskog učenja. Atributi „Polazni aerodrom“ i „Prevoznik“ uzimaju nominalne vrednosti, stoga ih je potrebno prilagoditi modelu. Korišćenjem biblioteke *Pandas*, i njene metode `get_dummies()` [3], konvertujemo kategoričke podatke u „dummy“ vrednosti. Izvršena je i modifikacija atributa „Brzina vetra“, „Udari vetra“ i „Pravac vetra“, kombinovanjem istih i konstruisanjem dva nova atributa: „Brzina vetra prema severu“ i „Udari vetra prema severu“.

Skup podataka je podeljen na dve grupe: skup podataka za trening, koji čini 80% uzorka i skup podataka za testiranje koji čini 20% uzorka. Odabir podataka za skupove je nasumično urađen. Uzorak sadrži podatke o 59.438 letova [4] [5]. Trening skup sadrži podatke o 47.550 letova, test skup sadrži podatke o 11.888 letova. Podela skupa podataka je izvršena pomoću *Pandas* biblioteke i metode `train_test_split()` [3].

Za proveru verodostojnosti modela korišćena je tehnika unakrsne validacije (eng. *cross-validation*) koja je primenjivana na trening skupu podataka. Kao mera za ocenjivanje uspešnosti predikcije korišćena je preciznost modela (eng. *Accuracy*) (1).

$$\text{Preciznost} = \frac{\text{Ukupan broj pogodnih predikcija}}{\text{Ukupan broj predikcija}} \quad (1)$$

Da bi obučili model potrebno je algoritme mašinskog učenja primeniti nad trening podacima. Ciljnu varijablu modela predstavlja ciljna varijabla trening uzorka. Da bismo odredili najbolji algoritam za obučavanje potrebno je ispitati performanse svakog od algoritama. Ispitivanje performansi algoritama se vrši na test uzorku. Test uzorak predstavlja podatke koji nisu učestvovali u formiranju modela. Međusobnim upoređivanjem performansi algoritama određen je algoritam koji daje najbolje performanse. Uvidom u vrednosti ciljne varijable, utvrđeno je da veći broj letova ne kasni. Smislenost modela određuje se njegovim poređenjem sa *null* modelom.

„*Null* model u klasifikacionim problemima za ishod predikcije uvek uzima najčešću vrednost.” [1] U ovom slučaju je to broj letova koji ne kasni. Verovatnoća pojave letova koji ne kasne predstavljena je *null* modelom (2). Poređenjem tačnosti kreiranog modela sa tačnošću *null* modela može se zaključiti da je kreirani model bolji, što je prvi korak u predikciji. Rezultati modela će biti prikazani preko pita grafikona (eng. *Pie chart*). Vrednosti prikazane na pita grafikonima su vrednosti iz matrice konfuzije (eng. *Confusion matrix*).

$$\text{null} = \frac{\text{Broj letova koji ne kasne}}{\text{Ukupan broj letova}} \quad (2)$$

### 3. REZULTATI I ANALIZA REZULTATA

Krajnji rezultati istraživanja podeljeni su u dva dela. Prvi deo se odnosi na rezultate sa ciljnom varijablom „Da li let kasni“. Drugi deo rezultata odnosi se na ciljnu varijablu „Da li je let kasnio duže od 15 minuta“.

U tabeli 1 prikazani su rezultati merenja performansi modela za svaki od algoritama mašinskog učenja za ciljnu varijablu „Da li let kasni“. Najbolje performanse ostvaruje model „Nasumična šuma“ (eng. *Random Forest*). Preciznost modela, prema unakrsnom vrednovanju, iznosi 0.84328. Preciznost modela je bolja od preciznosti *null* modela za 23,84 %. *Naive Bayes* se ne može koristiti u predviđanju, jer je model baziran na ovom algoritmu lošiji od *null* modela.

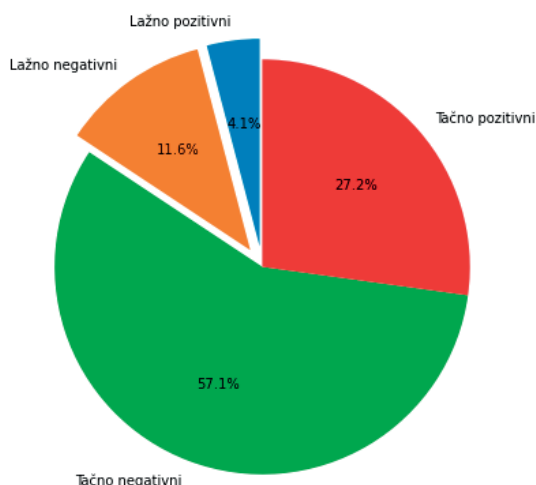
**Tabela 1:** Prikaz performansi modela mašinskog učenja za ciljnu varijablu „Da li let kasni“

	Preciznost trening uzorka	Preciznost Test uzorka	Preciznost unakrsnog vrednovanja (CV=10)	<i>Null</i> model
<i>Naive Bayes</i>	0,61247	0,58403	0,61056	0,60468
Drvo odlučivanja	0,82042	0,810313	0,81243	0,60468
K najbližih suseda	0,81659	0,79828	0,80116	0,60468
Nasumična šuma	0,99891	0,83025	0,84328	0,60468

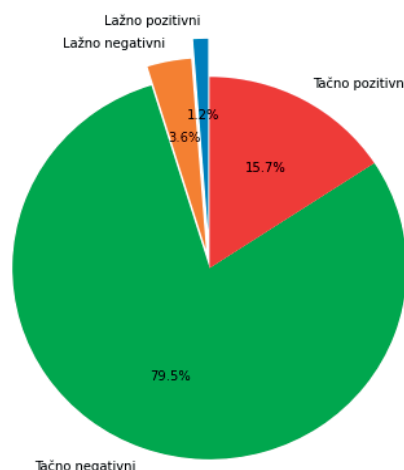
U tabeli 2 prikazani su rezultati merenja performansi modela za svaki od algoritama mašinskog učenja za ciljnu varijablu „Da li let kasni duže od 15 minuta“. Najbolje performanse ostvaruje model „Nasumična šuma“ (eng. *Random Forest*). Preciznost modela, prema unakrsnom vrednovanju, iznosi 0,95575. Preciznost modela je bolja od preciznosti *null* modela za 16,75 %. *Naive Bayes* se ni ovde ne može koristiti u predviđanju, jer je model lošiji od *null* modela.

**Tabela 2:** Prikaz performansi modela mašinskog učenja za ciljnu varijablu „Da li let kasni duže od 15 minuta“

	Preciznost trening uzorka	Preciznost Test uzorka	Preciznost unakrsnog vrednovanja (CV=10)	<i>Null</i> model
<i>Naive Bayes</i>	0,77213	0,75009	0,76032	0,78823
Drvo odlučivanja	0,95649	0,94877	0,95182	0,78823
K najbližih suseda	0,92177	0,91732	0,92101	0,78823
Nasumična šuma	0,99785	0,95224	0,95575	0,78823



Slika 3. Grafički prikaz rezultata dobijenih modelom predikcije baziranom na algoritmu „Nasumična šuma“ za ciljnu varijablu „Da li let kasni“



Slika 4. Grafički prikaz rezultata dobijenih modelom predikcije baziranom na algoritmu „Nasumična šuma“ za ciljnu varijablu „Da li let kasni duže od 15 minuta“

Na slici 3 i na slici 4 prikazan je najbolji rezultat modela predikcije. Rezultat je prikazan na osnovu matrice konfuzije (eng. *Confusion matrix*). Na slici 3 su prikazani rezultati modela za ciljnu promenljivu "Da li let kasni". Na slici 4 su prikazani rezultati modela za ciljnu promenljivu "Da li let kasni duže od 15 minuta". Na grafikonu su prikazana 4 tipa rezultata: tačno pozitivni rezultati (broj letova za koje je model predikcije pogodio da kasne), tačno negativni (broj letova za koje je model predikcije pogodio da ne kasne), lažno pozitivni (broj letova koji kasne, a model je predvideo da ne kasne) i lažno negativni (broj letova koji ne kasne, a model je predvideo da kasne). Obzirom da je preciznost modela mašinskog učenja za predikciju ciljne varijable "Da li let kasni" 84,328%, u 15,672% slučajeva model će davati loše rezultate. Netačni rezultati mogu dovesti do lošeg raspoređivanja ljudskih resursa, što povećava troškove. Aviokompanije, mogu na pogrešan način da planiraju letove, što može dovesti do povećanja troškova kompanija.

Međutim, menadžment kontrole letenja i menadžment aerodroma određuje smene zaposlenih na osnovu intervala koji su veći od 15 minuta. Prema tome kašnjenje koje je manje od 15 minuta neće značajno uticati na raspoređivanje ljudskih resursa. Aviokompanije treba da planiraju letove tako da kašnjenje ne bude veće od 15 minuta. Stoga, model mašinskog učenja sa ciljnom varijablom "Da li let kasni duže od 15 minuta", predstavlja model koji bi mogao da se koristi kao pomoć u planiranju letova i raspoređivanju osoblja. Njegova preciznost iznosi 95,575%, dok u 4,425% slučajeva daje loše rezultate, prema tome rizik od greške je manji, i troškovi koje snose aerodrom, kontrola letenja i aviokompanije su znatno manji nego u prethodnom modelu, ali bi i dalje postojali. Model mašinskog učenja, za predikciju ciljne varijable "Da li let kasni", mogao bi da pruži istraživačima bolji uvid u razloge kašnjenja kao i podlogu za razvoj sofisticiranijeg modela predviđanja.

#### 4. ZAKLJUČAK

U okviru istraživanja u radu, za predikciju kašnjenja letova, predložena je metoda prediktivne analize bazirana na nadgledanom mašinskom učenju. Implementacijom opisane metode pokazano je da algoritam *Naive Bayes* ne daje dobre rezultate. Ostali algoritmi klasifikacije, koji su navedeni u radu, daju dobre rezultate. Najbolje rezultate predikcije daje algoritam „nasumična šuma“. U radu je opisana primena metodologije primenjene u istraživanju. Radi izbora relevantnih atributa prediktivnih modela, u istraživanju je izvršena i deskriptivna statistika, korišćenjem Python-ovih biblioteka „Matplotlib“ i „Seaborn“. Manipulacija podacima je izvršena korišćenjem Python-ovih biblioteka „Pandas“ i „Numpy“. Primena algoritama mašinskog učenja na raspoloživom skupu podataka izvršena je korišćenjem Python-ove biblioteke Scikit-Learn. Predložena metodologija verifikovana je u realizovanoj studiji slučaja.

Uz pomoć prediktivnih modela opisanih u radu i na osnovu ulaznih atributa, moguće je odrediti broj letova koji kasne u određenom vremenskom periodu. Takođe, moguće je odrediti da li će određeni let kasniti. Korišćenjem ovih predikcija, zagušenja na aerodromima se mogu smanjiti. Model omogućava operativnom osoblju da se pripremi za scenarije u kojima određeni let kasni. Menadžment aerodroma i kontrole leta uz pomoć modela imaju mogućnost da ravnomerno rasporede smene zaposlenima, kako bi se izjednačila radna opterećenja zaposlenima. Kontrolori leta bi imali bolji uvid u situaciju. Aviokompanije bi mogle da izbegnu kašnjenje svojih letova, tako što će moći da predvide da li će određeni let kasniti. Model će istraživačima dati bolji uvid u razloge kašnjenja letova.

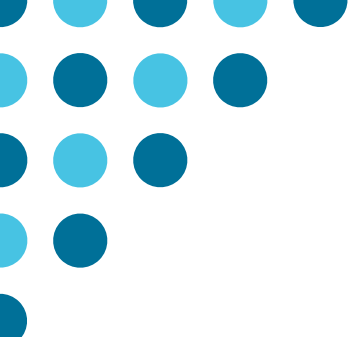
#### ZAHVALNICA

Ovaj rad delimično je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije, u okviru projekta pod brojem 036012.

#### LITERATURA

- [1] Marin, I., Shukla, A., et al., „Big Data Analysis with Python“, Packt, Birmingham – Mumbai, 2019.
- [2] Janković, S., Kukic, K., Uzelac, A. & Maraš, V. (2019). „Predikcija nivoa saobraćaja u lokalnoj računarskoj mreži primenom nadgledanog mašinskog učenja“. *XXXVII Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju – PosTel 2019, Beograd, 3. i 4. decembar 2019*
- [3] Ozdemir, S., (2016). „Principles of Data Science“ *Packt Publishing Ltd.*,
- [4] Podaci o vazдушnom saobraćaju Ministarstva Saobraćaja Sjedinjenih Američkih Država, Zavod za statistiku saobraćaja dostupni na „<https://transtats.bts.gov>“, datum pristupa 15. Maj 2022. godine.
- [5] Podaci o vremenu organizacije *Visual Crossing* dostupni na „<https://www.visualcrossing.com/weather>“, datum pristupa 15. Maj 2022. godine.





