

*University of Novi Sad  
Technical faculty "Mihajlo Pupin" Zrenjanin*

**Proceedings of the 11<sup>th</sup> International Conference  
on Applied Internet and Information Technologies  
AIIT 2021**

15 October, 2021, Zrenjanin, Serbia





**University of Novi Sad  
Technical faculty "Mihajlo Pupin"  
Zrenjanin, Republic of Serbia**



**XI INTERNATIONAL CONFERENCE ON  
APPLIED INTERNET AND INFORMATION TECHNOLOGIES**

**AIIT2021  
PROCEEDINGS**



October 15, 2021  
Zrenjanin, Serbia

**Publisher and organizer of the conference:**

**University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Republic of Serbia**

**For publisher:**

Dragica Radosav, PhD, Professor  
Dean of Technical faculty "Mihajlo Pupin",  
Zrenjanin,  
Republic of Serbia

**Proceedings editors:**

Visnja Ognjenovic, PhD – main editor  
Dalibor Dobrilovic, PhD  
Evgeny Cherkashin, PhD  
Andrijana Bocevaska, PhD

**Conference chairman:**

Visnja Ognjenovic, PhD

**Cover design:**

Visnja Ognjenovic, PhD

**Conference co-chairmen:**

Dalibor Dobrilovic, PhD  
Andrijana Bocevaska, PhD  
Evgeny Cherkashin, PhD

**Technical preparation of proceedings:**

Visnja Ognjenovic, PhD  
Dalibor Dobrilovic, PhD  
Sinisa Mihajlovic, MSc  
Marko Blazic, MSc

**e-Proceedings**

ISBN 978-86-7672-352-2

**Disclaimer:**

All rights reserved. No part of this proceeding may be reproduced in any form without written permission from the publisher.

The publisher and editors are not responsible either for the statements made or for the opinion expressed in this publication.

The authors solely are responsible for the content of the papers and any copyrights, which are related to the content of the papers.

CIP - Каталогизacija u publikaciji  
Библиотеке Матице српске, Нови Сад

004(082)(0.034.4)

**INTERNATIONAL Conference on Applied Internet and Information Technologies (11 ; 2021 ; Zrenjanin)**

Proceedings [Elektronski izvor] / XI International Conference on Applied Internet and Information Technologies AIIT 2021, 15 October, 2021, Zrenjanin, Serbia ; [organizer] Technical Faculty "Mihajlo Pupin", Zrenjanin. - Zrenjanin : Technical Faculty "Mihajlo Pupin", 2022. - 1 elektronski optički disk (CD-ROM) : tekst, ilustr. ; 12 cm

Sistemski zahtevi: Nisu navedeni. - Naslov sa nasl. ekrana. - Elektronska publikacija u formatu pdf opsega 215 str. - Bibliografija uz svaki rad. - Registar.

ISBN 978-86-7672-352-2

a) Информационе технологије -- Зборници

COBISS.SR-ID 62497289

## **INTRODUCTION**

---

The objectives of the International conference on Applied Internet and Information Technologies are aligned with the goal of regional economic development. The conference focus is to facilitate the implementation of Internet and Information Technologies in all areas of human activities. The conference provides a forum for discussion and exchange of experiences between people from government, state agencies, universities, research institutions, and practitioners from industry. Information technologies change during time and this year AIIT conference addressed the diversity of ICT application areas and relevant research topics such as:

- Information systems
- Software engineering and applications
- Data science and big data technologies
- Business intelligence and IT support to decision-making
- Communications and computer networks
- Data and system security
- Distributed systems
- Internet of Things
- Embedded systems
- Software quality
- Software maintenance
- Computer graphics
- IT management
- E-commerce
- E-Government
- E-Education
- Internet marketing
- ICT practice and experience

Information technologies enable collaboration across the globe. This year the conference was successfully co-organized by 5 institutions from 4 countries - Serbia, North Macedonia, Russia, and Bulgaria. It has been managed in collaboration with 4 co-chairmen from Serbia, North Macedonia, and Russia.

International Conference on Applied Internet and Information Technologies (AIIT) is an annual conference that was held since 2012, based on successful results of the International Conference on Information and Communication Technologies for Small and Medium Enterprises in 2011. This year, AIIT2021 was held on October 15, 2021, in Zrenjanin, Serbia.

Due to a COVID-19 pandemics, the conference is held in virtual form, with online presentations with Google Meet, and streaming video and poster presentations available at the web site of the conference (<http://www.tfzr.uns.ac.rs/aiit/>). There were 41 accepted papers and 2 accepted papers in abstract with 105 authors from 14 countries (Serbia, North Macedonia, Montenegro, Bosnia and Herzegovina, Croatia, Hungary, Romania, Bulgaria, Russia, India, Malaysia, Saudi Arabia, Egypt, Canada). The papers are presented online, or in the video stream and poster sessions. Within the video presentation session, there is a presentation of IT company Crater Training Centar, Belgrade, Serbia.

The AIIT 2021 organizing committee would like to thank the authors of the papers for their contribution. All submitted papers were peer-reviewed by the members of the AIIT2021 program committee. Each submitted paper was assigned to at least two reviewers from different countries and the paper analysis was conducted as a double-blind review.

Special gratitude is addressed to many reviewers from co-organizing institutions that made a great impact on the quality of papers. The AIIT organizing committee especially appreciates the IT company's efforts in supporting the conference by its participation.

Information technologies are integrated with every human activity. IT application enhancements are encouraged by university research, business organizations, public institutions, and the IT industry. The AIIT organizing committee welcomes future presentations of work in this field at the next AIIT conference, hoping that all of us will meet again in the real conference event.

**Conference chairs:**

**Visnja Ognjenovic**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**co-chairman Dalibor Dobrilovic**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**co-chairman Evgeny Cherkashin**, *Institute of High Technologies, Irkutsk, Russia*

**co-chairman Andrijana Bocevska**, *Faculty of Information and Communication Technologies - Bitola, North Macedonia*

## MAIN ORGANIZING INSTITUTION

---



Technical Faculty "Mihajlo Pupin" Zrenjanin  
University of Novi Sad  
SERBIA  
<http://www.tfzr.uns.ac.rs/>

## ORGANIZING PARTNER INSTITUTIONS

---



Faculty of Information and Communication Technologies - Bitola  
"St. Kliment Ohridski" University - Bitola  
North MACEDONIA  
<http://www.fikt.edu.mk/>



Irkutsk National Research Technical University  
Institute of High Technologies, Irkutsk  
Matrosov Institute for System Dynamics and Control Theory SB RAS  
Irkutsk, RUSSIA  
<http://www.istu.edu/>



Irkutsk State Transport University  
Irkutsk, RUSSIA  
[www.irgups.ru/en/about-university](http://www.irgups.ru/en/about-university)



Faculty of Engineering  
South-west university "Neophyte Rilsky"-Blagoevgrad  
BULGARIA  
<http://www.swu.bg/>

## CONFERENCE SUPPORTING INSTITUTIONS

---

Municipality of Zrenjanin, Serbia  
Regional Chamber of Commerce, Zrenjanin, Serbia  
Regional Center for development RCR Banat, Zrenjanin, Serbia  
Zrenjaninski IKT Klaster, Zrenjanin, Serbia  
Business Incubator, Zrenjanin, Serbia

## COMPANIES PRESENTED AT AIIT2021

---



Crater Training Center  
Belgrade, Serbia  
<https://school.craterstudio.com>



Levi 9  
Zrenjanin, Novi Sad, Serbia  
<https://www.levi9.com/>



Vega IT  
Zrenjanin, Serbia  
<https://www.vegait.rs/>

## ACKNOWLEDGMENT

---

Thanks to the tourist organization of the City of Zrenjanin on documentary movies of Zrenjanin, which were used at the conference AIIT2021.

## CONFERENCE CHAIRS

---

**Visnja Ognjenovic**, University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia

**co-chairman Dalibor Dobrilovic**, University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia

**co-chairman Evgeny Cherkashin**, Institute of High Technologies, Irkutsk, Russia

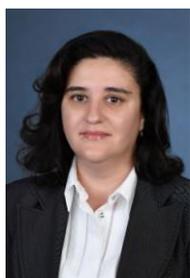
**co-chairman Andrijana Bocevska**, Faculty of Information and Communication Technologies - Bitola, North Macedonia



**Dr. Visnja Ognjenovic** is an assistant professor at the Information Technology department at the University of Novi Sad, Technical Faculty „Mihajlo Pupin“, Zrenjanin. She has received her PhD in Information technology in 2016. Her teaching areas are in the field of Artificial Intelligence, Data Science, and Computer graphics. Her research interests are in the area Data Mining, Machine Learning, Computer graphics, etc. Dr. Visnja Ognjenovic has more than 70 research articles published in international journals and conferences and she has participated in several EU and national funded projects.