# SIMULACIJA I STOHAŠTIČKI MODELI

## SIMULATION AND STOCHASTIC MODELS





## BOUNDS ON THE COUPON COLLECTOR PROBLEM

JELENA JOCKOVIĆ<sup>1</sup>, BOJANA TODIĆ<sup>2</sup>

<sup>1</sup> University of Belgrade – Faculty of Mathematics, jjocko@matf.bg.ac.rs

<sup>2</sup> University of Belgrade – Faculty of Mathematics, btodic@matf.bg.ac.rs

---

**Abstract:** *The coupon collector problem (CCP) can be formulated as follows: a company issues coupons of  $n$  different types, each with a particular probability of being drawn. The object of interest is the number of coupons that a collector needs to buy in order to obtain the full collection ( $W_n$ ). In the case of CCP with unequal probabilities, the exact distribution function of  $W_n$  is computationally expensive, which is important in some fields of application, and reported in the literature. A way to overcome this difficulty is obtaining upper and lower bounds for this quantity, which are easy to compute, even for large  $n$ . We will start from several known bounds for the cumulative distribution function of  $W_n$ , and further computationally simplify, and extend these results to some other cases of the CCP.*

**Keywords:** *Coupon collector problem, unequal probabilities, waiting time, bounds*



## METODOLOGIJA KREIRANJA VEŽBI NA BRODSKOM SIMULATORU

### METHODOLOGY OF CREATING EXERCISES ON A SHIP BRIDGE SIMULATOR

SLOBODAN RADOJEVIĆ<sup>1</sup>, MILAN KRESOJEVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet odbrane – Vojna akademija, Beograd, [slobodan.radojevic@va.mod.gov.rs](mailto:slobodan.radojevic@va.mod.gov.rs)

<sup>2</sup> Univerzitet odbrane – Vojna akademija, Beograd, [milan.kresojevic@va.mod.gov.rs](mailto:milan.kresojevic@va.mod.gov.rs)

**Rezime:** Metodologija kreiranja vežbi na brodskom simulatoru usko je povezana sa potrebama obuke pomoraca - kandidata. Da bi uspešno kreirali vežbu na simulatoru koja bi ispunila sa jedne strane tražene zahteve međunarodnih, evropskih i nacionalnih propisa, sa druge strane potrebe kandidata kao i detaljni nastavni plan i program neophodno je široko znanje i metodološko iskustvo nastavnika. Metodologija kreiranja vežbe zasniva se na definisanju parametara na samom brodskom simulatoru, odabiru plovidbenog područja i određenog tipa broda, zatim i definisanja scenarija vežbe sa svim pratećim parametrima. Upravo definisanje parametara predstavlja metodološki izazov nastavnika. Upotrebom naučenih lekcija, metode višestruke studije slučaja i eksperimenata na simulatoru nastavnik će kreirati svrsishodan i upotrebljiv scenario vežbe na brodskom simulatoru.

**Ključne reči:** metodologija; brodski simulator; višestruka studija slučaja; eksperiment na simulatoru.

**Abstract:** The methodology of creating exercises on a ship bridge simulator is closely related to the practical result of the training that is necessary for the seafarers - candidates. In order to successfully create a simulator exercise that would meet on the one hand the required requirements of International, European and national regulations and legislative, on the other hand, the needs of candidates and the detailed teaching syllabus require extensive knowledge and methodological experience of teachers. The methodology of creating the exercise is based on defining the parameters on the ship simulator itself, selecting the sailing area and a certain type of ship, and then defining the exercise scenario with all the accompanying parameters. Defining the parameters is a methodological challenge for the teacher. Using the lessons learned, multiple case studies method, and simulator-based experiments, the teacher will create a meaningful and usable exercise scenario on a ship simulator.

**Keywords:** methodology; ship bridge simulator; multiple case study; simulator-based experiment.



## PRIPREMA SIMULACIJE NA BRODSKOM SIMULATORU NAVI-TRAINER PROFESSIONAL 5000

## PREPARATION OF THE SIMULATION ON THE SHIP SIMULATOR NAVI-TRAINER PROFESSIONAL 5000

MILAN KRESOJEVIĆ<sup>1</sup>, SLOBODAN RADOJEVIĆ<sup>2</sup>

<sup>1</sup> Univerzitet Odbrane – Vojna akademija, Beograd, [milan.kresojevic@va.mod.gov.rs](mailto:milan.kresojevic@va.mod.gov.rs)

<sup>2</sup> Univerzitet Odbrane – Vojna akademija, Beograd, [slobodan.radojevic@va.mod.gov.rs](mailto:slobodan.radojevic@va.mod.gov.rs)

**Rezime:** Brodski simulator NAVI-TRAINER PROFESSIONAL 5000 predstavlja softversko i hardversko rešenje koje se uspešno može koristiti kao kompletan simulator mosta broda.. Sistem kontroliše instruktor sa instruktorske stanice gde kreira različite scenarije, vežbe ili zadatke koje trenirano lice izvršava na virtualnom komandnom mostu. Definisane ulaznih parametara na simulatoru radi obuke kadeta i studenata u oblasti pomorstva i rečnog brodarstva treba da budu što približniji realnim situacijama u kojima brod može da se nadje na moru ili reci. Iako brodski simulator poseduje različite mogućnosti kreiranja vežbi, sama postavka odnosno definisani parametri determinišu razliku između kvaliteta obuke. Pored mogućnosti odabira matematičkog modela broda koji konkretnom korisniku najviše odgovara, moguće je odabrati interesni rejon-akvatoriju, dok je poželjno da se parametri sredine kreiraju za svaku vežbu posebno. Pored definisanja parametara sredine koji se mogu preuzeti iz različitih dostupnih javnih izveštaja, posebno je interesantna upotreba modela predviđanja nastanka određenije pojave. Na primer, ukoliko govorimo o vodostaju upotreba autoregresivnog modela za predviđanje vodostaja donosi novi kvalitet obuke, odnosno podiže nivo znanja polaznika i daje mu mogućnost snalaženja u situaciji kada nema dostupne sve podatke koji su mu potrebni za donošenje odluke, posebno za kritične manevre broda.

**Ključne reči:** Brodski simulator; Priprema simulacije; Ulazni parametri; Autoregresivni model.

**Abstract:** Ship simulator NAVI-TRAINER PROFESSIONAL 5000 is a software and hardware solution that can be successfully used as a full bridge simulator. Defining the input parameters on the simulator for the training of cadets and students in the field of maritime and river shipping should be as close as possible to real situations in which the ship can be found at sea or river. Although the ship simulator has different possibilities for creating exercises, the setting itself or the defined parameters determine the difference in the quality of training. In addition to the possibility of choosing the mathematical model of the ship that suits the particular user the most, it is possible to choose an interesting area - the water area, while it is desirable that the parameters of the environment are created for each exercise separately. In addition to defining the parameters of the environment that can be taken from various available public reports, it is particularly interesting to use the model of predicting the occurrence of the phenomenon. For example, if we talk about the water level, the use of the autoregressive model for water level prediction brings a new quality of training, i.e. raises the level of knowledge of participants and gives them the opportunity to cope in a situation where they do not have all the information they need to make decisions, especially for critical ship maneuvers.

**Keywords:** Ship simulator; Simulation preparation; Input parameters; Autoregressive model.





## THE SIMULATION MODEL OF BLOCKCHAIN TRANSACTION

ANA SAVIĆ<sup>1</sup>, HANA JOVANOVIĆ<sup>2</sup>, GORAN BJELOBABA<sup>3</sup>, NIKOLA POPOVIĆ<sup>4</sup>

<sup>1</sup> Academy of Technical and Art Applied Studies – School of Electrical and Computer Engineering Belgrade, Serbia  
ana.savic@viser.edu.rs

<sup>2</sup> Comtrade Information Technology School of Applied Studies, Belgrade, Serbia, hana.stefanovic@its.edu.rs

<sup>3</sup> University of Belgrade – Faculty of organizational science, gbjelobaba@gmail.com

<sup>4</sup> Alfa BK University – Faculty of Mathematics and Computer Science Belgrade, Serbia, nikolap6901@gmail.com

**Abstract:** *In this paper the simulation blockchain-based model is created and tested through a few transactions among the participants. The participants are able to send and receive coins or to mine blocks to earn coins, while all transactions are stored in the blocks' data, including the address of sender and the address of the receiver. The address is a hash value of a public key for asymmetric cryptography. The RSA (Rivest-Shamir-Adleman) asymmetric algorithm is used, including public key and private key. The model is created in CrypTool2 software, including three blocks and smaller numbers for generating the RSA public and private key pairs than in real blockchain transactions, in order to minimize the simulation time. Some attacks are also simulated, and those transactions are not accepted because of invalid signature, since the attacker does not have someone's private key.*

**Keywords:** *Blockchain technology, RSA algorithm, Valid and Invalid Transactions, CrypTool2*

### 1. INTRODUCTION

Blockchain is a set of contemporary technologies that in synergy create a network that ensures trust among users. Blockchain is a decentralized and distributed database in which data cannot be changed or deleted and which enables the verification of transactions [1]. Three basic parts can be singled out: block, chain and network. It is based on distributed general ledger (DLT) technology to store cryptographically verified user group data, as agreed in a predefined network protocol. Cryptographic techniques enable encryption of all important data records in the blockchain, important for consistent data and record integrity [2].

A block is a list of transactions recorded in a book/register over a period of time. Transaction data is stored on various computers on the network, which are connected using the peer-to-peer protocol. This registry can be thought of as a record book that notes and stores all transactions between participants chronologically, and all online users, called nodes, have an identical copy of the book. Also, each node shares the same copy of the data, i.e. digital register (Digital Ledger) [3]. The network consists of "full nodes". Each node includes a complete record of all transactions that have ever been noted in that blockchain. The function of nodes is to continuously check the authenticity of records in the chain, and in the case of verification fail, to reject the proposed data blocks. Therefore, this approach to data storage is considered far safer than centralized databases.

The approach of selecting an initially valid block of transactions is called proof-of-work [4] and practically protects the network from misuse. In the end, the validated block gets its own unique timestamp and signature (hash), and as such is propagated to other nodes in the network [5]. When adding each subsequent validated block [8], the certificate counter for the previously entered blocks is incremented, and the probability that they are false is further reduced. Over time, the reliability of recorded data increases.

Changing data in the blockchain is done according to predefined rules. Changes are passed to all nodes to update the local copy of the data [6]. Once a transaction has been saved and validated by all nodes in the network, it is no longer possible to change the data of that transaction, or it is difficult to change it. The process of validating these transactions is called mining and is based on one of the consensus algorithms on the basis of which an agreement is reached between the nodes when adopting a new block [1]. After confirmation, there is a connection with other transactions in the new block which is joined to the blockchain. Procedure ensures that each block is formed as that the previous and the next are irrefutably connected, thus forming a chain of blocks or a blockchain. A chain is a hash that connecting blocks, or

mathematically "chaining". A hash [7] is formed from data that was existing in the block before. A hash can be viewed as a fingerprint of data that locks into blocks according to order and time. The hash forms a one-way function and it cannot be decrypted, while the hash function creates an algorithm to map data of any size to a series of fixed-size bits.

A blockchain consists of three layers: a protocol layer, a network layer, and an application or business layer. Each layer adds different components to the blockchain in order to develop it [8].

There are other divisions of blockchain networks, into public, private and consortium [10], where the consortium is a type of blockchain where the consensus process is controlled by a pre-selected set of nodes [11].

## 2. BLOCKCHAIN SIMULATION MODEL

A decimal system numbering of up to three levels is used for the headings of the sections:

The simulation model with three blocks (block ID 0, block ID 1 and block ID 2) and two participants, Ana and Goran, is presented in Fig1.

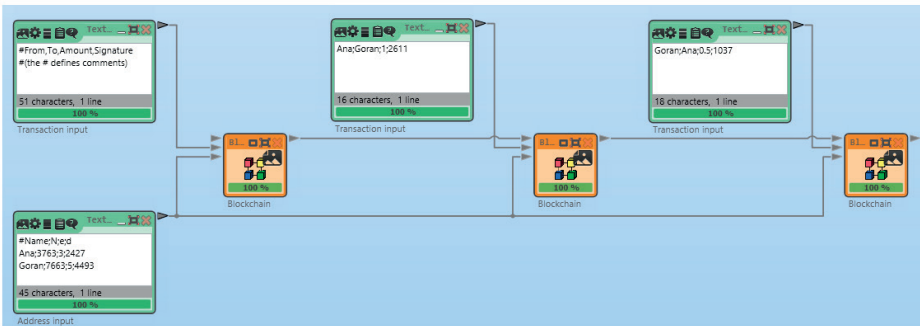


Fig.1. The simulation blockchain-based model with 3 blocks and 2 participants

The transactions and balance in each block is given in Figs 2.-5.

Blockchain											
<b>Block header</b>	Block Id: 0 Block hash: 00009DA60A28C8840CA7										
	Previous block hash: 0 Timestamp: 4/17/2022 11:40:13 AM										
	Nonce: 51,055										
<b>Statistics</b>	Transactions: 1 Failed transactions: 0										
	Hash algorithm: SHA256 (10 byte) Hashes/sec:										
	Mining difficulty: 16 bit										
<b>Transactions</b>	<table border="1"> <thead> <tr> <th>Sender</th> <th>Recipient</th> <th>Coins</th> <th>Signature</th> <th>Status</th> </tr> </thead> <tbody> <tr> <td>MINING_REWARD</td> <td>Ana (N=3763, e=3)</td> <td>2.55</td> <td>0</td> <td>Successful</td> </tr> </tbody> </table>	Sender	Recipient	Coins	Signature	Status	MINING_REWARD	Ana (N=3763, e=3)	2.55	0	Successful
Sender	Recipient	Coins	Signature	Status							
MINING_REWARD	Ana (N=3763, e=3)	2.55	0	Successful							

Fig2. Block ID 0 transactions

Blockchain							
<b>Block header</b>	Block Id: 0 Block hash: 00009DA60A28C8840CA7						
	Previous block hash: 0 Timestamp: 4/17/2022 11:40:13 AM						
	Nonce: 51,055						
<b>Statistics</b>	Transactions: 1 Failed transactions: 0						
	Hash algorithm: SHA256 (10 byte) Hashes/sec:						
	Mining difficulty: 16 bit						
<b>Transactions</b>	<table border="1"> <thead> <tr> <th>Name</th> <th>Balance</th> </tr> </thead> <tbody> <tr> <td>Ana</td> <td>2.55</td> </tr> <tr> <td>Goran</td> <td>0</td> </tr> </tbody> </table>	Name	Balance	Ana	2.55	Goran	0
Name	Balance						
Ana	2.55						
Goran	0						

Fig3. Block ID 0 balance



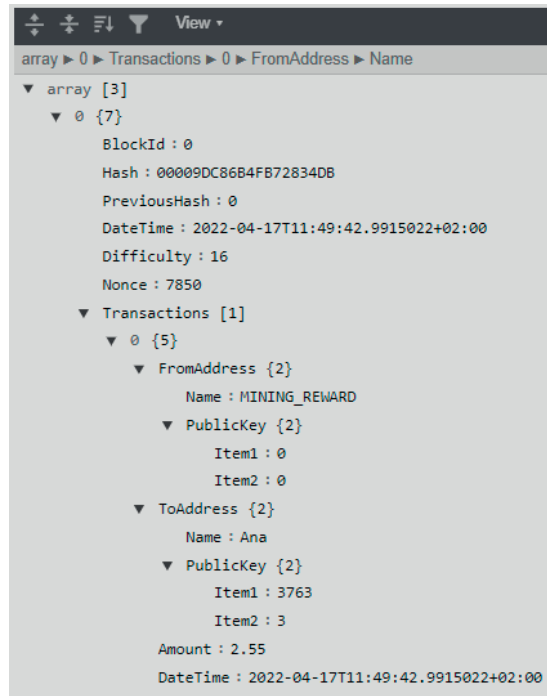


Fig7. Transaction data in JSON format

### 3. ADDING A NEW PARTICIPANT

The new participant's (Hana) address is generated, using the RSA key generator, as it is presented in Fig8.

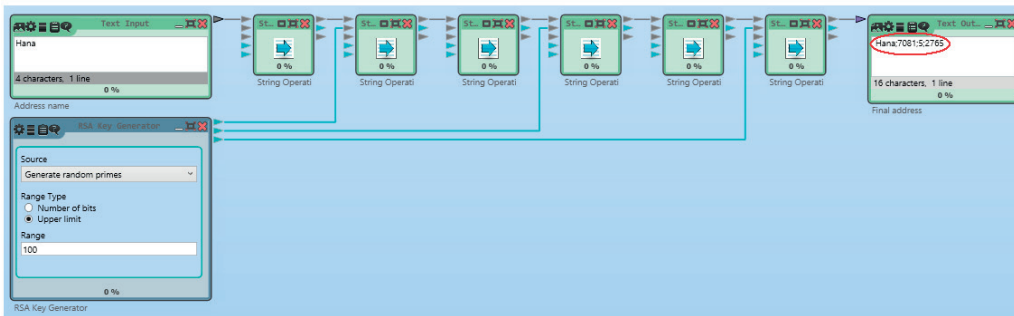


Fig8. Creating the new participant's address

The model with new participant added id presented in Fig9.

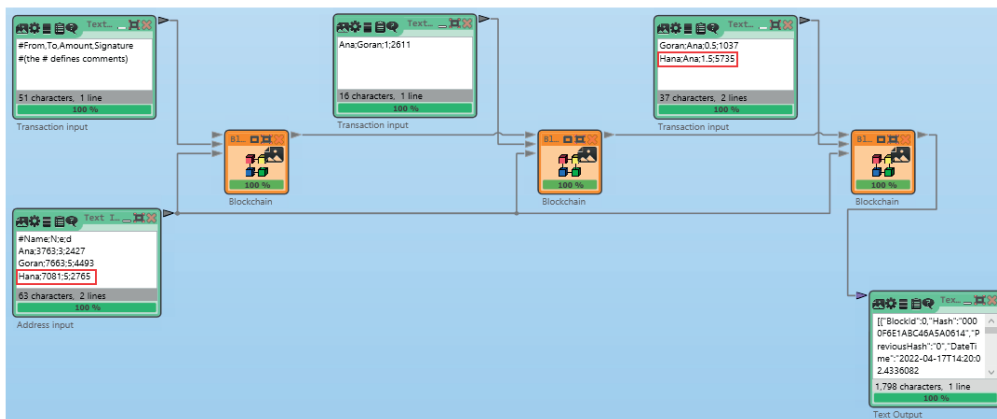


Fig.9. The simulation model with new participant (Hana) and her transactions

The transactions after adding the new participant are presented in Fig10 and Fig.11.

Block header			
Block Id:	0	Block hash:	0000A15393A2F59F6913
Previous block hash:	0	Timestamp:	4/17/2022 2:28:50 PM
Nonce:	12,462		
Statistics			
Transactions:	1	Failed transactions:	0
Hash algorithm:	SHA256 (10 byte)	Hashes/sec:	
Mining difficulty:	16 bit		
Transactions			
Name		Balance	
Ana		2.55	
Goran		0	
Hana		0	

Fig10. Block ID 0 balance after adding the new participant

Block header			
Block Id:	2	Block hash:	00007F7CD73C0AA391C9
Previous block hash:	0000D91CAB7C8C3E3C99	Timestamp:	4/17/2022 2:28:51 PM
Nonce:	201,686		
Statistics			
Transactions:	3	Failed transactions:	0
Hash algorithm:	SHA256 (10 byte)	Hashes/sec:	
Mining difficulty:	16 bit		
Transactions			
Name		Balance	
Ana		3.55	
Goran		1.5	
Hana		3.5	

Fig11. Block ID 2 balance after adding the new participant

#### 4. CONCLUSION

The paper presents the basic models of blocking technologies and the simulation model presented in the paper illustrates several examples of transactions between users. Specifically, the paper presents three users. The simulation model illustrates a successful and unsuccessful financial transaction. Extended simulation model could be applied in practice.

#### REFERENCES

- [1] Z. Misic and B. Mrazovac, "Implementacija Programske Podrške za Manipulaciju Podacima Decentralizovane Aplikacije za Upravljanje Pristupom Zaštićenom Zoni," *Zb. Rad. Fak. Teh. Nauk. u Novom Sadu*, vol. 35, no. 10, pp. 1818–1821, Oct. 2020, doi: 10.24867/09IH03Misic.
- [2] J. Menez, P. van Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*, 5th ed. CRC press Series on Discrete Mathematics and Its Applications, 2001.
- [3] W. Zhao, K. Liu, and K. Ma, "Design of Student Capability Evaluation System Merging Blockchain Technology," in *Journal of Physics: Conference Series*, Mar. 2019, vol. 1168, no. 3, doi: 10.1088/1742-6596/1168/3/032123.
- [4] G. Hileman and M. Rauchs, *GLOBAL BLOCKCHAIN BENCHMARKING STUDY*, vol. 10, no. 2. 2017, pp. 37–38.
- [5] V. J. Morkunas, J. Paschen, and E. Boon, "How blockchain technologies impact your business model," *Bus. Horiz.*, vol. 62, no. 3, pp. 295–306, May 2019, doi: 10.1016/j.bushor.2019.01.009.
- [6] A. Alammery, S. Alhazmi, M. Almasri, and S. Gillani, "Blockchain-based applications in education: A systematic review," *Applied Sciences (Switzerland)*, vol. 9, no. 12. MDPI AG, Jun. 01, 2019, doi: 10.3390/app9122400.
- [7] J. H. Seo, "Information Security and Cryptology – ICISC 2019: 22nd International Conference,

- Seoul, South Korea, December 4–6, 2019, Revised Selected Papers,” 2020, vol. 11975, doi: 10.1007/978-3-030-40921-0.
- [8] G. Bjelobaba, M. Paunovic, A. Savic, H. Stefanovic, J. Doganjic, and Z. M. Bogavac, “Blockchain Technologies and Digitalization in Function of Student Work Evaluation,” *Sustain.*, vol. 14, no. 9, pp. 1–22, 2022, doi: 10.3390/su14095333.
- [9] Antonio Ramón Bartolomé, Carles Bellver, Linda Castañeda, and Jordi Adell, “Blockchain in Education: Introduction and Critical Review of the State of the Art,” *EDUTEC. Rev. Electrónica Tecnol. Educ.*, no. April 2018, pp. 1–14, 2017, doi: 10.21556/edutec.2017.61.
- [10] H. Hyvärinen, M. Risius, and G. Friis, “A Blockchain-Based Approach Towards Overcoming Financial Fraud in Public Sector Services,” *Bus. Inf. Syst. Eng.*, vol. 59, no. 6, pp. 441–456, 2017, doi: 10.1007/s12599-017-0502-4.
- [11] L. Min and G. Bin, “Online teaching research in universities based on blockchain,” *Educ. Inf. Technol.*, Jan. 2022, doi: 10.1007/s10639-022-10889-w.
- [12] J. Sedlmeir, H. U. Buhl, G. Fridgen, and R. Keller, “The Energy Consumption of Blockchain Technology: Beyond Myth,” *Bus. Inf. Syst. Eng.*, vol. 62, no. 6, pp. 599–608, 2020, doi: 10.1007/s12599-020-00656-x.
- [13] S. Alam *et al.*, “A Blockchain-based framework for secure Educational Credentials,” 2021.
- [14] M. Finck, “Blockchains and Data Protection in the European Union,” *Eur. Data Prot. Law Rev.*, vol. 4, no. 1, pp. 17–35, 2018, doi: 10.21552/edpl/2018/1/6.
- [15] R. Chatterjee and R. Chatterjee, “An Overview of the Emerging Technology: Blockchain,” in *2017 3rd International Conference on Computational Intelligence and Networks (CINE)*, Oct. 2017, pp. 126–127, doi: 10.1109/CINE.2017.33.



# STATISTIČKI MODELI

---

# STATISTICAL MODELS



UNIVERZITET U BEOGRADU  
**Ekonomski fakultet**  
UNIVERSITY OF BELGRADE  
Faculty of Economics  
and Business







## ANALIZA BROJA OBOLELIH OD VIRUSA COVID-19 U SRBIJI PRIMENOM ARIMA MODELA<sup>1</sup>

### ANALYSIS OF COVID-19 CASES IN SERBIA USING ARIMA MODELS

EMILIJA MAKSIMOVIĆ<sup>1</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, [emilija.maksimovic@ekof.bg.ac.rs](mailto:emilija.maksimovic@ekof.bg.ac.rs)

**Rezime:** Cilj ovog rada predstavlja konstruisanje ARIMA modela za opisivanje dinamike broja obolelih od virusa COVID-19 u Srbiji. Primenjena je Boks-Dženkinsova strategija modeliranja prilikom odabira adekvatnog ARIMA modela. Prilikom analize korišćeni su kumulativni podaci o broju obolelih na nedeljnom nivou za period 28.12.2020 – 9.5.2022. godine. U radu su ocenjeni ARIMA(1,2,1) model i redukovani ARIMA(4,2,0) model. Dobijeni rezultati za dva različita horizonta predviđanja (6 i 17 nedelja) ukazuju na to da ocenjeni ARIMA(1,2,1) model ima precizniju prognozu stope rasta broja obolelih od virusa COVID-19 u Srbiji.

**Ključne reči:** pandemija COVID-19, ARIMA modeli, predviđanje

**Abstract:** The aim of this paper is to construct an ARIMA model to describe the dynamics of COVID-19 cases in Serbia. The Box-Jenkins modelling approach is applied for selecting an ARIMA model. Weekly cumulative data of COVID-19 cases from 28th December 2020 to 9th May 2022 are used in the analysis. ARIMA (1,2,1) model and reduced ARIMA (4,2,0) model are estimated in the paper. The results obtained for two different forecast horizons (6 and 17 weeks) indicate that estimated ARIMA (1,2,1) model has a more accurate forecasts of the growth rate of COVID-19 cases in Serbia.

**Keywords:** pandemic COVID-19, ARIMA models, forecasting

#### 1. UVOD

Pandemija COVID-19 pogodila je sve zemlje sveta uzrokujući veliki broj ljudskih žrtava. Prema podacima Svetske zdravstvene organizacije, od početka pandemije, obolelo je preko 500 miliona ljudi, dok je oko 6 miliona umrlo od posledica virusa. Tokom poslednje dve godine zemlje su uvodile razne mere u cilju sprečavanja daljeg širenja virusa i očuvanja ljudskih života. Sa druge strane, ubrzani razvoj novih mutacija virusa uticao je na rast broja obolelih, kao i potrebu za njihovu hospitalizaciju, opterećujući zdravstvene sisteme širom sveta.

Od početka pandemije, na nivou pojedinačnih zemalja, prati se dnevni razvoj broja novozaraženih, hospitalizovanih i umrlih ljudi. Ovi podaci obezbeđuju informacije koje su od ključnog značaja za dalje sprovođenje epidemioloških mera pružajući sliku o aktuelnim epidemiološkim kretanjima. Cilj ovog rada predstavlja formiranje ARIMA modela za opisivanje kretanja vremenske serije broja obolelih od COVIDa-19 u Srbiji. Prilikom izbora odgovarajućeg ARIMA modela korišćena je Boks-Dženkinsova strategija modeliranja. Empirijska analiza zasniva se na kumulativnim podacima o broju obolelih na nedeljnom nivou za period 28.12.2020 – 9.5.2022. godine (ukupno 72 nedelje). Na bazi ocenjenih modela, u radu je formirana statička prognoza stope rasta broja obolelih od COVIDa-19 za dva različita horizonta predviđanja (6 i 17 nedelja).

Tema modeliranja i predviđanja kretanja broja obolelih od COVIDa-19 primenom ARIMA modela istražena je u nekoliko radova [1], [3], [5] i [6]. U njima su prilikom formiranja modela korišćeni dnevni podaci koji se odnose na sam početak pandemije, pri čemu nijedan od njih ne analizira podatke za Srbiju. Doprinos ovog rada ogleda se u tome što je prvi koji empirijski obrađuje ovu temu za našu zemlju.

Ostatak rad je strukturiran na sledeći način. Osnovni metodološki koncepti su predstavljeni u delu 2. Podaci i empirijski rezultati su izloženi u delu 3. Zaključni komentari su dati u delu 4.