**Dr. Dalibor Dobrilovic** is an associate professor at the Information Technology department at the University of Novi Sad, Technical Faculty „Mihajlo Pupin“, Zrenjanin. He has received his PhD in Information technology in 2012. His teaching areas are in the field of Computer Networking, Communication Systems, and Data and computer systems security. His research interests are in the area of IoT, Smart Cities, Wireless communications, Wireless sensor networks, Computer Networking, Engineering education, etc. Dr. Dalibor Dobrilovic has more than 120 research articles published in international journals and conferences and he has participated in several EU and national funded projects. He is a member of IEEE and ACM societies. Since 2019 he is the president of the Council of the Technical Faculty "Mihajlo Pupin" Zrenjanin.



**Dr. Andrijana Bocevska** is an Associate Professor at the Faculty of Information and Communication Technologies, "St. Kliment Ohridski" University – Bitola, R. North Macedonia. She received her MSc and PhD degrees in Mechanical Engineering in December 2001 and October, 2012, respectively. Her research areas include: Integrated computational methods and applications, Computer integrated manufacturing, Product engineering, technology and systems. Dr. Andrijana Bocevska has published 6 books and more than 35 research articles published in international journals, conferences and congresses and she has participated in several EU and domestic funded projects. Dr. Andrijana Bocevska currently teaches subjects in: Application software, Solid modeling, Computer integrated manufacturing, Scientific visualization in virtual environments, Product Lifecycle Management. Associate Professor Andrijana Bocevska was appointed to the position of the Vice-dean for teaching and international cooperation on 01 March 2018.



**Dr. Evgeny Cherkashin** has graduated from Irkutsk State Technical University at 1996, at 1999 defended dissertation "Quant/2 system for automatic theorem proving" on application new logical calculus for control technical systems. After that, he mostly deals with application first-order logical inference systems for model identification algorithm synthesis, software model transformations. Most of the scientific activity is carried on in Institute for Systems Dynamics and Control theory of Siberian Branch of Russian Academy of science, at Laboratory of Complex information systems. E.Cherkashin instructs students of two Irkutsk universities programming, software design, real-time system engineering and artificial intelligence. He is author more of 160 scientific papers.

## TABLE OF CONTENTS

### INVITED PAPERS

Customized image processing as a solution for compensating color vision deficiencies in the digital environment <i>N. Milić Keresteš</i>	1
Artificial Intelligence Implemented in Covid-19 Detection <i>Kostandina Veljanovska</i>	2
Web GUI Upgrade for Manual Semantic Extraction i E-Learning <i>M. Jovanovic</i>	7

### REGULAR PAPERS

An overview of 4D medical image compression <i>Maja Gaborov, Marko Blažić, Dijana Karuović, Mila Kavalić, Igor Vecštejn, Dragana Milosavljev</i>	9
Formula One Lap Time Data Visualization and Prediction Software <i>F. Đorđević, D. Ivetić</i>	15
Energetics Aspect of Smart Buildings Based on Internet Of Things Architecture – An overview <i>Maja Gaborov, Dragan Ivetić, Srđan Popov</i>	20
An Android-based Application for Reading Serbian Identity Cards – Challenges and Design Considerations <i>J. Jovanović, D. Dragan, D. B. Gajić, and V. B. Petrović</i>	26
The Influence of Audio vs. Multimedia Classroom Instruction on Critical Thinking of EFL Learners <i>Lela Ivanovska</i>	32
Modelling Business Intelligence Systems for Effective Decision Making <i>Mihalj Bakator, Dragan Čočkal, Dejan Đorđević, Melita Čočkal-Hronjec, Dragana Milosavljev</i>	36
Improving CRM with Internet of Things and Big Data <i>Mihalj Bakator, Dragica Radosav, Mila Kavalić, Dragana Milosavljev, Edit Terek Stojanović</i>	41
Achieving Sustainable Development Through Information Systems <i>Mihalj Bakator, Dragica Radosav, Nataša Đalić, Mila Kavalić, Dragana Milosavljev</i>	46
Integration of Heterogeneous Data into Electronic Patient Records <i>Snezana Savoska, Blagoj Ristevski, Natasha Blazheska-Tabakovska, Ilija Jolevski, Andrijana Bocevaska, Vladimir Trajkovic</i>	50
Derivation, Analysis and Simulation of Outage Performance of MIMO Multi-branch SC Diversity System in $\alpha$ - $\mu$ Fading and Co-Channel Interference Environment <i>D. Krstic, S. Suljovic, N. Petrovic, Z. Popovic, and S. Minic</i>	55
The impact of air pollution on bacteriological flora <i>D. Jovanovski, E.M. Jovanovska, K. Popovska, A. Naumoski</i>	60
Application of the Blockchain Technology in Medicine and Healthcare <i>Panche Tashevski, Blagoj Ristevski, Snezana Savoska, Ilija Jolevski, Mimoza Mijoska</i>	65
A Survey of Energy-efficient Solutions for 5G Networks <i>M. Ilić, V. Mikić, A. Zakić and D. Zlatković</i>	70
Review Analysis of E-Government in the Republic of Serbia <i>M. Mazalica, B. Radulovic and A. Stojkov</i>	74
A Survey of Machine Learning Techniques Used in Recommender Systems <i>Nora PireciSejdiu, Blagoj Ristevski, Ilija Jolevski</i>	79
Storage, selection and visualization raster models in spatial databases <i>A. Vasiljević, M. Čeliković, S. Popov</i>	85

Secure Communication in Early Fire Detection Systems <i>R. Radišić, D. Dinu and S. Popov</i>	90
Fog Computing architecture for IoT Smart traffic applications <i>M. Mazalica, D. Ivetić</i>	95
Choosing the best Python web framework for beginner according to experienced users <i>Dejan Viduka, Boris Ličina and Luka Ilić</i>	100
Analysis Of Students' Learning And Achievement Based On Data From The University Information Systems <i>Ilker Ali, Natasha Blazeska-Tabakovska , Igor Nedelkovski, Blagoj Ristevski</i>	104
Improving Test Execution Phase through Diversity of Approaches: A Systematic Literature Review <i>Sara Gračić and Vuk Vuković</i>	110
Comparison of Dart and JavaScript Programming Languages <i>I. Vecštejn, V. Ognjenović, E. Brtko, T. Milić, M. Gaborov</i>	116
Mitigating Covid-19 Impact on Small Businesses and Startups Using Digital Technologies <i>Z. Kotevski and A. Shijakova Kotevski</i>	120
Conceptual Data Model Design for Adaptable Web-Based Museum Information System <i>Ljubica Kazi, Dragica Radosav, Zoltan Kazi, Dejan Masliković, Natalija Vulikić, Tijana Stanković Pešterac, Biljana Radulović and Ivana Berković</i>	126
Clean Code Quality Attributes and Measurements: an Initial Review <i>Lj. Kazi, S. Mihajlović and M. Bhatt</i>	133
Synthetic media (Deepfake) generation and detection methods and challenges <i>Sasa Arsovski, Angely Sim Jia Wun, Branko Markoski, Aleksandar Sofić, Velibor Premcevski</i>	138
Review of software architecture patterns in traffic systems <i>Z. Stojanov, G. Jotanovic, G. Jausevac and D. Perakovic</i>	142
The Application of Semi-Linguistic Summaries in Traffic Data Analysis <i>V. Brtko, M. Sisak, V. Makitan, G. Jotanović, G. Jauševac</i>	148
Recommender systems for carer guidance <i>Trần Đức Thế , Viktoria Kopylova , Evgeny Cherkashin, Nikita Lukyanov</i>	152
Machine Learning Techniques for Smart Digital Technologies <i>H. M. Said</i>	157
Utilization Of Different Approaches For Data Security In Business Intelligence <i>V. Naneva and K. Stefanova</i>	162
Adopting AR and Deep Learning for Gamified Fitness Mobile Apps: Yoga Trainer Case Study <i>M. Radenkovic, V. Nejkovic, N. Petrovic</i>	167
Lean Production and Industry 4.0 <i>Sanja Stanisavljev, Zlatko Košut, Saša Zec, Branko Markoski, Željko Stojanović</i>	172
Use of CNNs on mobile devices to protect data from malware and unauthorized attacks <i>S. Mihajlović, D. Ivetić, I. Berković</i>	175
Creating a mobile application using the Kotlin programming language <i>B. Babić, E. Brtko, I. Vecštejn</i>	180
Review of challenges in identifying microservices from software artifacts <i>A. Stojkov and Ž. Stojanov</i>	185
Green Cloud Computing in the Purpose of Energy Efficiency <i>V. Mikić, M. Ilić, A. Zakić and D. Zlatković</i>	190
Designing the prototype of a scalable smart gardening system for testing and evaluation <i>S. Felbab, D. Dobrilović, Z. Čović, J. Simonc</i>	194
A Model for Integration of Internet of Things Systems in a Smart City <i>H. Dimova Popovska, T. Dimovski, and I. Hristoski</i>	199

Application of CNNs on mobile devices for emotion recognition based on facial expressions <i>S. Mihajlović, D. Ivetić, I. Berković</i>	203
---	-----

**APPENDIX**

Organizing Committee	208
Program Committee	210
List of Reviewers	213
Index of Authors	214

# **INVITED PAPERS**

# Customized image processing as a solution for compensating color vision deficiencies in the digital environment

N. Milić Keresteš

Faculty of Technical Sciences, Department of Graphic Engineering and Design, Novi Sad, Serbia  
milicn@uns.ac.rs

**Abstract** - Colour vision deficiency is a functional disorder, predominantly genetic, manifested with reduced or complete loss of colour discrimination ability.