<sup>1</sup> Projekat: *Macroeconomic implications of COVID-19 and effectiveness of policy response in Europe: Empirical evidence and econometric modelling* finansiran od strane Fonda za nauku Republike Srbije.

## 2. METODOLOGIJA

Prilikom opisivanja kretanja vremenske serije broja obolelih od COVIDa-19 koristimo autoregresione modele pokretnih proseka za integrisane vremenske serije. Ove modele zapisujemo na sledeći način: ARIMA(p,d,q), gde p predstavlja red autoregresione komponente, d nivo intergrisanosti vremenske serije i q red komponente pokretnih proseka. Oni se mogu predstaviti formom sledećeg oblika:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L)^d X_t = \theta_0 + (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) e_t \quad (1)$$

gde su  $\Phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)$  i  $\Theta(L) = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)$  polinomi koji opisuju redom autoregresionu komponentu i komponentu pokretnih proseka stacionarne vremenske serije  $(1 - L)^d X_t$ . Sa L je označen operator doznje, dok je  $e_t$  proces beli šum.

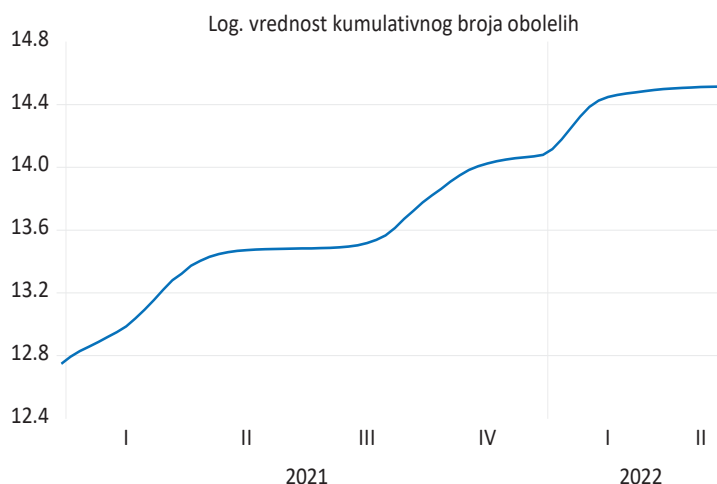
U radu je korišćena Boks-Dženkinsova strategija modeliranja kako bi se izvršio izbor odgovarajućeg ARIMA modela. Ona se sprovodi u tri faze. Prva podrazumeva identifikaciju modela u okviru koje se određuje stepen integrisanosti vremenske serije (d), kao i red autoregresione komponente i komponente pokretnih proseka modela (p i q). Prilikom utvrđivanja stepena integrisanosti vremenske serije korišćen je Diki-Fulerov test jediničnog korena (DF test) i KPSS test jediničnog korena. Druga faza podrazumeva ocenjivanje modela, dok se u trećoj fazi proverava adekvatnost modela.

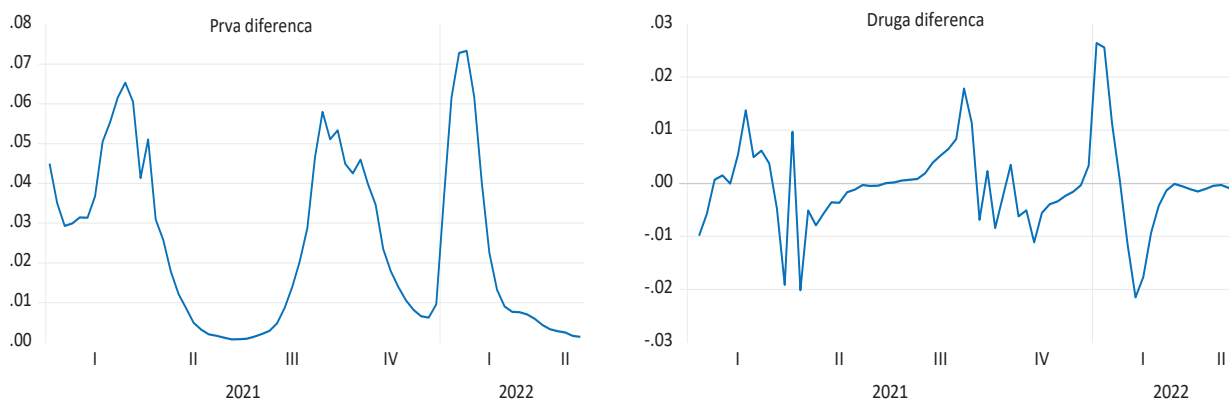
Jedan od glavnih ciljeva formiranja ovih modela jeste njihova upotreba prilikom predviđanja. U radu je korišćena statička prognoza koja se može primeniti samo u okviru perioda koji je obuhvaćen uzorkom, odnosno u okviru perioda za koji su nam dostupni stvarni podaci analizirane vremenske serije. Prilikom poređenja preciznosti prognoze konkurentnih ARIMA modela koriste se sledeća tri pokazatelja: koren srednje kvadratne greške prognoze, srednja apsolutna greška prognoze i srednja apsolutna procentualna greška prognoze.

## 3. PODACI I EMPIRIJSKI REZULTATI

Podaci o dnevnom broju obolelih za Srbiju preuzeti su sa sajta *Our World in Data*. Na osnovu njih formirani su kumulativni podaci na nedeljnom nivou za period 28.12.2020 – 9.5.2022. godine. Prilikom analize korišćene su logaritmovane vrednosti. Svi rezultati su dobijeni primenom programskog paketa *EViews 11*.

Na narednom grafiku predstavljeno je kretanje logaritmovane vrednosti kumulativnog broja obolelih u posmatranom periodu (polazna serija), kao i kretanje prve i druge difference polazne serije. Možemo primetiti da polaznu seriju karakteriše rastući trend nestacionarnog tipa koji je verovatno prisutan i u prvoj diferencijalnoj seriji. Druga diferencijalna serija oscilira oko nulte srednje vrednosti.





**Grafik 1:** Polazna serija, prva i druga diferencija polazne serije

Prilikom utvrđivanja nivoa integrisanosti polazne vremenske serije analiziramo kretanje obične i parcijalne autokorelacione funkcije. Ocene njihovih koeficijenata date su u Tabeli 1. Možemo primetiti da ocene običnih autokorelacionih koeficijenata sporo opadaju od vrednosti koja je blizu jedan (0.958), dok je kod parcijalne autokorelacione funkcije značajan samo koeficijent na prvoj doznji. Prethodno izloženi rezultati sugerišu prisustvo jediničnog korena u polaznoj seriji.

**Tabela 1:** Obična i parcijalna autokorelaciona funkcija

Doznje	1	2	3	4	5	6	7	8
AC	0.958	0.914	0.869	0.822	0.774	0.726	0.677	0.628
PAC	0.958	-0.039	-0.042	-0.042	-0.038	-0.034	-0.032	-0.026

Napomena: Autokorelacioni koeficijenti su statistički značajni na nivou značajnosti 5% ukoliko se nalaze van granica intervala ( $\pm 0.23$ ).

U tabelama 2 i 3 su redom dati rezultati primene A(DF) testa jediničnog korena i KPSS testa jediničnog korena. Rezultati primene A(DF) sugerišu postojanje dva jedinična korena u polaznoj seriji, dok rezultati KPSS testa daju isti zaključak na nivou značajnosti 5%. Na osnovu izloženih rezultata možemo zaključiti da je polazna serija integrisana drugog reda. U daljem modeliranju koristićemo transformaciju druge diference.

**Tabela 2:** Rezultati primene A(DF) testa jediničnog korena

$H_0: X_t \sim I(1)$	$H_0: X_t \sim I(2)$	$H_0: X_t \sim I(3)$
$H_1: X_t \sim I(0)$	$H_1: X_t \sim I(1)$	$H_1: X_t \sim I(2)$
ADF(2) = -3.023	ADF(1) = -3.09	DF = -4.61

Napomena: Sa  $X_t$  je označena polazna serija. Kritične vrednosti u modelu sa konstantom i trendom na nivou značajnosti 1%, 5% i 10% su redom: -4.1, -3.48 i -3.17.

**Tabela 3:** Rezultati primene KPSS testa jediničnog korena

$H_0: X_t \sim I(0)$	$H_0: X_t \sim I(1)$	$H_0: X_t \sim I(2)$
$H_1: X_t \sim I(1)$	$H_1: X_t \sim I(2)$	$H_1: X_t \sim I(3)$
KPSS = 0.162	KPSS = 0.19	KPSS = 0.10

Napomena: Sa  $X_t$  je označena polazna serija. Kritične vrednosti u modelu sa konstantom i trendom na nivou značajnosti 1%, 5% i 10% su redom: 0.216, 0.146 i 0.119. Broj doznji je odabran na osnovu rezultata primene ADF testa jediničnog korena.

Na osnovu analize obične i parcijalne autokorelacione funkcije druge diference polazne serije ocenjena su dva modela: ARIMA(1,2,1) model koji sadrži pet veštačkih promenljivih (u nastavku rada Model 1) i redukovani ARIMA(4,2,0) model sa tri veštačke promenljive (u nastavku rada Model 2). U tabelama 4 i 5 redom su prikazani ocenjeni modeli. Možemo primetiti da Model 1 ima nižu vrednost standardne greške regresije (S), kao i vrednost Švarcovog informacionog kriterijuma (SC). Rezultati primene Žark-Bera testa sugerišu da su reziduali normalno raspodeljeni kod oba ocenjena modela (data vrednost JB test-statistike i odgovarajuća p-vrednost). Takođe, reziduali su i neautokorelisani (data vrednost Boks-Ljungove test-statistike i odgovarajuća p-vrednost). Oba ocenjena modela dobro opisuju dinamiku druge diference polazne serije (Grafik 2).

**Tabela 4:** Ocena jednačine druge diference polazne serije (Model 1)

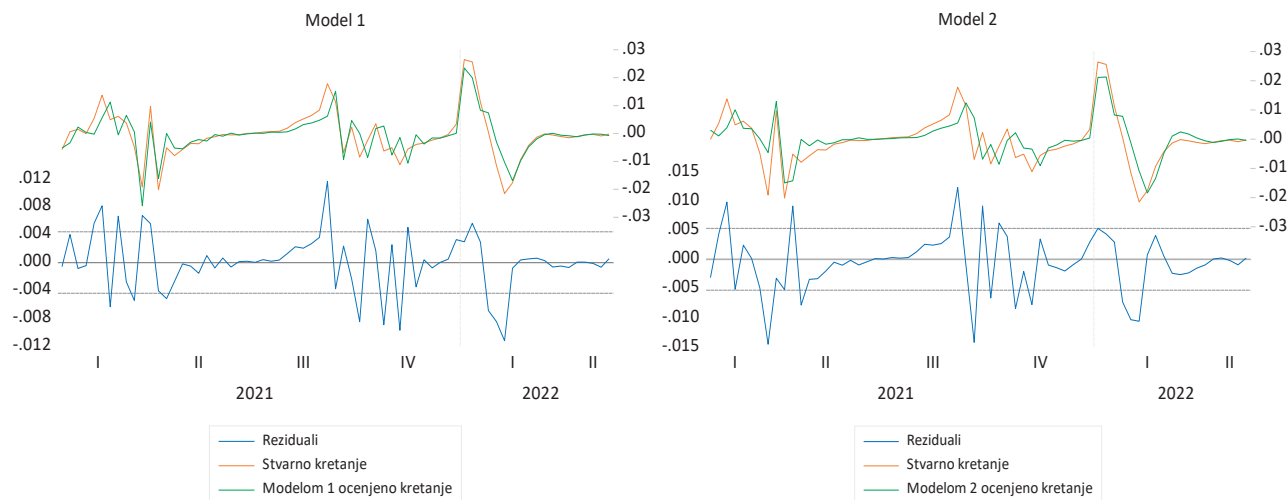
Promenljiva	Ocena	t-statistika
AR(1)	0.528	4.025
MA(1)	0.492	3.413
V <sub>1</sub>	0.020	4.909
V <sub>2</sub>	-0.021	-6.679
V <sub>3</sub>	-0.024	-7.756
V <sub>4</sub>	-0.013	-4.426
V <sub>5</sub>	0.015	3.707
S=0.0044 SC=-7.692 Q(7)=4.77 (0.44) JB=2.97 (0.23)		
Ocene korena karakteristične jednačine: AR (0.53) i MA (-0.49).		

Napomena: Prva veštačka promenljiva uzima vrednost 1 u prvoj nedelji januara 22. godine, 0 u ostalim; Druga uzima vrednost 1 u poslednjoj nedelji marta 21. godine, 0 u ostalim; Treća uzima vrednost 1 u drugoj nedelji aprila 2021. godine, 0 u ostalim; Četvrta uzima vrednost 1 u trećoj nedelji septembra 21. godine, 0 u ostalim; Peta uzima vrednost jedan u drugoj nedelji januara 22. godine, 0 u ostalim.

**Tabela 5:** Ocena jednačine druge diference polazne serije (Model 2)

Promenljiva	Ocena	t-statistika
AR(1)	0.754	8.829
AR(4)	-0.198	-2.363
V <sub>1</sub>	0.018	3.838
V <sub>2</sub>	0.029	6.832
V <sub>3</sub>	0.015	3.113
S=0.0055 SC=-7.398 Q(7)= 1.89(0.86) JB=3.58(0.17)		
Ocene korena karakteristične jednačine: AR (0.71±0.38i; -0.33±0.44i)		

Napomena: Prva veštačka promenljiva uzima vrednost 1 u prvoj nedelji januara 22. godine, 0 u ostalim; Druga uzima vrednost 1 u prvoj nedelji aprila 21. godine, 0 u ostalim; Treća uzima vrednost jedan u drugoj nedelji januara 22. godine, 0 u ostalim.

**Grafik 2:** Stvarne i modelom ocenjene vrednosti druge diference polazne serije

Na bazi ocenjenih modela formiramo statičku prognozu za prvu diferencu polazne serije (stopa rasta broja obolelih). Modeli su ocenjeni zaključno sa drugom nedeljom januara 2022. godine, a prognoziramo vrednosti za period 17.1.2022. – 9.5.2022. godine (17 nedelja). U Tabeli 6 su dati pokazatelji preciznosti prognoze za oba modela. Možemo primetiti da prvi model ima niže vrednosti sva tri pokazatelja, što ukazuje na to da ima precizniju prognozu.



Ukoliko skratimo horizont predviđanja na 6 nedelja, odnosno modele ocenimo zaključno sa poslednjom nedeljom marta 2022. godine dok se prognoziraju vrednosti za period 4.4.2022 – 9.5.2022. godine, prvi model takođe ima niže vrednosti sva tri pokazatelja preciznosti prognoze (Tabela 7). Dobijeni rezultat potvrđuje prethodni zaključak da prvi ocenjeni model, ARIMA(1,2,1) sa pet veštačkih promenljivih, ima precizniju prognozu.

**Tabela 7:** Pokazatelji preciznosti prognoze (horizont predviđanja 6 nedelja)

	<b>Model 1</b>	<b>Model 2</b>
Koren srednje kvadratne greške prognoze	0.000463	0.000575
Srednja apsolutna greška prognoze	0.000366	0.000438
Srednja apsolutna procentualna greška prognoze	17.65837	18.97891

#### 4. ZAKLJUČAK

U radu smo istraživali primenu ARIMA modela prilikom opisivanja dinamike broja obolelih od virusa COVID-19 u Srbiji. Na osnovu kumulativnih podataka o broju obolelih na nedeljnom nivou za period 28.12.2020 – 9.5.2022. godine ocenjena su dva ARIMA modela sa zadovoljavajućim statističkim svojstima: ARIMA(1,2,1) model i redukovani ARIMA(4,2,0) model. Dobijeni rezultati o pokazateljima preciznosti prognoze, za dva različita horizonta predviđanja (6 i 17 nedelja), ukazuju na to da ocenjeni ARIMA(1,2,1) model ima precizniju prognozu stope rasta broja obolelih od virusa COVID-19 u Srbiji.

#### LITERATURA

- [1] Alzahrani, S. I., Aljamaan, I. A., & Al-Fakih, E. A. (2020). Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *Journal of infection and public health*, 13(7), 914-919.
- [2] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- [3] Ceylan, Z. (2020). Estimation of COVID-19 prevalence in Italy, Spain, and France. *Science of The Total Environment*, 729, 138817.
- [4] Mladenović, Z., & Nojković, A. (2021). *Primenjena analiza vremenskih serija*. Centar za izdavačku delatnost Ekonomskog fakulteta.
- [5] Sahai, A. K., Rath, N., Sood, V., & Singh, M. P. (2020). ARIMA modelling & forecasting of COVID-19 in top five affected countries. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(5), 1419-1427.
- [6] Singh, R. K., Rani, M., Bhagavathula, A. S., Sah, R., Rodriguez-Morales, A. J., Kalita, H., ... & Kumar, P. (2020). Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model. *JMIR public health and surveillance*, 6(2), e19115.

## ANALIZA VARIJABILITETA PRIHODA U ZDRAVSTVENIM USTANOVAMA - NA PRIMERU PRIVATNIH BOLNICA U BEOGRADU

### THE REVENUE VARIABILITY ANALYSIS IN THE HEALTH INSTITUTIONS - ON THE EXAMPLE OF THE PRIVATE HOSPITALS IN BELGRADE

VESNA RAJIĆ<sup>1</sup>, TATJANA RAKONJAC-ANTIĆ<sup>2</sup>, IVANA IVKOVIĆ<sup>3</sup>, JASNA BABIĆ<sup>4</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, [vesna.rajic@ekof.bg.ac.rs](mailto:vesna.rajic@ekof.bg.ac.rs)

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, [tatjana.rakonjac@ekof.bg.ac.rs](mailto:tatjana.rakonjac@ekof.bg.ac.rs)

<sup>3</sup> Univerzitet u Beogradu – Ekonomski fakultet, [ivana.ivkovic@ekof.bg.ac.rs](mailto:ivana.ivkovic@ekof.bg.ac.rs)

<sup>4</sup> Univerzitet u Beogradu – Ekonomski fakultet, [jasna.babic@ekof.bg.ac.rs](mailto:jasna.babic@ekof.bg.ac.rs)

**Rezime:** Bolnička zdravstvena delatnost predstavlja sastavni deo sekundarnog nivoa zdravstvene zaštite. Donosioci odluka poređenjem prihoda i rashoda dobijaju informacije o uspešnosti poslovanja bolnica. Analiza varijabiliteta prihoda im daje početnu osnovu za optimalno upravljanje prihodima. U radu je analiziran varijabilitet prihoda na primeru 44 privatne bolnice u Beogradu. Korišćene su različite intervalne ocene koeficijenta varijacije kao relativne mere varijabiliteta. Rezultati empirijske analize su pokazali da prilikom intervalnog ocenjivanja ove mere, prednost treba dati percentil intervalima baziranim na bootstrap rangiranom uzorkovanju po redovima. Intervalne ocene dobijene ovom metodom su najuže u poređenju sa ostalim analiziranim intervalnim ocenama.

**Ključne reči:** zdravstvene ustanove, privatne bolnice, prihodi, koeficijent varijacije, intervali poverenja

**Abstract:** The hospital health activity is an integral part of the secondary level of the health care. Decision makers get the information on the business performance of the hospital comparing the revenues and the expenditures. The revenue variability analysis provides an initial basis for an optimal revenue management. The variability of the revenue on the example of 44 private hospitals in Belgrade is analyzed. Different interval estimates of the coefficient of variation as the relative measure of the variability were used. The results of the empirical analysis which refers to the interval estimation of this measure showed that advantage should be given to the percentile intervals based on the bootstrap ranked set sampling by rows. The interval estimates obtained using this method are the narrowest in comparison with the other analyzed interval estimates.

**Keywords:** health institutions, private hospitals, revenues, coefficient of variation, confidence intervals

## 1. UVOD

Zdravstvena zaštita predstavlja organizovanu delatnost koja ima za cilj očuvanje i unapređenje zdravlja stanovništva [11]. U okviru zdravstvenih ustanova se obavlja zdravstvena zaštita na primarnom, sekundarnom i tercijarnom nivou. Sekundarni nivo zdravstvene zaštite obuhvata specijalističko konsultativnu i bolničku zdravstvenu delatnost koja se sprovodi u okviru opštih i specijalizovanih bolnica koje mogu biti u javnoj i privatnoj svojini [12]. Finansijski izveštaji pružaju informacije o rashodima, prihodima i poslovnom rezultatu. Osnovni izvori prihoda bolnica u privatnom vlasništvu su redovni prihodi od pružanja medicinskih usluga. Prilikom poređenja prihoda i rashoda upravljači dobijaju informacije o uspešnosti poslovanja na osnovu kojih mogu da preduzimaju neophodne akcije sa ciljem jačanja poslovnog sistema.

Na osnovu analize varijabiliteta prihoda stvaraju se preduslovi za adekvatno upravljanje prihodima koje predstavlja osnovni metod za usaglašavanje ponude i tražnje za bolničkim uslugama. Ovo usaglašavanje je uslovljeno i brojem i strukturom zahteva potencijalnih korisnika i bolničkim kapacitetima.

Predmet rada je analiza varijabiliteta prihoda u zdravstvenim ustanovama na primeru 44 privatne bolnice u Beogradu. Cilj rada je da se na osnovu različitih intervalnih ocena koeficijenta varijacije ukaže na značaj ove mere prilikom analize varijabiliteta prihoda.

Koeficijent varijacije je relativna mera disperzije na osnovu koje se utvrđuje koliko u proseku pojedinačne vrednosti promenljive odstupaju od njene prosečne vrednosti izraženo u procentima prosečne vrednosti. Ona omogućava sagledavanje mogućeg rizika u poređenju sa očekivanom vrednošću investicija i pruža donosiocima odluka relevantne informacije za upravljanje prihodima u privatnim bolnicama.

Ukoliko osnovni skup sledi normalnu raspodelu, prilikom konstrukcije intervala poverenja za koeficijent varijacije osnovnog skupa koristi se metoda bazirana na  $t$ -statistici. Međutim, ako je pretpostavka o normalnosti narušena, moraju se razmatrati druge metode.

## 2. POSTOJEĆI INTERVALI POVERENJA

U ovom delu pažnja će biti posvećena različitim intervalnim ocenama koeficijenta varijacije koje su često predmet razmatranja u literaturi (videti [7, 8]). Koeficijent varijacije osnovnog skupa će u nastavku biti označen simbolom  $CV$ , dok će njegova ocena biti obeležena sa  $CV$ .

### Interval poverenja baziran na $t$ statistici

Statistika  $t = \frac{CV - CV}{S_{CV}}$  ima Student-ovu raspodelu sa  $n-1$  stepeni slobode, gde  $S_{CV}$  predstavlja ocenu standardne greške statistike  $CV$ . Interval poverenja za koeficijent varijacije određen sa nivoom pouzdanosti  $(1 - \alpha)100\%$  je sledećeg oblika (videti [1]):

$$I_{CV} = \left[ CV - t_{1-\alpha/2, n-1} S_{CV}, CV + t_{1-\alpha/2, n-1} S_{CV} \right], \quad (1)$$

gde je  $t_{1-\alpha/2}$  percentil  $t$  raspodele sa  $n-1$  stepeni slobode.

### Miller-ov interval poverenja

Miller [9] je utvrdio da raspodela statistike  $CV$  pod određenim uslovima sledi približno normalnu raspodelu. Interval poverenja za koeficijent varijacije određen sa nivoom pouzdanosti  $(1 - \alpha)100\%$  je (videti [9]):

$$I_{CV} = \left[ CV - Z_{1-\alpha/2} \sqrt{\frac{1}{(n-1)} CV^2 (0,5 + CV^2)}, CV + Z_{1-\alpha/2} \sqrt{\frac{1}{(n-1)} CV^2 (0,5 + CV^2)} \right], \quad (2)$$

gde je  $Z_{1-\alpha/2}$  percentil standardizovane normalne raspodele.

### Curto-ov i Pinto-ov interval poverenja

Curto-ov i Pinto-ov  $(1 - \alpha)100\%$  interval poverenja za koeficijent varijacije koji će u nastavku biti označen sa C&P, definiše se na sledeći način (videti [3]):

$$I_{CV} = \left[ CV - Z_{1-\alpha/2} \sqrt{\frac{1}{n} (CV^4 + 0,5 CV^2)}, CV + Z_{1-\alpha/2} \sqrt{\frac{1}{n} (CV^4 + 0,5 CV^2)} \right], \quad (3)$$

gde je  $Z_{1-\alpha/2}$  percentil standardizovane normalne raspodele.

## 3. INTERVALI POVERENJA BAZIRANI NA METODAMA PONOVLJENIH UZORAKA

*Bootstrap* i *jackknife* predstvljaju najčešće korišćene metode ponovljenih uzoraka. Za konstrukciju *bootstrap* intervalnih ocena, potrebno je iz originalnog uzorka izabrati  $B$  *bootstrap* uzoraka (odnosno uzoraka sa ponavljanjem). Dalje, u svakom od njih se izračunava vrednost statistike  $CV$ , označena simbolom  $CV^*$ .

### *Bootstrap-t* interval poverenja

U svakom od  $B$  uzoraka sa ponavljanjem određuje se vrednost statistike  $T^* = \frac{CV^* - CV}{S_{CV^*}}$  ( $S_{CV^*}$  je ocena standardne greške statistike  $CV^*$ ). Ako je nivo pouzdanosti  $(1 - \alpha)100\%$ , onda je *bootstrap-t* interval poverenja za koeficijent varijacije oblika:



$$I_{CV} = \left[ CV - \hat{t}_{1-\alpha/2} S_{CV}, CV - \hat{t}_{\alpha/2} S_{CV} \right], \quad (4)$$

gde su  $\hat{t}_{1-\alpha/2}$  i  $\hat{t}_{\alpha/2}$  percentili odgovarajuće *bootstrap* raspodele.

### Percentil interval poverenja

Ovaj interval se konstruiše korišćenjem percentila *bootstrap* raspodele razmatrane statistike. Percentil interval poverenja za koeficijent varijacije određen sa nivoom pouzdanosti  $(1 - \alpha)100\%$  definisan je sledećom formulom (videti [4]):

$$I_{CV} = \left[ CV_{\alpha/2}^*, CV_{1-\alpha/2}^* \right], \quad (5)$$

gde su  $CV_{\alpha/2}^*$  i  $CV_{1-\alpha/2}^*$  percentili *bootstrap* raspodele.

### BCa interval poverenja

Neka je  $(1 - \alpha)100\%$  odgovarajući nivo pouzdanosti. BCa (skraćeno od engl. bias corrected and accelerated) interval poverenja za koeficijent varijacije je (videti [4]):

$$I_{CV} = \left[ CV_{\alpha_1}^*, CV_{\alpha_2}^* \right], \quad (6)$$

gde su vrednosti  $\alpha_1$  i  $\alpha_2$  izračunate pomoću:

- konstante  $\hat{a}$  koja se naziva akceleracija;
- korekcije pristrasnosti  $\hat{z}_0$ ;
- percentila standardizovane normalne raspodele ( $Z_{\alpha/2}$  i  $Z_{1-\alpha/2}$ ).

Više o ovim formulama pogledati u [4].

### Jackknife interval poverenja

Neka  $(X_1, \dots, X_n)$  predstavlja slučajan uzorak. Uzorci veličine  $n-1$  izabrani iz originalnog uzorka  $(X_1, \dots, X_n)$  predstavljaju *jackknife* uzorke. Kako bi se formirao  $i$ -ti *jackknife* uzorak, izostavlja se element  $X_i$ , odnosno *jackknife* uzorak je oblika  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$ .