The research objective is to resolve the accessibility issue of visual information encoded with colour for the population with colour vision deficiencies. Considering the shortcomings of previous image enhancement methods for colour deficient observers, so-called daltonization methods, as well as the prevalence of the target population (over 200 million users), there is still an open search for a new daltonization method that optimally compensates weak colour discrimination and new ways for its implementation in a digital environment.

The content-independent daltonization methods do not take into account the image content and the spatial distribution of confusing colours in images. As a result, content-independent colour mapping does not assure colour differentiation by colour deficient observers. This way, it often happens that recolouring solves the initial problem, but creates a new confusing pair or, even worse, colours that are clearly different can be remapped into indistinguishable combinations. The content-dependent daltonization methods include more complex and computationally demanding image processing where the final pixel value depends on other image colours and/or its location. Although providing more successful differentiation of coloured elements, these methods are usually focused only on severe, dichromatic deficiencies causing problems with exaggeratedly changed colours and unnatural results.

The primary research objective is to define an image optimization model, customised for different types and severities of colour deficiency and different image content. The purpose of proposed image adaptation is to correct confusingly coloured image segments that users with anomalous colour vision perceive as the same while preserving the image naturalness by restricting, for each colour, an area of admissible remapping. This achieves a recolouring balance where resulting colours are made sufficiently distinctive from each other but do not deviate too much from the initial values. Within the defined model, the following image enhancement methods (daltonization methods) have been proposed: enhancement based on the deficiency type, enhancement based on the deficiency severity and universal enhancement.

The subjective evaluation confirmed that the severity of colour deficiencies affects the assessment of the image quality improvement, where users with mild deficiencies better evaluate the severity-based daltonization compared to

users with more severe deficiency. It was also revealed that the image content affects the choice of preference — the severity-based daltonization is preferred in case of natural scene images while the type-based daltonization provides better chromatic contrast for the artificial images.

**Keywords:** colour perception, colour vision deficiencies, digital image processing, colour gamut

# Artificial Intelligence Implemented in Covid-19 Detection

Kostandina Veljanovska\*

\* Department of Intelligent Systems, Faculty of ICT, University "St. Kliment Ohridski",  
Bitola, North Macedonia  
kostandinav@gmail.com

**Abstract** – Health professionals and scientists are at the forefront of the fight against the virus that humanity is currently battling. The next generation of scientists and physicians must be trained and prepared for a future pandemic, to respond to unexpected epidemics, and they must also learn to strengthen scientific communication. Scientists need to keep researching. Our attempt is in that direction to contribute to the study of a rapid way of diagnosing covid-19 that uses artificial intelligence. Namely, a machine learning algorithm for patterns recognition has been modeled in diagnosing the disease caused by covid-19 virus that uses a review of X-ray images.

**Keywords** - artificial neural network, covid-19, machine learning, deep learning, convolutional neural network

## I. INTRODUCTION

The pandemic we have been facing for more than a year and a half is a widespread risk to humans and animals, and the effects are felt in everyday life, especially in the health and economic sector. Health professionals and scientists are at the forefront of the fight against the virus, together with people around the world managing the situation as best as they can. The scientific community is experiencing a negative impact from the outbreak of the virus, facing the closure of universities, research centers and laboratories, as well as the cancellation and postponement of some scientific events, professional conferences, symposia, workshops and training programs [1,2]. Universities, research centers and laboratories as a fertile ground for the development of new ideas and the promotion of advances in the science and the scientific community are in an inevitable situation. As scientists try to understand a new coronavirus and reduce the chaos it has caused, the epidemic has created chaos in science itself [2]. Reducing the number of scientific events and closing down scientific jobs will result in extended research time and, in some cases, will have to start again with full experiments, or put the experiments on hold, or reduce them at a minimum. The results of this will produce an economic burden on the researcher, eventually prompting psychological stress, anxiety, tension or depression, which will culminate in reduced scientific success [3].

Covid-19 is a viral disease caused by the SARS-CoV-2 virus that first appeared in the city of Wuhan, China, in December 2019, and then has spread around the world,

leading to the declaration of a global pandemic. The virus has devastating effects on the lives of people around the world and to date it has infected over 226 million people. About 90% of them were cured, and 4.66 million people died, according to the Worldometers reference website. The general symptoms of the virus are: runny nose, sore throat, fever, cough, and headache. Some of the patients do not show strong symptoms of the virus, and some of them develop a severe form of the disease, especially combined with pneumonia of the lungs. Finding a way to quickly and accurately detect the disease is very important in order to reduce the negative effects of this disease and save human lives. Detection of coronavirus can be done in several ways: PCR - test, blood check and chest X-ray imaging. The first two are based on a simple comparison of numbers and reference values, but the third method requires more detailed consideration and analysis by experienced physicians. Our idea is to help in that direction and to develop a system that uses artificial intelligence (AI) to detect coronavirus in patients from radiographic images of their chest.

## II. CORONAVIRUS DETECTION

Covid-19 is most commonly diagnosed with polymer chain reaction (PCR) and serological testing, for which there is a lack of necessary material and specialized personnel when performing these tests in regions and at a time when those regions are quite affected. In addition, PCR may have relatively low sensitivity. Therefore, alternative methods are needed to support the diagnosis of covid-19, such as noninvasive imaging. For example, computed tomography (CT) images can be used to detect certain manifestations in the lungs associated with covid-19, and in general, CT is a more accurate chest imaging technique and has a higher sensitivity and efficiency than X-rays of chest. However, the use of CT to detect covid-19 places a significant burden on diagnostic departments at times when they need to respond quickly to help minimize the risk of infection spreading. In fact, X-rays are part of routine screening of patients and remain the primary way to detect pneumonia due to shorter time and lower cost and, normally, lower radiation exposure compared to CT. Hence, our interest was directed in that area.

### III. ARTIFICIAL INTELLIGENCE IN MEDICAL DIAGNOSTICS

#### A. General Applications

Artificial intelligence (AI) models are already yielding successful results in the analysis of medical data. In fact, new machines with artificial intelligence models have already been developed and work similarly to experts in specific diagnostic tasks. In addition, AI systems are adopted to extract information from medical imaging with the ultimate goal of creating tools to reduce diagnostic errors, improve efficiency, and reduce costs. These systems are typically embedded in image-based decision support systems to assist shooting professionals. Medical recording enables characteristic activities such as risk assessment, diagnosis, prognosis, detection of response to a particular therapy. Automatic processing requires less laboratory infrastructure and supplies, as well as fewer healthcare personnel. In this context, AI-based models that use medical X-rays as input are an alternative worth exploring for the automatic detection of covid-19 in patients for clinical analysis and diagnosis.

#### B. Artificial Intelligence in Coronavirus Detection

To study the impact and dynamics of the pandemic that is relevant to AI, various academic databases on clinical applications of machine learning and deep learning should be studied, including clinical features, electronic medical records, medical imaging (CT, X-ray, ultrasound images, etc.) to diagnose covid-19. This field has already being researched in scientific circles [5, 6]. Artificial intelligence techniques have been used to detect pneumonia by computed tomography of the lungs [7, 8]. Machine learning models have also been applied to covid-19 case data to predict infected cases and recovery rates using chest X-rays [9]. The k-NN model has been studied to distinguish imaging from a positive patient and imaging from a patient with another disease [10], or the AI model which, in addition to these two types of imaging, also includes imaging of healthy patients [11].

### IV. MODEL DEVELOPMENT

For the purposes of this research, deep learning techniques as well as general machine learning techniques and algorithms were studied, and it was concluded that one of the most effective ways to analyze visual images when it comes to deep learning are convolutional neural networks (CNN). Experiments are already known for the diagnosis of coronavirus using CNN from CT scan [12, 13, 14] and from X-ray scan [15, 16, 17, 18, 19]. Analysis of the functioning of the database biases in the diagnosis of the virus is also part of scientific experiments with CNN [20]. Convolutional neural network has been applied on audio recordings of patients coughing to make an accurate diagnosis of covid-19 [21].