U cilju konstruisanja *jackknife* intervala poverenja potrebno je generisati  $n$  *jackknife* uzoraka iz originalnog uzorka. Dalje, u svakom *jackknife* uzorku računa se vrednost statistike  $CV$ , koja se obeležava simbolom  $CV_{(i)}$ . Na kraju, definisana je vrednost  $\tilde{CV}_i$  koja se određuje kao  $nCV - (n-1) CV_{(i)}$ . Na osnovu formula za prosek i *jackknife* standardnu grešku:

$$\tilde{CV} = \sum_{i=1}^n \tilde{CV}_i / n, \quad (7)$$

$$S_{jack} = \left\{ \sum_{i=1}^n \left( \tilde{CV}_i - \tilde{CV} \right)^2 / [(n-1)n] \right\}^{1/2}, \quad (8)$$

može se konstruisati  $(1 - \alpha)100\%$  interval poverenja za koeficijent varijacije (videti [4]):

$$I_{CV} = \left[ \tilde{CV} - t_{1-\alpha/2, n-1} S_{jack}, \tilde{CV} + t_{1-\alpha/2, n-1} S_{jack} \right], \quad (9)$$

gde je  $t_{1-\alpha/2}$  odgovarajući percentil  $t$  raspodele sa  $n-1$  stepeni slobode.

#### 4. INTERVALI POVERENJA BAZIRANI NA METODAMA PONOVLJENIH UZORAKA, RANGIRANOM UZROKOVANJU I TRANSFORMACIJE INTERVALA POVERENJA

Procedura balansiranog rangiranog uzorkovanja se sastoji u sledećem (videti [5]). Bira se  $k^2$  jedinica i na osnovu njih se formira  $k$  uzoraka koji se sastoje od  $k$  jedinica. U svakom od  $k$  uzoraka potrebno je rangirati jedinice. Iz  $k$ -tog uzorka se bira jedinica sa  $k$ -tim rangom. Dakle, raspolaže se uzorkom koji sadrži  $k$  jedinica  $X_{(1)}, X_{(2)}, \dots, X_{(k)}$ . Kako bi se generisao balansirani rangirani uzorak veličine  $n$ , navedeni postupak se ponavlja  $m$  puta.

U radu će biti analizirani Miller-ov i Curto-ov i Pinto-ov interval poverenja bazirani na proceduri rangiranog uzorkovanja (oni će biti obeleženi sa RSS Miller-ov i RSS C&P interval).

##### Intervali poverenja bazirani na metodama ponovljenih uzoraka u kombinaciji sa procedurom rangiranog uzorkovanja

Procedura *bootstrap* rangirano uzorkovanje po redovima se sprovodi tako što se iz  $k$  nezavisnih slučajnih uzoraka veličine  $m(X_{(1)1}, \dots, X_{(1)m}), \dots, (X_{(k)1}, \dots, X_{(k)m})$  generiše  $k$  *bootstrap* uzoraka koji se sastoje od  $m$  elemenata (videti [10]).

U nastavku će se analizirati percentil i BCa intervali poverenja koji su bazirani na ovoj proceduri (oni će biti označeni sa RSS percentil i RSS BCa).

##### Intervali poverenja bazirani na parcijalnom rangiranom uzorkovanju

Procedura izbora parcijalnog rangiranog uzorka koji se sastoji od  $k$  elemenata vrši se na sledeći način (videti [6]). Na osnovu formule  $r = \alpha k, 0 \leq \alpha < 0,5$  se određuje veličina prostog slučajnog uzorka koji će zajedno sa rangiranim uzorkom formirati parcijalni rangirani uzorak. Zatim se iz osnovnog skupa bira  $2r$  prostih slučajnih uzoraka sa jednim elementom. Dalje, vrši se izbor  $k - 2r$  uzoraka sa  $k$  elemenata. U svakom od njih se vrši rangiranje jedinica i iz  $i$ -tog uzorka (za  $i = r+1, \dots, k-r$ ) se vrši izbor jedinice sa  $i$ -tim rangom. Ako je potrebno konstruisati parcijalni rangirani uzorak veličine  $n$ , navedeni postupak je potrebno ponoviti  $m$  puta.

U radu će se analizirati Miller-ov i Curto-ov i Pinto-ov interval poverenja koji su bazirani na opisanoj proceduri (oni će biti obeleženi sa PRSS Miller-ov i PRSS C&P interval).

##### Transformacije intervala poverenja

U ovom delu će biti reči o transformacijama baziranim na odsečenoj sredini primenjujući prosto slučajno uzorkovanje i rangirano uzorkovanje, kao i o *bootstrap* transformaciji postojećih intervala poverenja (videti [7, 8]).

Prilikom transformacija postojećih intervala poverenja baziranih na odsečenoj sredini, varijansa uzorka se određuje na osnovu odsečene sredine. Prvo se koristi odsečena sredina koju je uveo Bonett, a kojom je moguće odstraniti procenat vrednosti određen po formuli  $1/\left[2(n-4)^{1/2}\right]$  (videti [2]). Zatim se primenjuje odsečena sredina kojom se eliminiše 10% vrednosti.

Za konstruisanje Miller-ovog i Curto-ovog i Pinto-ovog intervala koristeći Bonett-ovu odsečenu sredinu primenjuju se formule (2) i (3), gde  $\bar{X}$  predstavlja Bonett-ovu odsečenu sredinu uzorka (u nastavku će navedeni intervali biti obeleženi sa Miller-ov trim i C&P trim interval). Na sličan način će biti transformisani Miller-ov i Curto-ov i Pinto-ov interval na osnovu odsečene sredine koja eliminiše 10% vrednosti (u nastavku će navedeni intervali biti obeleženi sa Miller-ov trim 10% i C&P trim 10% interval).

Dalje, u radu će biti analizirani Miller-ov i Curto-ov i Pinto-ov interval poverenja bazirani na proceduri rangiranog uzorkovanja uz primenu odsečene sredine (u nastavku će ovi intervali biti obeleženi sa RSS Miller-ov trim, RSS Miller-ov trim 10%, RSS C&P trim i RSS C&P trim 10% interval).

Na kraju, biće analizirani Miller-ov *bootstrap-t* i Curto-ov i Pinto-ov *bootstrap-t* interval poverenja na osnovu percentila koji se primenjuju u slučaju formiranja *bootstrap-t* intervala.

#### 5. EMPIRIJSKA ANALIZA

U svrhe empirijske analize, posmatrane su 44 privatne bolnice na teritoriji grada Beograda. Promenljiva čiji varijabilitet je analiziran je poslovni prihod u 2020. godini. Podaci su dobijeni iz dostupnih finansijskih

izveštaja na sajtu Agencije za privredne registre. U Tabeli 1 nalaze se razmatrani intervali poverenja za koeficijent varijacije. Svi rezultati dobijeni su korišćenjem programa *R*.

Posmatrajući Tabelu 1, zaključuje se da je najuži interval RSS Percentil, odnosno percentil baziran na *bootstrap* rangiranom uzorkovanju po redovima. Ovom intervalu treba dati prednost prilikom ocenjivanja koeficijenta varijacije.

**Tabela 1:** Intervali poverenja za koeficijent varijacije

Interval	Granice 95% intervala
<i>t</i>	(1,329; 2,057)
Percentil	(1,214; 2,053)
RSS Percentil	(1,232; 1,807)
<i>Bootstrap-t</i>	(1,405; 2,445)
BCa	(1,362; 2,320)
RSS BCa	(1,454; 2,353)
<i>Jackknife</i>	(1,868; 2,744)
Miller-ov	(1,338; 2,945)
Miller-ov trim	(0,764; 2,698)
Miller-ov trim 10%	(0,763; 2,725)
RSS Miller-ov	(1,453; 3,021)
RSS Miller-ov trim	(0,712; 1,774)
RSS Miller-ov trim 10%	(0,765; 2,603)
Miller-ov <i>bootstrap-t</i>	(0,928; 3,715)
PRSS Miller-ov	(0,751; 3,011)
C&P	(0,775; 2,611)
C&P trim	(0,775; 2,687)
C&P trim 10%	(0,774; 2,714)
RSS C&P	(0,772; 2,388)
RSS C&P trim	(0,775; 2,616)
RSS C&P trim 10%	(0,771; 2,841)
C&P <i>bootstrap-t</i>	(0,955; 3,504)
PRSS C&P	(0,769; 2,321)

Izvor: Izlaz iz programskog paketa *R*

Zaključuje se da se prosečno odstupanje prihoda od proseka nalazi u intervalu od 1,232 do 1,807 sredine, odnosno da se mogući rizik nalazi u intervalu od 1,232 do 1,807 očekivane vrednosti.

## 6. ZAKLJUČAK

U radu je analiziran varijabilitet prihoda na primeru privatnih bolnica u Beogradu. Važna mera varijabiliteta je koeficijent varijacije jer pruža donosiocima odluka relevantne informacije za upravljanje приходima u bolnicama. Korišćeni su različiti intervali poverenja za koeficijent varijacije osnovnog skupa kao relativne mere disperzije: postojeći intervali poverenja (*t* interval, Miller-ov interval, Curto-ov i Pinto-ov interval), intervali poverenja bazirani na metodama ponovljenih uzoraka (percentil interval, *bootstrap-t* interval, BCa interval i *jackknife* interval), intervali poverenja bazirani na procedurama rangiranog uzorkovanja i parcijalnog rangiranog uzorkovanja, kao i transformacije postojećih intervala poverenja korišćenjem odsečene sredine. Ustanovljeno je da je prilikom analize varijabiliteta prihoda zdravstvenih ustanova poželjno koristiti percentil intervale bazirane na *bootstrap* rangiranom uzorkovanju po redovima. Ovi intervali obezbeđuju najuže ocene u poređenju sa ostalim razmatranim intervalima.

## LITERATURA

- [1] Abdi, H. (2010). Coefficient of Variation. In: *Encyclopedia of Research Design*, Salkind, N. J. (ed.), Thousand Oaks, CA, Sage, 169-171.
- [2] Bonett, D. G. (2006). Approximate confidence interval for standard deviation of nonnormal distributions. *Computational Statistics and Data Analysis*, 50(3), 775-782.
- [3] Curto, J. D., & Pinto, J. C. (2009). The coefficient of variation asymptotic distribution in the case of non-iid random variables. *Journal of Applied Statistics*, 36(1), 21-32.
- [4] Efron, B., & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. New York: Chapman&Hall, ISBN: 0-412-04231-2.

- [5] Ghosh, S., Chatterjee, A., & Balakrishnan, N. (2017). Nonparametric confidence intervals for ranked set samples. *Computational Statistics*, 32(4), 1689-1725.
- [6] Haq, A., Brown, J., Moltchanova, E., & Al-Omari, A.-I. (2013). Partial Ranked Set Sampling Design. *Environmetrics*, 24(3), 201-207.
- [7] Ivković, I. (2021). Neparametarske statističke tehnike ocenjivanja regresionih koeficijenata i koeficijenta varijacije u korporativnim finansijama. Ekonomski fakultet, Univerzitet u Beogradu.
- [8] Ivković, I., Rajić, V. (2021). Better confidence intervals for the population coefficient of variation. *Communications in Statistics – Simulation and Computation* 50 (12), 4215-4262.
- [9] Miller, E. G. (1991). Asymptotic test statistics for coefficient of variation. *Communication in Statistics – Theory and Methods*, 20(10), 3351-3363.
- [10] Modarres, R., Hui, T. P., & Zheng, G. (2006). Resampling methods for ranked set samples. *Computational Statistics and Data Analysis*, 51(2), 1039-1050.
- [11] Rakonjac-Antić T. (2018). *Penzijsko i zdravstveno osiguranje*. Beograd: Centar za izdavačku delatnost, Ekonomski fakultet, Univerzitet u Beogradu, ISBN: 978-86-403-1543-2.
- [12] Zakon o zdravstvenom osiguranju, Sl. glasnik RS, br. 25/2019.

## **BENFORD'S LAW APPLICATION ON ATTRACTING INVESTMENT AND NEW EMPLOYMENT IN THE REPUBLIC OF SERBIA DURING THE PERIOD OF 2016-2022**

JELENA STANOJEVIĆ<sup>1</sup>, MILENA LUTOVAC DJAKOVIĆ<sup>2</sup>, ŽELJKO JOVIĆ<sup>3</sup>

<sup>1</sup> University of Belgrade – Faculty of Economics and Business, jelena.stanojevic@ekof.bg.ac.rs

<sup>2</sup> University of Belgrade – Faculty of Economics and Business, milena.lutovac@ekof.bg.ac.rs

<sup>3</sup> University of Belgrade – Faculty of Economics and Business, zeljko.jovic@ekof.bg.ac.rs

**Abstract:** *In this paper we use Benford's law to detect whether considered reports contain fraud or irregularity in terms of some errors. In a line with our intuition, the leading digits in data follow uniform distribution, but that is not true. The Benford's law claims something else. Namely, by the law leading digits 1-9 from all large set of data appears with decreasing logarithmic distribution. A frequency of a digit 1 is the largest (30%), and the digit 9 appears with the smallest frequency (4.58%). In order to detect does reports contain some irregularity, we consider incentives for attracting investment and new employment in the Republic of Serbia during the period of 2016-2022.*

**Keywords:** *Benford's law, analysis of irregularity, Z and K-S tests, state incentives*

### **1. INTRODUCTION**

Benford's law first time appears in 1881. American astronomer Simon Newcomb (1835.-1909.) proposed the law, [15]. He noticed in logarithmic tables that the first pages were used more than the others. The consequence of that was the fact that the leading digit one appears more often than in the uniform distribution. After 60 years, American electrical engineer and physicist Frank Albert Benford (1883.-1948.) gave the proof of the law by using a large data sets and their analysis. He collected more than 20,000 observations from difference data sets. After first steps, Benford's law found its application in the area of accounting, auditing and taxation, in a purpose to detect of fraud or some errors in the financial report. That application was suggested by Nigrini, at 1992. He find out that when humans make up the fraud the digits are not in according with Benford's law. The probability frame of the law was given by Hill and in 1995 he gave empirical proof of the law.

In this paper Benford's law was applied to detect irregularity and possible errors in reports of incentives for attracting investment and new employment in the Republic of Serbia during the period of 2016-2022. Fitting with the law was considered with Z test and K-S (Kolmogorov-Smirnov) test for the first place in a number. The purpose of this paper is to suggest one of possible applications of Benford's law as a digital analysis of the financial reports and their validity.

The rest of the paper is organized as follows. Section 2 presents the relevant literature, describes the theoretical background of Benford's law and some limitation of application. In section 3 we describe data and the tools we use in our empirical analysis. Section 4 contain our main results and Section 5 concludes.

### **2. LITERATURE REVIEW AND BACKGROUND OF BENFORD'S LAW**

#### **2.1. Literature review**

First application of Benford's law in economic forecast was introduced by Varian in 1972. Application which detects errors in macroeconomic data was given in [14] and application in international and governmental macroeconomic statistics is given in the papers [20], [10] and [8].

Further, one modification of Benford's law was used in the purpose to show that some firms in New Zeland did not conform to the law, [3]. Similar results showed in [23] for US firms.

In the area of accounting and tax settings main results Nigrini and his coauthors proposed most significant results. For example, [16] this paper analyzed interest paid and interest received on individual tax returns. After Nigrini's initial paper, many researchers investigated application of Benford's law in the area of forensic accounting and in analysis of financial report, for example [17], [5], [13], [19], [22], [25], [21], [24] and many others. A guide for detection of errors in data was given in [18], [19] and [5].

Another application of the law was made by the authors in countries of Latin America, on the topic in tracking exchange rate management, [2]. The authors found out that the law holds when the euro is considered but does not hold when they considered the US dollar exchange rate.

Also Benford's law was used in the area of social networks in the paper [7] and application on prices in certain eBay auctions was given in the paper [6]. Further, it was used for testing psychological barriers in stock markets and connection with the law, was given in the paper [4], while application of the law on stock prices was made in [12].

In last two years, the most popular usage of the Benford's law was to detect whether the countries manipulate with COVID-19 data during the pandemics. Many papers were written on that topic, see for example [27] and [1].

This paper analyse Benford's law on incentives for attracting investment and new employment in Republic of Serbia during the period of 2016-2022. Namely, according to the Law on State Aid Control, it is defined as: "Any actual or potential public expenditure, or reduced realization of public revenue, by which the beneficiary of state aid acquires a more favorable position on the market compared to competitors, thereby distorting or endangering market competition", [28].

## 2.2. Background of Benford's law

Benford's law is one example of very complex mathematical tool with simple application. Since 1990 till recent time, the law was presented as unique method and process of detecting manipulation, fraud or some errors with the data, [26]. It is important fact that if there is not coincide digits with Benford's law, it is necessary to use deeper analysis and testing of data with different tests. Special remark is that positive and negative values have the same treatment in testing, and also a values with decimal numbers.

In the paper [19] is given the probability of an occurrence of the first digit, by the next formula:

$$P\{D_1 = k\} = \log\left(1 + \frac{1}{k}\right), k \in \{1, 2, \dots, 9\}, \quad (1)$$

where  $D_1$  represents the first position of a number and  $k$  is a digit on the first position.

Also in the same paper gives the probability of an occurrence of the second digit, by the next formula:

$$P\{D_2 = k\} = \sum_{j=1}^9 \log\left(1 + \frac{1}{10j+k}\right), k \in \{0, 1, \dots, 9\}, \quad (2)$$

where  $D_2$  represents the second position of a number and  $k$  is a digit on the second position

Further, the probabilities of appearance of any digit on the appropriate position are derived in the next equation, where higher-order digits appear with uniform distribution, with equal probability of 0.1:

$$P\{D_k = j_k\} = \sum_{j_1=1}^9 \sum_{j_2=0}^9 \dots \sum_{j_{k-1}=0}^9 \log\left(1 + \frac{1}{\sum_{i=1}^k 10^{k-i} j_i}\right), \quad (3)$$

$D_k$  represents  $k$ -th position of a number and  $j_k$  is a digit on that position,  $j_k \in \{0, 1, \dots, 9\}$ , [11].

The probabilities of appearance for the first five digits are presented in Table 1.

From Table 1 we may see the probabilities of an appearance of digits on the first position based on Benford's law. The digit one has the highest probability (30.10%), digit nine has the lowest probability (4.58%). And higher-order digits appear with uniform distribution with equal probability of 10%.

In some cases, Benford analysis is useful if the results comes from two distributions and it is not necessary that previous distributions complain with the Benford distribution ([5]), on large data sets (full year's transaction), accounts that appear to conform, etc. Also, the cases when it is not possible to apply Benford's law, are the ones when set of data compose of assigned numbers (for example, zip codes), when the digits are formed with some intention (for example, prices), when the digits are limited with minimal or maximal values (for example, height or weight), etc.

## 3. DATA AND METHODOLOGY

Data from reports of incentives for attracting investment and new employment in the Republic of Serbia during the period of 2016-2022 are the ones used in this paper. First digit of a numbers ware tested, to detect possible irregularity or errors in a data, with two tests: Z test and K-S test.

**Table 1:** The probability of an appearance of a digit in different places (percentage)

Digit	1st place	2nd place	3rd place	4th place	5th place
0	-	11.968	10.178	10.018	10.000
1	30.103	11.389	10.138	10.014	10.000
2	17.609	10.882	10.097	10.010	10.000
3	12.494	10.433	10.057	10.006	10.000
4	9.691	10.031	10.018	10.002	10.000
5	7.918	9.668	9.979	9.998	10.000
6	6.695	9.337	9.940	9.994	10.000
7	5.799	9.035	9.902	9.990	10.000
8	5.115	8.757	9.864	9.986	10.000
9	4.576	8.500	9.827	9.982	10.000

Source: [9]

First test which was used to check fitting with Benford’s distribution is the Z-test. For the appropriate digit Z-statistics with normal distribution was calculated, to obtain the statistical significance of the deviations in the expected and the observed proportion. The statistic is given by the next equation:

$$Z_k = \frac{|p_o - p_k| - \frac{1}{2n}}{\sqrt{\frac{p_k(1-p_k)}{n}}}, \tag{4}$$

where  $Z_k$  is Z-statistics for the digit  $k$  ( $k = 1, 2, \dots, 9$ ),  $p_k$  represent expected proportion of the digit  $k$  based on Benford’s law,  $p_o$  is observed proportion,  $n$  is the number of observations of the examined variable, the term  $\frac{1}{2n}$  is Yates’ correction factor and it is used when it is smaller then the absolute difference  $|p_o - p_k|$  in the numerator, see [19].

Null hypothesis for this test is that observed proportion does not statistically differ from the expected proportion based on Benford law. From the table of normal distribution, we read that Z-statistics above 1.64 indicate that the observed frequency of occurrence of a specific digit is significantly difference from the expected digit frequency with a 10% level of significance, Z-statistics above 1.96 indicate significantly difference with a 5% level of significance and Z-statistics above 2.57 indicate significantly difference with a 1% level of significance. Further, we will give all calculations with a 5% level of significance. The lack of a Z-test is reflected in the test frequencies of occurrence of single digit and compression with expectations frequency according to Benford’s law for that digit, not whether it is an occurrence all digits in line with expectations.

Second test which was used to check fitting with Benford’s distribution is K-S test. Appropriate statistics represents the maximum value of the deviation from Benford’s law, by summarizing the differences between the frequencies of the digits 1 to 9 that are obtained by testing sample data and frequencies that are theoretically expected. Appropriate statistics is given by the next formula:

$$KS = \max\{|f_{o1} - f_1|, |(f_{o1} + f_{o2}) - (f_1 + f_2)|, \dots, |(f_{o1} + f_{o2} + \dots + f_{o9}) - (f_1 + f_2 + \dots + f_9)|\}, \tag{5}$$

where  $f_{ok}$  represents frequency of occurrence of the digit  $k$  and it is determined by examining the data (observed frequency), while  $f_k$  represents the theoretical frequency of occurrence of the digit  $k$  and which is expected by Benford’s law.

During K-S testing, the obtained value of KS statistics is compared to critical value. The critical value with a 95% confidence level is  $c = 1.36/\sqrt{n}$ , where  $n$  denotes the number of considering values in a data set. The lack of the K-S test is noticed when examining large-scale sets. As the value of  $n$  increases, the appropriate p-value decreases, and we will not reject the null hypothesis if the distribution fit Benford’s distribution almost perfectly.

#### 4. MAIN RESULT

In 190 positions of investment, 34 start with the digit one, what is 17.89% of all positions, 46 start with the digit two, what is 24.21% of all positions and 23 start with the digit five, what is 12.11% of all positions. The differences which are presented in the next table are not statistically significant at most digits (Z-statistics is less

**Table 2:** Z test for the first digit of investment

Digit	Count	Frequency	Benford's law	Absolute difference	Z-stat
1	34	0.1789	0.3010	0.1221	3.5895
2	46	0.2421	0.1761	0.0660	2.2938
3	26	0.1368	0.1249	0.0119	0.3865
4	19	0.1000	0.0969	0.0031	0.0214
5	23	0.1211	0.0792	0.0419	2.0032
6	13	0.0684	0.0670	0.0015	0.0811
7	9	0.0474	0.0580	0.0106	0.4712
8	7	0.0368	0.0512	0.0143	0.7306
9	13	0.0684	0.0458	0.0227	1.3212
$\Sigma$	190	1	1		

Source: Authors' calculations.

then 1.96), except for the digits one, two and five. The result of the test for the first digit of investment is given in Table 2.

In 179 positions of new employment, 23 start with the digit five, what is 12.85% of all positions, 18 start with the digit seven, what is 10.06% of all position. The differences which are presented in the next table are not statistically significant at most digits (Z-statistics is less then 1.96), except for the digits five and seven. The result of the test for the first digit of new employment is given in Table 3.

**Table 3:** Z test for the first digit of new employment

Digit	Count	Frequency	Benford's law	Absolute difference	Z-stat
1	58	0.3240	0.3010	0.0230	0.5891
2	23	0.1285	0.1761	0.0476	1.5738
3	30	0.1676	0.1249	0.0427	1.6130
4	11	0.0615	0.0969	0.0355	1.4772
5	23	0.1285	0.0792	0.0493	2.3049
6	8	0.0447	0.0670	0.0223	1.0419
7	18	0.1006	0.0580	0.0426	2.2769
8	3	0.0168	0.0512	0.0344	1.9189
9	5	0.0279	0.0458	0.0178	0.9625
$\Sigma$	179	1	1		

Source: Authors' calculations.

The numeric analysis with K-S test for the first digits of investment is given in addition. According to the calculations, the most common first digit is two. It appeared in 24.21% of total cases. On the other hand, the digit eight had the lowest appearance, it appeared in 3.68% of total cases. However, in order to be sure, statistical K-S test is going to be applied. The result of the test for the first digit of investment is given in Table 4.

K-S test indicates that the null hypothesis can be reject at a 5% significant level, because empirical KS value is equal to 0.1221, and it is larger then theoretical KS value of 0.0987. So, the conclusion is that the first digit distribution of investment does not follow the Benford's law distribution.

Further, the numeric analysis with K-S test for the first digits of new employment is given in addition. According to the calculations, the most common first digit is one. It appeared in 32.40% of total cases. On the other hand, the digit nine had the lowest appearance, it appeared in 2.79% of total cases. However, in order to be sure, statistical K-S test is going to be applied. The result of the test for the first digit of new employment is given in Table 5.

K-S test indicates that the null hypothesis will not be rejected at 5% significant level, because empirical KS value is equal to 0.0522, and it is less then theoretical KS value of 0.1016. So, the conclusion is that the first digit distribution of new employment follows the Benford's law distribution.



**Table 4:** K-S test for the first digit of investment

Digit	Count	Frequency	Benford law	Cumulative density	Cumulative density of Benford's law	Absolute difference
1	34	0.1789	0.3010	0.1789	0.3010	0.1221
2	46	0.2421	0.1761	0.4211	0.4771	0.0561
3	26	0.1368	0.1249	0.5579	0.6021	0.0442
4	19	0.1000	0.0969	0.6579	0.6990	0.0411
5	23	0.1211	0.0792	0.7789	0.7782	0.0008
6	13	0.0684	0.0670	0.8474	0.8451	0.0023
7	9	0.0474	0.0580	0.8947	0.9031	0.0084
8	7	0.0368	0.0512	0.9316	0.9542	0.0227
9	13	0.0684	0.0458	1.0000	1.0000	0
$\Sigma$	190	1	1			

Source: Authors' calculations.

**Table 5:** K-S test for the first digit of new employment

Digit	Count	Frequency	Benford law	Cumulative density	Cumulative density of Benford's law	Absolute difference
1	58	0.3240	0.3010	0.3240	0.3010	0.0230
2	23	0.1285	0.1761	0.4525	0.4771	0.0246
3	30	0.1676	0.1249	0.6201	0.6021	0.0181
4	11	0.0615	0.0969	0.6816	0.6990	0.0174
5	23	0.1285	0.0792	0.8101	0.7782	0.0319
6	8	0.0447	0.0670	0.8547	0.8451	0.0096
7	18	0.1006	0.0580	0.9553	0.9031	0.0522
8	3	0.0168	0.0512	0.9721	0.9542	0.0178
9	5	0.0279	0.0458	1.0000	1.0000	0
$\Sigma$	179	1	1			

Source: Authors' calculations.