The artificial intelligence model includes machine learning (ML) and deep learning (DL) in order to automatically detect significant patterns in data and solve problems that are impossible (or impractical) to represent and solve with conventional algorithms [22, 23]. Deep

Learning (DL) learns high-level abstractions in data using hierarchical architectures. It combines several layers of nodes to build a gradually more abstract representation of the data, allowing concepts such as categorization or classification of objects to be learned directly from raw data collected by onsite-mounted sensors. The current success of DL is directly related to the production of inexpensive graphics cards with multiple processors or graphics processing units that increase speed and reduce training time to create a deep learning model. Convolutional neural networks (CNN) are used to recognize shapes / forms / patterns and are the dominant DL architecture for image classification that may even compete with the human ability to perform recognition and classification tasks.

#### A. Artificial Neural Network – Convolutional Neural Network

CNN brought a revolution in the field of computer vision by increasing the accuracy of image classification, but also many times improved scene classification, object detection, semantic segmentation of biological images and face detection, text recognition and human body recognition in natural images. The main practical success of CNN is the face recognition and automatic driving of cars.

The CNN uses hierarchical layers of convolutional filters to mimic the effects of sensory fields in the visual cortex of animals, taking advantage of the local spatial correlations present in the images [22]. The CNN-based model generally requires a large set of training samples to achieve good generalization skills.

Convolutional neural networks as one of the most powerful deep learning algorithms designed for image processing contain three types of layers (levels): convolutional, pooling layers and fully connected layers.

CNN is a multilayer perceptron that uses a bit of pre-processing unlike other image classification algorithms. The network learns to optimize filters through automatic learning as opposed to conventional algorithms where such filters are manually designed. This means that CNNs have a key advantage because they are independent of prior knowledge and manual property extraction. They use convolution in at least one network level instead of general matrix multiplication [24].

The following figure (Figure 1.) shows an example of a convolutional calculation where a block of pixels through a filter generates a map of properties, i.e. an image that will be used for further processing.

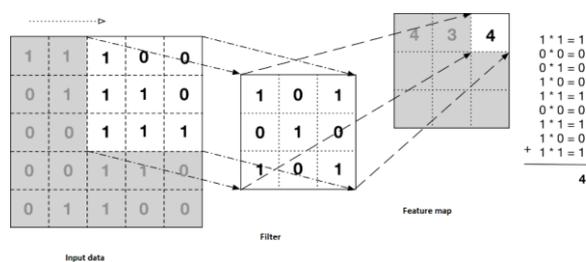


Figure 1. Method of convolutional calculation

The convolutional layer contains a set of filters and their parameters have to be learned. Every individual filter is convolved with the input data to calculate an activation map which is made of neurons. Normally, the height and weight of the filters are smaller compared to the same of the input volume.

The way convolutional calculation transforms size of data, we can see in following calculation. If the size of the image is  $N \times N$  and the size of the filter is  $F \times F$ , then after convolution the result will be matrix with dimension:

$$(N \times N) \text{ convolve } (F \times F) = (N-F+1) \times (N-F+1)$$

Each step at the convolution level is followed by a non-linear activation function.

The output from the convolutional layer contains high-level features extracted from the data. A fully-connected layer in order to learn non-linear combinations of these features is added after cycles of convolution and reduction of the size of data features.

Pooling layer is usually used to reduce the spatial volume of input images after convolution. It is used between two convolutional layers. The max pooling or average pooling or L2-norm pooling is used in order to reduce the spatial volume of the input image.

At the end, a fully connected layer which involves weights, biases, and neurons connects neurons from one layer to neurons in another layer. It is used to classify images between different categories as process of training the network goes on.

#### B. Covid-19 Detection Algorithm

For the purposes of this study, we developed a model that, based on collected X-ray images of the chest, recognizes positive and negative patients on covid-19, that is, images of patients who tested positive for PCR and patients who tested negative for coronavirus or patients without any diseases. A set of chest X-ray images of people that have been processed and divided into two sub-sets for training and testing is used as an input database. The data were taken from a database available online and were properly prepared for processing and designing the model. We used Keras library. The experiments were performed in two phases: training of the model in 10 epochs as the first phase and second phase, training in 20 epochs. In both phases, almost 90% of the data were used (224 images) for training and 12% for testing.

### V. DISCUSSION OF THE RESULTS

Our model contains four convolutional levels. The first level accepts pictures and does filtration and activation. The next level is pooling level where the pictures are processed and divided into blocks. Overfitting of the model is avoided using dropout of certain percentage of data. The last level of the network uses only one dimensional data, so three dimensional data are flattened with specific function.

Database has 224 images for model training. As the process of training was going on, it was evident that the model was improving in terms of increasing the accuracy and lowering the loss in predicting the result of diagnosis

for patients. Still, there were epochs where the model entered in over-fitting, stagnating the process of increasing the accuracy of prediction.

As it can be seen from Table I., in the first phase, the maximum result that is reached in the 10<sup>th</sup> epoch is 95.982% classification accuracy.

It was noticed that after testing the model with 30 images, four mistakes were done, i.e. four images that are from positive covid-19 patients, the model classified them as normal.

TABLE I. TRAINING PROCES FOR THE FIRST PHASE

N. of Epoch	Discussion of the training process	
	Discussion of the training process	Accuracy (%)
1	Trained after 7 iterations	56
2	Trained after 6 iterations	64
3	The accuracy increased, loss decreased	73
4	Loss decreased	85
5	The accuracy increased, loss decreased	91
6	Model stoped increasing accuracy (overfitting)	90
7	Slightly increased accuracy	94
8	Due to the size of the database, overfitting	93
9	The loss increased, and the accuracy is lower than in 7 <sup>th</sup> epoch	93
10	Model reached maximum accuracy	96

During the second phase we noticed that the model does over-fitting again, due to the size of the database, reaching a maximum percentage of accuracy of 98.25%, which is not enough.

The maximum result was reached at the 20<sup>th</sup> epoch of 98.25% accuracy. Table II. shows the whole process of training. There were epochs where the model is overfitting, after 10<sup>th</sup>, 13<sup>th</sup> and after 17<sup>th</sup> epoch. Our supposition is that it was due to the database size. Still, if the database size is big enough, dropout could improve the situation. But, at this time due to difficulties in functioning in hospitals and radiographic departments it was impossible to gather database that has proper size in order to make this research more accurate.

For the second phase after testing the model with 30 images, it made two mistakes, i.e. two images that are from patients negative on covid-19 model predicted as positive on covid-19. It can be seen that the model after 20 epoch training corrects the error of the model that was trained in 10 epochs. It was noticed that the model in the learning process starts with about 50% accuracy and 1.2% loss. After that as the training process continued and as the weights were adjusted the accuracy increased and the loss decreased.

TABLE II. TRAINING PROCESS FOR THE SECOND PHASE

N. of Epoch	Accuracy (%)
1	50
2	71
3	76
4	85
5	87
6	90
7	92
8	93
9	94
10	92
11	93
12	94
13	93
14	95
15	96
16	97
17	95
18	97
19	98
20	98

## VI. CONCLUSIONS AND FURTHER RESEARCH

In the hope that the world will continue to learn more about covid-19 and control the virus effectively, we need to continue to research, and re-examine and only then will science advance again and the scientific community function normally again.

This research was conducted in order to help health professionals to effectively, accurately and fast detect each single positive case on covid-19 using X-ray images. The method that we propose is convolutional neural network as the most effective deep learning technique for machine learning visual patterns recognition. Network was trained on a small database of X-ray images. We can conclude that after training the model with 20 epochs, we got a maximum accuracy of 98 percentage which means that the model has a high degree of accuracy. Nevertheless, further upgrades and increases in the collection of recordings are needed for future training and testing. By applying a larger database, and additional tools and libraries the model training process will improve and testing will show better results. After that the well-designed model could be implemented to use in a real situation.

Combining the model of convolutional neural network for covid-19 detection from X-ray images of the patients

with other data gathered from the medical record of the patients, such as coughing for example, could be one of future steps in our research that will improve the precision and speed up the process of establishing the diagnosis for the patients and simplify the procedures for the medical staff.

Implementing another technique from artificial intelligence combined with convolutional neural network could speed up the process of recovering for the patients and save lives, at the same time, save money, save time and save energy for paying attention to other diseases and patients. We have started analysis of this issues and that will be focus on our future research.