## 5. CONCLUSION

Benford's law is the phenomenon which was initially research as random. Today, there are many applications of that law in various field. It's primary application is in detecting possible fraud, irregularity or some errors in financial reports or other set of data. Modern software tools give us the opportunity easier application of the law, which is one of it's advantages in comparison with other tools and methods. Using the Benford's law we analyzed the validity of investment and new employment in the Republic of Serbia during the period of 2016-2022. In the testing of first digit with K-S test, we came to the conclusions that investments are not in accordance with Benford's law, while new employments are in accordance with Benford's law. In the testing of first digit with Z-test, we concluded that in investment the distributions of the digits one, two and five are statistically differ from expected distribution and in new employment the distributions of the digits five and seven are statistically differ from expected distribution. In the line with previous analysis, our conclusion is that more attention must be given to these digits.

The limitations of this method are situations when it is not possible to applied the law. That is when set of data is comprised of assigned numbers, or when the digits are formed with some intention, or when the digits are limited with minimal or maximal values, etc.

Future research could be engaged Benford's law with new set of data, for example from the field of insurance, and could be consider comparison with some others statistical methods. Application of new tests for fitting could be also one direction of the future research.

## REFERENCES

- [1] Balashov, V.S, Yan, Y. & Zhu, X. (2020). Who Manipulates Data During Pandemics? Evidence from Newcomb-Benford Law
- [2] Carrera, C . (2015). Tracking exchange rate management in Latin America. *Review of Financial Economics*, 25(10), 35-41.

- [3] Carslaw, C.A. (1988). Anomalies in income numbers: Evidence of goal oriented behavior. *Accounting Review*, 321-327.
- [4] De Ceuster, M.J.K., Dhaene, G. & Schatteman, T. (1998). On the hypothesis of psychological barriers in stock markets and Benford's Law. *Journal of Empirical Finance*, 5(3), 263-279.
- [5] Durtschi, C., Hillison, W. & Pacini, C. (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. *Journal of forensic accounting*, 5(1), 17-34.
- [6] Giles, D.E. (2007). Benford's law and naturally occurring prices in certain ebaY auctions. *Applied Economics Letters*, 14(30), 157-161.
- [7] Golbeck, J. (2015). Benford's law applies to online social networks. *PloS one*, 10(80), 135-169.
- [8] Gonzalez-Garcia, M.J. & Pastor, M.G.C. (2009). Benford's law and macroeconomic data quality. *International Monetary Fund*
- [9] Gorenc, M. (2019). Benford's Law As a Useful Tool to Determine Fraud in Financial Statements. *Management (18544223)*, 14(1).
- [10] Hindls, R. & Hronova, S. (2015). Benford's Law and Possibilities for Its Use in Governmental Statistics. *Statistika-Statistics and economy journal*, 95(2), 54-64.
- [11] Jošić, H. & Žmuk, B. (2021). Assessing the quality of Covid-19 data: Evidence from Newcomb-Benford law, 135-156.
- [12] Krakar, Z. & Žgela, M. (2009). Evaluation of Benford's low Application in stock prices and stock turnover. *Informatologia*, 42(3), 158-165.
- [13] Krakar, Z. & Žgela, M. (2009). Application of Benford's law in payment systems auditing. *Journal of Information and Organizational Sciences*, 33(1), 39-51.
- [14] Michalski, T. & Stoltz, G. (2013). Do countries falsify economic data strategically? Some evidence that they might. *Review of Economics and Statistics*, 95(2), 591-616.
- [15] Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of mathematics*, 4(1), 39-40.
- [16] Nigrini, M.J. (1996). A taxpayer compliance application of Benford's law. *The Journal of the American Taxation Association*, 18(1), 72.
- [17] Nigrini, M.J. & Mittermaier, L.J. (1997). The use of Benford's law as an aid in analytical procedures. *Auditing*, 16(2), 52.
- [18] Nigrini, M.J. & Miller, S.J. (2009). Data diagnostics using second-order tests of Benford's law. *Auditing: A Journal of Practice & Theory*, 28(2), 305-324.
- [19] Nigrini, M.J. (2012). *Benford's Law: Applications for forensic accounting, auditing, and fraud detection*. John Wiley & Sons, Vol 586.
- [20] Nye, J. & Moul, C. (2007). The political economy of numbers: on the application of Benford's law to international macroeconomic statistics. *The B.E Journal of Macroeconomics*, 7(1).
- [21] Papić, M., Vudrić, N. & Jerin, K. (2017). Benfordov zakon i njegova primjena u forenzičkom računovodstvu. *Zbornik sveučilišta Libertas*, 1(1-2), 153-172.
- [22] Slijepčević, S. & Blašković, B. (2014). Statistical detection of fraud in the reporting of Croatian public companies. *Financial theory and practice*, 38(1), 81-96.
- [23] Thomas, J.K. (1989). Unusual patterns in reported earnings. *Accounting Review*, 773-787.
- [24] Todorović, M. & Milosavljević, M. (2016). Kvantitativni revizijski alati za otkrivanje obmanjujućeg finansijskog izveštavanja. *XLIII Simpozijum o operacionim istraživanjima*, 127-130.
- [25] Tota, I., Aliaj, A. & Lamčja, J. (2016). The use of Benford's law as a tool for detecting fraud in accounting data. *Interdisciplinary Journal of Research and Development*, 3(1).
- [26] Warshavsky, M.S. (2010). Applying Benford's law in financial forensic investigations. *National litigation consultant's review*, 10(2), 1-4.
- [27] Zhang, J. (2020). Testing case number of coronavirus disease 2019 in China with Newcomb-Benford law. *arXiv preprint arXiv:2002.05695*
- [28] Izveštaj o dodeljenoj državnoj pomoći u Republici Srbiji 2019. Republika Srbija, Komisija za kontrolu državne pomoći.

## QUALITY EVALUATION OF CLUSTERING RESULTS: SILHOUETTE COEFFICIENT

MILAN STAMENKOVIĆ<sup>1</sup>, MARINA MILANOVIĆ<sup>2</sup>

<sup>1</sup> University of Kragujevac – Faculty of Economics, m.stamenkovic@kg.ac.rs

<sup>2</sup> University of Kragujevac – Faculty of Economics, milanovicm@kg.ac.rs

---

**Abstract:** *The quality evaluation of extracted series of possible solutions, resulting from the hierarchical clustering procedure, represents a critical analytical activity in conducting cluster analysis, since it enables the objective selection of an optimal division of multivariate observations in the available data set. Hence, the purpose of this Paper is to emphasize the applicative significance and potential of the criteria for evaluating the quality of cluster analysis results and selecting the optimal number of clusters, with a focus on the methodological-applicative properties of the silhouette coefficient.*

**Keywords:** *Cluster analysis, classification, quality evaluation, optimality criteria, silhouette coefficient*



## STATISTIČKA KLASIFIKACIJA NEEKONOMSKIH FAKTORA U FUNKCIJI UNAPREĐENJA POSLOVNE EFIKASNOSTI U BANKARSKOM POSLOVANJU

### STATISTICAL CLASSIFICATION OF NON-ECONOMIC FACTORS IN THE FUNCTION OF IMPROVING BUSINESS EFFICIENCY IN BANKING OPERATIONS

NINA UREMOVIĆ<sup>1</sup>, MIRJANA LANDIKA<sup>2</sup>, JANA ALEKSIĆ<sup>3</sup>, MAJA ĐURIĆ<sup>4</sup>

<sup>1</sup> Panevropski univerzitet „Apeiron“ – Fakultet poslovne ekonomije, [nina.d.uremovic@apeiron-edu.eu](mailto:nina.d.uremovic@apeiron-edu.eu)

<sup>2</sup> Panevropski univerzitet „Apeiron“ – Fakultet poslovne ekonomije, [mirjana.f.landika@apeiron-edu.eu](mailto:mirjana.f.landika@apeiron-edu.eu)

<sup>3</sup> Panevropski univerzitet „Apeiron“ – Fakultet poslovne ekonomije, [jana.s.aleksic@apeiron-edu.eu](mailto:jana.s.aleksic@apeiron-edu.eu)

<sup>4</sup> Panevropski univerzitet „Apeiron“ – Fakultet poslovne ekonomije, [maki20111989@gmail.com](mailto:maki20111989@gmail.com)

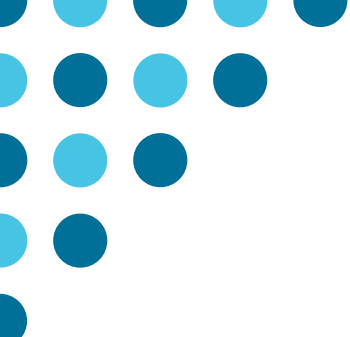
**Rezime:** Zadatak i cilj provedenog istraživanja odnosi se na ekstrakciju faktora iz domena neekonomske motivacije zaposlenih u bankarskom sektoru Republike Srpske kojima bi se mogla modelirati njihova radna efikasnost u realnim okolnostima i na konkretnim radnim zadacima. Istraživački projekat proveden je anketiranjem ispitanika na relevantnom uzorku zaposlenih u ciljanom sektoru na osnovu čega je iz skupa od petnaest promjenljivih izdvojeno četiri faktora kojima je moguće iskazati neekonomske motivišuće faktore zaposlenih u bankarskom sektoru, a obuhvataju: razumijevanje, komunikaciju, olakšavanje i brigu. Nastavak istraživanja mogao bi ići u pravcu formiranja adekvatnog regresionog modela kojima bi se iskazivala vrijednost učinka pomoću izdvojenih faktora kao nezavisnih promjenljivih.

**Ključne reči:** Statističke metode – Faktorska analiza, Obrazovanje, Obrazovanje za rad, Ekonomski razvoj

**Abstract:** The task and the goal of the conducted research were related to the extraction of factors from the domain of non-financial motivation of the banking sector employees of Republika Srpska, which could be used to model their work efficiency in real circumstances and specific work tasks. The research project was carried out by surveying respondents on a relevant sample of employees in the target sector, on the basis of which four factors were selected from a set of fifteen variables that can be used to express the noneconomic motivating factors of employees in the banking sector, and include: understanding, communication, facilitation and care. The further research could go in the direction of forming an adequate regression model, which would show the performance value using extracted factors as independent variables.

**Keywords:** Statistical Methods - Factor Analysis, Education, Business Education, Economic Growth





# VIŠEKRITERIJUMSKA ANALIZA I OPTIMIZACIJA

## MULTIPLE CRITERIA ANALYSIS AND OPTIMIZATION







## OMNI-CHANNEL RETAIL: MULTIPLE CRITERIA DECISION AIDING APPROACH

MLADEN STAMENKOVIĆ<sup>1</sup>, DRAGAN STOJKOVIĆ<sup>2</sup>, ALEKSA DOKIĆ<sup>3</sup>

<sup>1</sup> Univerzitet u Beogradu – Ekonomski fakultet, mladen.stamenkovic@ekof.bg.ac.rs

<sup>2</sup> Univerzitet u Beogradu – Ekonomski fakultet, dragan.stojkovic@ekof.bg.ac.rs

<sup>3</sup> Univerzitet u Beogradu – Ekonomski fakultet, aleksa.dokic@ekof.bg.ac.rs

---

**Abstract:** *The prevalence of the new type of customers seeking a seamless shopping experience created new multiple channel market challenges. Traditional retailers are under pressure to respond to growing customer demands by incorporating new digital touchpoints into their physical sales channels. However, the ever-present question of integrating offline and online sales channels remains. Contemporary literature does not provide a uniform, objective, and quantitative approach to multiple channel classification. The main research goal of this paper was to develop a strategic roadmap for implementing multiple channel strategies based on a quantitative approach, which both academics and practitioners can use. We propose a methodological framework to surpass the gap. The paper focuses on specific dimension attributes and their relations in different multiple channel scenarios. Attributes are analyzed using a dominance-based rough set approach (DRSA) to provide a set of decision rules to identify individual retailers' multiple channel strategy. We develop a multiple channel strategic roadmap with a multiple channel-related attribute prioritization and a predictive capacity for brick-and-click newcomers. Derived "if-then" decision rules help retail managers focus on channel aspects relevant to implementing a specific multiple channel strategy. A fully quantified roadmap also simplifies the optimization process, allowing retailers to allocate limited resources to improving dimension attributes most important for the specific multiple channel strategy. Finally, the roadmap aids industry newcomers by providing unique strategic pathways toward the desired multiple channel strategy.*

**Keywords:** *Omni-channel retail, Multiple Criteria Decision Aiding, Dominance-based Rough Set Approach*

CIP - Каталогизација у публикацији  
Народна библиотека Србије, Београд

519.8(082)

СИМПОЗИЈУМ о операционим истраживањима (49 ; 2022 ;  
Врњачка Бања)

Zbornik radova = Proceedings / XLIX Simpozijum o  
operacionim istraživanjima = XLIX International Symposium  
on Operational Research, SYM-OP-IS 2022, Vrnjačka Banja,  
19-22. septembar 2022. ; urednici, editors Zorica Mladenović,  
Mladen Stamenković ; [izvršni organizator, executive organizers  
Ekonomski fakultet, Univerzitet u Beogradu]. - Beograd :  
Ekonomski fakultet Univerziteta, Centar za izdavačku delatnost  
= Belgrade : The Publishing Centre of the Faculty of Economics,  
2022 (Beograd : Službeni glasnik). - XVII, 639 str. : ilustr. ; 30 cm

Radovi na srp. i engl. jeziku. - Tiraž 50. - Str. VII: Predgovor /  
urednici = Preface / editors. - Napomene i bibliografske reference  
uz tekst. - Bibliografija uz svaki rad. - Abstracts.

ISBN 978-86-403-1750-4

a) Операциона истраживања -- Зборници

COBISS.SR-ID 85350665