## REFERENCES

- [1] Subramanya, S.H. et al, Impact of COVID-19 pandemic on the scientific community, *Qatar Med J.* 2020; 2020(1): 21.
- [2] Drake N. How the Coronavirus is Hampering Science. *Scientific American.* March 10 2020.
- [3] Malureanu, A. et al., The Relationship between Self-Confidence, Self-Efficacy, Grit, Usefulness, and Ease of Use of eLearning Platforms in Corporate Training during the COVID-19 Pandemic, *Sustainability* 2021, 13, 6633
- [4] Worldometers website, Dadax Limited, accessed 15.09.2021 [https://www.worldometers.info/coronavirus/?utm\\_campaign=homeAdvegas1?](https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1?)
- [5] Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in the diagnosis of COVID-19: challenges and perspectives. *Int J Biol Sci* 2021; 17(6):1581-1587.
- [6] DeGrave, A.J., Janizek, J.D. & Lee, SI. AI for radiographic COVID-19 detection selects shortcuts over signal. *Nat Mach Intell* 3, 610–619 (2021).
- [7] Harmon, S.A., Sanford, T.H., Xu, S. et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat Commun* 11, 4080 (2020).
- [8] Kaheel, H., et al., AI-Based Image Processing for COVID-19 Detection in Chest CT Scan Images, *Front. Comms. Net.*, 09 August 2021.
- [9] Abdul Salam M, Taha S, Ramadan M (2021) COVID-19 detection using federated machine learning. *PLoS ONE* 16(6).
- [10] Brunese, L. et al., Machine learning for coronavirus covid-19 detection from chest x-rays, *Procedia Computer Science*, Volume 176, 2020, pp 2212-2221.
- [11] Yousefi, B., et al. (2021) Impartially Validated Multiple Deep-Chain Models to Detect COVID-19 in Chest X-ray Using Latent Space Radiomics. *Journal of Clinical Medicine*.
- [12] Zhao, W., Jiang, W. & Qiu, X. Deep learning for COVID-19 detection based on CT images. *Sci Rep* 11, 14353 (2021).
- [13] Ghaderzadeh M, Asadi F, Jafari R, Bashash D, Abolghasemi H, Aria M., Deep Convolutional Neural Network–Based Computer-Aided Detection System for COVID-19 Using Multiple Lung Scans: Design and Implementation Study, *J Med Internet Res* 2021.
- [14] Huang, L. et al., Covid-19 classification with deep neural network and belief functions, *Université de technologie de Compiègne, CNRS, Heudiasyc, Compiègne, France*, 2021
- [15] Shorfuzzaman, M., et al., "Artificial Neural Network-Based Deep Learning Model for COVID-19 Patient Detection Using X-Ray Chest Images", *Journal of Healthcare Engineering*, vol. 2021, Article ID 5513679, 16 pages, 2021.
- [16] Reshi, A.A., et al., "An Efficient CNN Model for COVID-19 Disease Detection Based on X-Ray Image Classification", *Complexity*, vol. 2021, Article ID 6621607, 12 pages, 2021.
- [17] Pedro R. A. S. Bassi, Romis Attux, A Deep Convolutional Neural Network for COVID-19 Detection Using Chest X-Rays, *Research on Biomedical Engineering*, Springer Science and Business Media LLC, April, 2021

- [18] Mahesh, P. et al., Covid-19 Detection from Chest X-Ray using Convolution Neural Networks, J. Phys.: Conf. Ser, 2021
- [19] Veljanovska, K., et al., Algorithm for Covid-19 Diagnosis, Department of Intelligent Systems, FICT – Bitola, 2021.
- [20] Maior CBS, Santana JMM, Lins ID, Moura MJC (2021) Convolutional neural network model based on radiological images to support COVID-19 diagnosis: Evaluating database biases. PLoS ONE 16(3).
- [21] Coppock, H., et al., End-to-end convolutional neural network enables COVID-19 detection from breath and cough audio: a pilot study, BMJ Innovations 2021;7:356-362.
- [22] Veljanovska, K. Analysis of Machine Learning Algorithms Performance for Real World Concept, International Journal of Advanced Research in Science and Engineering (IJARSE), Vol. 6, Issue 12, 2017.
- [23] Veljanovska, K., Machine Learning Algorithms Comparison, International Journal of Engineering and Technical Research, Vol. 7, Issue 11, 2017.
- [24] Mnih, V., et al., Human-level control through deep reinforcement learning. Nature. 2015; 518:529–533. pmid:25719670
- [25] Goodfellow, I., et al., Deep Learning, MIT Press, 2016, p. 326.

# Web GUI Upgrade for Manual Semantic Extraction i E-Learning

M. Jovanovic

University of Niš, Faculty of Electronic Engineering, Niš, Serbia  
martin.jovanovic@elfak.ni.ac.rs

**Abstract** - This paper presents the new structural design on the e-learning aid system currently in development at Computer Science Department of the University of Niš, Faculty of Electronic Engineering. The system enables the student to query the lecture in plain or rich text format for any existing relations between keywords, thus reducing the learning time and eliminating the student's need to periodically back up in order to review definitions of terms in order to infer their mutual relation. This is achieved through the graphical user interface, by the drag and drop operation performed on the chosen keywords, which triggers the query on the database of relations. The database can be developed manually or inferred based on any existing set of relations, and any database of relations can be applied to any chosen textual document, thus enabling cross-domain application. The paper describes the current state of development of the system, as well as possible directions for further research. Research presented supported by the University of Niš.

**Keywords:** e-learning, manual semantic extraction, aid system

# **REGULAR PAPERS**

# An overview of 4D medical image compression

Maja Gaborov\*, Marko Blažić\*, Dijana Karuović\*, Mila Kavalić\*\*, Igor Vecštejn\*, Dragana Milosavljev\*

\* University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia

\*\* University of Novi Sad, Faculty of Technical Science, Novi Sad, Serbia

majagaborov26@gmail.com, markoblastic93@gmail.com, dijanakaruovic@gmail.com, milazakin@gmail.com,  
igor.vecstejn@gmail.com, dragana.milosavljev@tfzr.rs

**Abstract** – As a result of using newest technical opportunities in contemporary medical diagnostics, there is need for compression of medical images that requires special approach. In order to achieve to best possible archiving and transporting of medical data, special algorithms have been developed. Modern medicine cannot be imagined without utilization of informational technologies. In this paper, 4D compression technique are analyzed based on the reviewed literature. The main goal was to determine where 4D compression is applied. 4D compression has been found to be used for diagnosis, including 4D MRI, 4D CT and PET scans.

**Keywords:** 4D compression, medical image, lossy, lossless, MRI, PET, CT, SPECT

## I. INTRODUCTION

With the increasing use of medical imaging in clinical diagnosis, there is a need for a fast and safe method for sharing a large number of medical imaging among healthcare professionals [1]. Medical imaging is playing an increasingly important role in modern healthcare where is used for noninvasive diagnostics, therapy planning, and follow-up evaluations. The most updated imaging technologies offer enhanced features including additional levels of morphological and anatomical details as well as functional mapping. Among these advanced technologies, 4D imaging is experiencing growing popularity for its capability of producing a three-dimensional view of the object while at the same time monitoring the object over time (time is the fourth dimension)[2]. The lossless approach ensures that the output image (reconstructed) to be equivalent to the original input image after performing compression. Unfortunately, the scheme can only attain minimal compression. Even in lossy approach, the output image possess certain level of degradation as compared to original input image. The prime reason for this is because the redundancies lying within the image is rejected by the compression technique. Fortunately, lossy technique can accomplish higher degress of compression along with better retention of perceptual quality compared to lossless schemes[3].

Four-dimensional (4D) medical images are increasingly used in diagnosis, including 4D MRI, 4D CT, and PET scans [4][5]. Without efficient compression, large amounts of data would easily overwhelm storage and transmission systems. Doctors want lossless compression for accurate diagnosis and treatment, as well as in accordance with legal and

regulatory requirements. Although there has been extensive work on the compression of still (2D) and volume (3D) images without loss and loss, the problem of 4D image compression is a relatively new research area [4].

This paper analyzes the 4D compression technique. The aim of this paper is to show what the 4D compression technique is used for and to explore all the potential uses of 4D compression.

## II. BACKGROUND OF MEDICAL IMAGES

All modern medical imaging devices (imaging modalities) such as X-ray, CT (computed tomography) scan, and MRI (magnetic resonance imaging) use DICOM. The DICOM standard (Digital Imaging and Communications in Medicine) is a set of rules that allows trading of medical images and information between hospitals and computers [6][7]. Image archiving and communication systems (PACS) enable archiving, review and distribution of medical images. It is a system used for manipulation of medical images and information, communicates with hospital information systems, radiological information systems, departmental information systems etc[8]. Detecting and analysing the diseases has become Significant role played by the medical imaging devices in detecting, analyzing the diseases. Many medical imaging equipment's like X-Ray, CT SCAN, PET SCAN, SPECT etc., are in the market by different vendors like GE, SIEMENS, PHILIPS etc., Many of the equipment's are 2D and even 3D/4D. All medical modalities have built in native image file format usually DICOM [9]. Based on the dimensionality medical images are classified as 2D, 3D, 4D. 2D medical images representation of object in 2 axis i.e. X axis and Y axis. 3D medical images representation of object in 3 axis i.e. X, Y and Z axis. 4D medical images represents 3D and time series [9].

### A. CT (COMPUTED TOMOGRAPHY)

CT scanning uses special X-ray equipment to take multiple images from different angles around the body. The computer processes the information from the images and creates an image that shows a cross section of the area being examined. CT scanning is mainly suitable for studying parts of the chest and abdomen of the human body. Four-dimensional computed tomography is a dynamic system for imaging the

extent of moving organs with image quality comparable to conventional CT [9].

#### *B. MRI (MAGNETIC RESONANCE IMAGING)*

An MRI machine or scanner uses a powerful magnet and radio waves linked to a computer to create remarkably clear and detailed cross sectional images of the body. MRI can be used to view, monitor and diagnose spinal, joint or muscle problem, abdominal tumors and disorders, brain tumors and abnormal breast cancer, heart or blood vessel problems. 4D MRI can be used to visualization of complex flow patterns associated with healthy and pathologic of blood movement[9].

#### *C. PET (POSITRON EMISSION TOMOGRAPHY)*

Scan uses radiation or nuclear medicine imaging, to produce 3 dimensional, color images of the functional processes within the human body. It is mainly used in diagnosing epilepsy cancer. 4D PET scans and outlined a PET imaging technique that reduces image blurring arising from respiratory motion[9].

#### *D. SPECT (SINGLE PHOTON EMISSION COMPUTED TOMOGRAPHY)*

The scan provides highly functional information about the patient's specific organ or body system. SPECT is one of the best tools used to diagnose the brain, heart, kidneys, etc., and also to detect tumors in the body, because it provides information about blood flow, body temperature, etc[9].

### III. APPLICATION OF 4D MEDICAL IMAGE

4D imaging is finding application in various branches of medicine, the most common being oncology, cardiology, and neurology, and is contributing to understand disease progression, locate abnormalities, fine-tune radiation dosage, assess therapy efficacy, and monitor disease remission or spread. 4D imaging is also becoming a useful tool in image-guided surgery (IGS). Also known as surgical navigation, IGS is a method for performing a surgery with the aid of one or more imaging systems. In IGS, typically, 4D imaging is utilized to obtain preoperative images. Preoperative images show a view of the patient's anatomy and are combined to generate a digital map, which is used by the surgeon to precisely position and orientate his surgical tools. Lately, 4D imaging systems for intraoperative imaging have also been introduced. These instruments are for the most part multimodal systems that not only provide images but also supply a therapeutic agent [2].

As technology continues to advance and evolve, innovators are constantly creating new ways to detect, diagnose, and treat disease with the help of medical imaging technology-including 3D and 4D models that push the frontiers of medical science. Advancements in 3D and 4D imaging technology facilitate efficient and accurate real-time visualization of the human body while minimizing distortion for the patient [10]. Advanced of 4D imaging method makes image analysis much faster and more accurate than ever [11].

Dynamic volumetric (four dimensional- 4D) medical images are typically huge in file size and require a vast amount of resources for storage and transmission purposes[12]. Dynamic medical imaging modalities enable the examination of functional and mechanical properties of the human body and are used for clinical applications, e.g., 4D CT for respiratory organ motion modelling, 4D MR imaging for functional heart analysis and 4D ultrasound (US) for echocardiography analysis. These 4D modalities have high spatial (volumetric) and temporal (time sequence) sampling rate to capture the periodic motion cycles of organ activities, and this information is used for clinical decision making. However, the acquisition of these dynamic images requires larger radiation doses which may cause harm to humans, and longer image scanning and reconstruction times; these factors limit the use of 4D imaging modalities to broader clinical applications[13]. 4D medical images consist of 3D images (volumes) captured at different time points, with each volume comprised of 2D images (slices). These data are usually huge in file size and pose a big burden on the resources needed to store them for future study and follow up. With the increasing use of telemedicine and the picture archiving and communications system (PACS), there is also a need to quickly transmit these data over limited band width channels. Hence, it has become essential to design efficient lossless compression methods for storage and transmission of 4D medical images. Since lossless compression of 4D medical images is still a relatively new area of research, 3D and 2D lossless compression algorithms are often used to compress volumes or slices independently. Current state-of-the-art 2D and 3D compression methods use mainly wavelet transforms or prediction coding to decorrelate the data and improve the compression performance. However, these compression methods fail to exploit redundancies in all four dimensions. Few methods that exploit redundancies in all four dimensions have been proposed using either 4D wavelet transforms or 3D motion compensation. However, the lossless compression ratios achieved by these approaches remain comparable to those achieved by state-of-the-art 3D compression techniques, such as 3D-JPEG2000, thus leaving room for much needed improvement [12].

Recent advances in computer science and image technology have led to the innovative development of cardiovascular imaging. Recent three-dimensional volumetric computed tomography (CT) has increased the diagnostic accuracy of aortic and coronary artery disease, and scintigraphic imaging on color maps provides functional information on ventricular muscle viability, and computer-controlled film films are much more convenient than use. film images. Blood flow imaging is a new trend in cardiovascular imaging. Two types of blood flow imaging tools are available today: measurement-based flow visualization, including 4D flow magnetic resonance imaging, or echocardiography flow visualization software and computer flow simulation based on CF dynamics). MRI and echocardiography flow visualization provides measured blood flow, but has limitations in temporal

and spatial resolution, while CFD flow calculates flow according to assumptions instead of flow measurement and has fine enough resolution to limit computer memory, and allows even virtual surgery combined with computer graphics. Blood flow imaging can provide useful information for the development of predictive medicine in cardiovascular diseases and can lead to discoveries in cardiovascular surgery, especially in the decision-making process. Tools based on flow measurement are 4D magnetic resonance imaging and VFM echocardiography, both of which have limitations in temporal and spatial resolutions, but can visualize complicated flow by moving the heart chamber or valve. CFD simulation is a method of flow calculation with sufficiently fine resolution, but it depends on the calculation assumption. CFD simulation enables virtual surgery on a computer in combination with computer graphics. Blood flow imaging provides several predictive parameters of mechanical stress caused by diseased flow. Choosing the appropriate way to visualize the flow for each characteristic clinical situation is essential for the efficient use of blood flow images [14].

External beam radiation therapy (RT) is one of the main cancer therapies for lung cancer. Breathing and heart motion during irradiation causes significant variations in organ and target geometry in the order of several centimeters. This increases the dose to healthy tissue and reduces the dose to the target area, impairing the balance between complications and cure. 4D RT aims at compensating the deformation uncertainty by incorporating the motion characteristics into the dose calculation or gating the treatment device in phase with the motion pattern. These techniques require a patient-specific motion model. With the advent of multi-slice CT, 4D image acquisition of dynamic processes such as breathing is now becoming possible [15].

The new 4D image reconstruction methods incorporate additional models of the respiratory and cardiac motion of the patient to reduce image blurring due to respiratory motion and image noise of the cardiac-gated frames of the 4D cardiac-gated images. They describe respiratory motion estimation and gating method based on patient PET list-mode data. The estimated respiratory motion is applied to the respiratory gated data to reduce respiratory motion blur. The gated cardiac images derived from the list-model data are used to estimate cardiac motion. They are then used in the cardiac-gated images summing the motion-transformed cardiac-gated images for significant reduction in the gated images noise. Dual respiratory and cardiac motion compensation is achieved by combining the respiratory and cardiac motion compensation steps. The results are further significant improvements of the 4D gated cardiac PET images. The much improved gated cardiac PET image quality increases the visibility of anatomical details of the heart, which can be explored to provide more accurate estimation of the cardiac motion vector field and cardiac contractility [16]. The development of quantitative image reconstruction in medical imaging, including emission computed tomography (ECT) and

x-ray CT, has recently shifted from three-dimensional to four-dimensional, i.e., the inclusion of the time dimension. There are two major goals for this development. First is to reduce reconstructed image artifacts due to patient motion. In particular, compensation of involuntary patient motion, e.g., respiratory motion, that causes resolution loss has received much attention. Second is to improve the temporal resolution of dynamic images for improved detection of global and regional motion abnormalities [16].

In recent years, development in non-invasive and painless medical imaging such as CT or MRI, has improved the process of diseases diagnosis and clarification, including tumours, cysts, injuries, and cancers. The fullbody scanner with superior spatial resolution provides essential details of complicated anatomical structures for effective diagnostics. However, it is challenging for a physician in glance over a large data-set of over hundreds or even thousands of images (2D "slices of the body"). When a doctor wants to view a patient's CT or MRI scans for analysing, he needs to review and compare among many layers of 2D image stacks (many 2D slices make a 3D stack). If the patient is scanned multiple time (three consecutive months, for instance) to confirm the growth of the tumours, the dataset is turned to be 4D (time-stamp added). The manual analysing process is time-consuming, troublesome and labour intensive. The innovation of Augmented Reality (AR) in the last few decades allows to illuminate this problem. In one study, authors propose an AR technique which assists the doctor in instantly accessing and viewing a patients set of medical images quickly and easily. The doctor can use an optical head-mounted display such as the Google Glass, or a VR Headset such as the Samsung Gear VR, or a general smartphone such as the Apple iPhone X. He looks at one palm-sized AR Tag with patients document embedded with a QR code, and the smart device could detect and download the patients data using the decrypted QR code and display layers of CT or MRI images right on top of the AR tag. Looking in and out from the tag allows the doctor to see the above or below of the current viewing layer. Shifting the looking orientation left or right allows the doctor to see the same layer of images but in different timestamp (e.g. previous or next monthly scans). Their obtained results demonstrated that this technique enhances the diagnosing process, save cost and time for medical practice [17].

Medical image data (Ultrasonography, Computed Tomography, Magnetic Resonance Imaging etc.) consumes maximum storage and utilize maximum bandwidth for transmission that often results in degradation of image quality. Due to these inherent issues in such type of images, compression is the only applicable technique explored [3].

#### IV. 4D ALGORITHMS

In the text that follows 4D algorithm in medicine will be presented.

##### A. 4D compression

This model is presented as one of the possible solutions for the problem, a large amount data that cannot be stored so easily. This model can be divided into two main parts: exploring temporal redundancies and exploring frequency redundancies. Analyzing the shape of the original 4D dataset based on edge detection is the purpose of temporal analysis. Kernel then decides how the dataset is going to be segmented by the contours and characteristics. CNN convolutional neural networks) enables every segment subvolume, to have their own motion field, and every subvolume to be independent from others. Frequency redundancies, are removed after finishing with temporal redundancies, using 3D wavelet transform[18].

##### B. Temporal redundancies

For temporal redundancies, time analysis is of the great importance. Time analysis requires registration of object and static analysis. Static analysis is based on dividing data into time frames that can be observed how each frame is trying to match shape that is presented in kernel. In real time, every object is moving different motion in the space and time, so it is important to look on every object separately[7].

##### C. 3D canny edge detectorbased segmentation

3D canny edge detector can be used for detecting edges. “The edge gradient of volumetric object can be by gradient components for each voxel. Second step is to estimate edge strength with the orientation of edge normal”[18].

##### D. Surface Extraction

Every 3D object can be described based on shape of his surface. This description are used for classification object and movement predictions while compressing. HK segmentation is well-known method that allows finding patches of various shape for objects that as a result is giving visible object surface [18]. HK method is dividing object into homogeneous regions that are called homogenous surface patches. “For a given range object, the goal is to compute a new object registered with and with the same size in which each voxel is associated with a local shape class selected from a given dictionary” [18]. To complete the job, two tools are required: the dictionary of shape classes and the algorithm, that will determine which class give the best approximation of the surface for every voxel separately[ 18].

##### E. Estimating the motion field

“In estimating the motion field of segmented subvolume the basic assumption is that the motion field is well approximated by a constant vector field within any small region of the object plane. The optical flow is the approximation of the motion field which can be computed from time-varying volumetric sequences [18].”

##### F. Motion parameters for neural network

The gold is to recreate motion of object, but that requires tie analysis to synthesis. Semantics of the system that will do visualization the motion needs to be known. Action and parameters should be linked to the 3D model, in order to recreate the motion. This model converts result into visual parameters that will build motion field, which is previously mentioned[18].

##### G. Frequency redundancies

Fast wavelet transform is used to explore frequency redundancies of still 3D object. Using wavelet transformation, the original data is decomposed on low and high coefficient frequency. One-dimensional wavelet is used on every row, column and in depth separately, for each voxel values. As result, average value is presented with coefficient[18].

##### H. Fast 3D wavelet transform

“The fast wavelet transform is computed with filters that are separable products of one-dimensional low-pass  $h[n]$  and high-pass  $g[n]$  filters”[18].

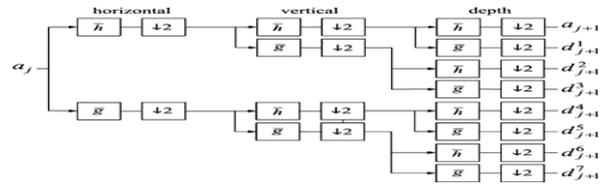


Fig 1. Three-dimensional decomposition and subsampling along rows, columns and depth[18]

#### V. DISCUSSION

In one study authors extend initial ideas presented in by proposing a novel lossless compression method that fully exploits redundancies in all dimensions of 4D medical images. Their new method decorrelates the data in the spatial and temporal dimensions by recursively applying a multi-frame motion compensation process that employs a full 4D search with variable block-sizes and bi directional prediction. Redundancies in the resulting motion vectors are also reduced using a novel differential coding algorithm, while the residual data is losslessly compressed using a new context based adaptive binary arithmetic coder. Performance evaluations show that proposed method provides a significant improvement in compression ratio compared to current stat-of-the-art such as JPEG2000, 3D-JPEG2000 and H.264/AVC[19].

In second study it is presented a lossless compression technique for 4D medical images which is based on the H.264/AVC video coding standard. This compression technique discover efficiently temporal and spatial redundancies between 2D image slices and 3D images in 4D medical images and eliminates any concerns that can harm the quality of the image while compressing the same. Result of this technique shows the performance that is better for 70% percent from current 4D compression methods. This paper also shows that this compression technique is most valuable for compressing images from big scans such as MRI, PET, CT etc. In the following paper written by Hui

Yun et al, it is presented another solution in form of MP4 video coder for compression of 4D images in medicine. The size of 4D-CBCT is larger with every day and that became a burden to hard drive memory space of the hospital. For example, lung cancer patient with daily 4D-CBCT scan (33 fractions, 3-mm slice thickness, and 10 breathing phases), the total size of 4DCBCT images in the whole treatment courses would be approximately 8 GB. For a cancer center treated with 300 patients per day, if only 10% of patients were using 4D-CBCT scan, the total size of images would be 240 GB per day and approximately 5 TB per month. To transfer and backup such huge amount of 4D-CBCT data would be a challenge for a busy clinic. MP4 shown better result from MJ2 and AVI video coders, the size of the file that is compressed can be reduce to the 99% percent from the original size [20].

Recent work of 4D lossless medical image compression is based on the application of techniques derived from video compression to effectively eliminate redundant parts in different image dimensions. In this context, some authors presented a new approach to 4D lossless medical image compression consisting of applying 2D wave transformation in spatial directions, followed or not by lifting the transformation or compensating for movement in the direction between sections, and the resulting sections are encoded by 3D SPIHT. Their approach was compared to 3D SPIHT with / without motion compensation. The results show that their approach offers better lossless compression performance [21].

Diagnostic imaging, especially brain imaging today widely developed involving more and more modalities. The size of the image data set then becomes very large. This is not only a problem due to problems with storage space, but also due to reliable data communication as in cloud architecture. The author's goal was to develop a new method called sequential -preserving differences to compress a set of brain image data using redundancy in 4D format. They tested their method with actual clinical patient data produced by CT Perfusion and measured the compression ratio as well as the strength of the compression rate to assess performance. With 10 data sets, they have an average compression rate of 0.53 and space savings of more than 47%. This method is lossless, and the degree of compression was acceptable, which makes it suitable for application in cloud architecture [22].

## VI. CONCLUSION

4D image compression is a relatively new research area. 4D data is commonly used for compression of medical images that are recorded for the patients with the cancer. By this it was meant the images from the big scans such as MRI, PET, CT. This data compression is very simple, choosing folder with the 3D images that we want to transform into 4D and the folder where we want to save it. Since 4D compression is a newer area, the research could be conducted in a few years when this data compression technique is more applied. Currently, there exists abundant

research works on medical image compression considering lossy and lossless types, but the need of medical images to be compressed efficiently with optimal compression ratio is yet a question mark.

## REFERENCES

- [1] Rahmat, R.F., Andreas, T. S. M., Fahmi, F., Pasha, M.F., Alzahrani, M.Y., & Budiarto, R. (2019). Analysis of DICOM Image Compression Alternative Using Huffman Coding. *Journal of Healthcare Engineering*, 2019. <https://doi.org/10.1155/2019/5810540>
- [2] Research and markets – The world's largest market research store. (2019). *4D Imaging in Healthcare*. Link available: <https://www.researchandmarkets.com/reports/4876630/4d-imaging-in-healthcare>. Accessed on 29.06.2021.
- [3] Bhagat, A., & Atique, M. (2012). Medical images: Formats, compression techniques and DICOM image retrieval a survey. *2012 International Conference on Devices, Circuits and Systems (ICDCS)*, 172-176.
- [4] Yan, P., & Kassim, A. (2004). Lossless and near-lossless motion-compensated 4D medical image compression. *IEEE International Workshop on Biomedical Circuits and Systems*, 2-13.
- [5] Kanaga, E. G. M., Anitha, J., & Juliet, D. S. (2021). 4D medical image analysis: a systematic study on applications, challenges, and future research directions. *Advanced Machine Vision Paradigms for Medical Image Analysis*, 97-130. <https://doi.org/10.1016/B978-0-12-819295-5.00004-4>
- [6] Prvi Telemedicinski sajt Srbije. (n.d.). *DICOM standard*. Link available: <http://www.telemed.rs/dicom-standard/>. Accessed on 29.06.2021.
- [7] DICOM – Digital Imaging and Communications in Medicine. Link available: <http://medical.org>. Accessed on 29.06.2021.
- [8] Trkuljić, N. (2011). IT standards in medicine. *INFO M, Časopis za informacione tehnologije i multimedijalne sisteme*, 38/2011.
- [9] Girija, J., Murthy, G. K., & Reddy, P. C. (2017). 4D medical image registration: A survey. *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, 539-547. [10.1109/ISSI.2017.8389230](https://doi.org/10.1109/ISSI.2017.8389230)
- [10] MITA: Medical Imaging & Technology Alliance (n.d.). *3D, 4D & Beyond*. Link available: <https://www.medicalimaging.org/medical-imaging/principle-details/3d-4d-and-beyond>. Accessed on 29.06.2021.
- [11] Parisot, C. (1995). The DICOM standard. *The International Journal of Cardiac Imaging*, 11(3), 171-177. <https://doi.org/10.1007/BF01143137>
- [12] Sanchez, V., Nasiopoulos, P., & Abugharbieh, R. (2008). Efficient 4D motion compensated lossless compression of dynamic volumetric medical image data. *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 549-552. [10.1109/ICASSP.2008.4517668](https://doi.org/10.1109/ICASSP.2008.4517668)
- [13] Guo, Y., Bi, L., Ahn, E., Feng, D., Wang, Q., & Kim, J. (2020). A Spatiotemporal Volumetric Interpolation Network for 4D Dynamic Medical Image. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4726-4735.
- [14] Itatani, K., Miyazaki, S., Furusawa, T., Numata, S., Yamazaki, S., Morimoto, K., Makino, R., Morichi, H., Nishino, T., & Yaku, H. (2017). New imaging tools in cardiovascular medicine: computational fluid dynamics and 4D flow MRI. *General thoracic and cardiovascular surgery*, 65(11), 611–621. <https://doi.org/10.1007/s11748-017-0834-5>
- [15] Kaus, M., Netsch, T., Kabus, S., Pekar, V., McNutt, T., & Fischer, B. (2004). Estimation of Organ Motion from 4D CT for 4D Radiation Therapy Planning of Lung Cancer. *MICCAI*, 1017-1024. [10.1007/978-3-540-30136-3\\_123](https://doi.org/10.1007/978-3-540-30136-3_123)
- [16] Tsui, B. M., Feng, T., Wang, J., Xu, J., Abraham, M. R., Zimmerman, S. L., & Schindler, T. H. (2016). Advances in 4D gated cardiac PET imaging for image quality improvement and cardiac motion and contractility estimation. *Perspectives on*

- Nuclear Medicine for molecular diagnosis and integrated therapy*, 3-16. [https://doi.org/10.1007/978-4-431-55894-1\\_1](https://doi.org/10.1007/978-4-431-55894-1_1)
- [17] Le, H., Nguyen, M., & Yan, W. Q. (2020). A Web-Based Augmented Reality Approach to Instantly View and Display 4D Medical Images. *Asian Conference on Pattern Recognition ACPR 2019: Pattern Recognition*, 691-704. [https://doi.org/10.1007/978-3-030-41299-9\\_54](https://doi.org/10.1007/978-3-030-41299-9_54)
- [18] Žagar, M., Kovač, M., & Hofman, D. (2012). Framework for 4D medical data compression. *Tehnički Vjesnik*, 19(1), 99-105.
- [19] Sanchez, V., Nasiopoulos, P., & Abugharbieh, R. (2006). Lossless Compression of 4D Medical Images using H.264/AVC. *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, II-II, doi: 10.1109/ICASSP.2006.1660543
- [20] Yan, H., Li, Y. & Dai, J. (2020). Four-Dimensional Cone-Beam Computed Tomography Image Compression Using Video Encoder for Radiotherapy. *Journal of Digital Imaging* 33, 1292–1300. <https://doi.org/10.1007/s10278-020-00363-9>
- [21] Belhadeh, L., & Maaza, Z. M. (2018). Lossless 4D Medical Images Compression Using Adaptive Inter Slices Filtering. *International Journal of Advances in Applied Sciences (IJAAS)*, 7(4), 361-368. <http://doi.org/10.11591/ijaas.v7.i4.pp361-368>
- [22] Fahmi, F., Sagala, M.A., Nasution, T., & Anggraeny (2016). Sequential — Storage of differences approach in medical image data compression for brain image dataset. *2016 International Seminar on Application for Technology of Information and Communication (ISEMANTIC)*, 122-125. [10.1109/ISEMANTIC.2016.7873822](https://doi.org/10.1109/ISEMANTIC.2016.7873822)

# Formula One Lap Time Data Visualization and Prediction Software

F. Đorđević\*, D. Ivetić\*\*

\* Faculty of Technical Sciences/Department of Power Electronic and Telecommunication Engineering, University of Novi Sad, Serbia

\*\* Faculty of Technical Sciences /Department of Computing and Control Engineering, University of Novi Sad, Serbia  
filip.djordjevic@uns.ac.rs, ivetic@uns.ac.rs

**Abstract** – In this paper we present the implemented software solution to visualize lap time data of Formula One Drivers on a chosen Circuit using Windows Presentation Foundation (WPF) Framework. We introduce three ways of comparing and visualizing Formula One lap time data, as well as a simple lap time prediction algorithm based on selected driver's previous performance on the chosen circuit, with its results also visualized to be compared with actual race lap times.

## I. INTRODUCTION

Formula One is a racing World Championship established in 1950. Today, ten race teams (constructors) with two cars and drivers each compete during a, on average, twenty race (Grand Prix) long season for the Constructor's and the Driver's Championship.

Drivers are awarded championship points depending on their finishing positions in a race, with the 10 best placed drivers score points, and each team gets the number of points which is the total of what their drivers scored in that particular race.

The sport itself is governed by the Fédération Internationale de l'Automobile (FIA), that decides the regulations, both sporting and technical for its single seater open wheel and open cockpit racing cars.

Formula One teams and drivers compete over a race weekend which consists of two hour-long free practice sessions on Friday, one hour long free practice and one qualifying session on Saturday and a race on Sunday. During a free practice session, the racing circuit is free for the drivers to post lap times and set up the cars to their liking, but in accordance with circuit configuration and weather conditions.

With every outing of a single car on the circuit, the teams gather large amounts of data for almost every car component, and next to that, every car is fitted with a transponder in order to follow the car's position as well as the time it takes for the car to complete a lap.

Using this data, the engineers of Formula One teams can get an understanding of their driver's pace during a stint (a set of consecutive laps driven) and how the different sets of tires perform. These are the factors at the center of the Race Strategy, which is the team's plan on what lap their driver should change tires. All this data can be visualized in different ways, with one of them being a lap chart, displayed in Figure 1.

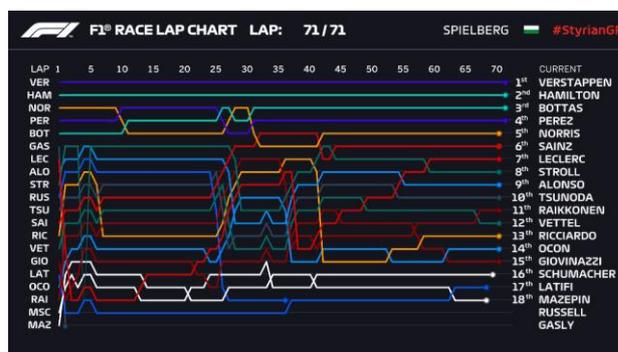


Figure 1. Official Formula One lap chart for the 2021 Styrian Grand Prix

## II. BACKGROUND AND RELATED WORK

### A. Sport Data Visualization

As with Formula One, other sports have a long tradition of data collection and reporting. Sport data is becoming more and more interesting for information visualization, visual analytics and data science, since through it, static and dynamic patterns can be identified [1].

Visualization can be characterized as a technique of creating images, diagrams or animations to convey a message. Data Visualization is a sub-category of visualization that enables graphic representation of data, utilizing charts, graphs or maps. Users can see, recognize and quickly recollect images, but also detect changes in color, size, shapes, movement etc.

### B. Related Work

In [1], Perin et al. identified three types of sports data that are used in visualization:

- *box-score* designates the statistical summary of a game. The authors use this term to describe any discrete data referencing in-game events. These include summary statistics, and finer-grained data. This type of data is often simple and small-scale, but it is diverse and results in a wide variety of visualization approaches. The most interesting approach for this paper with *box-score* data is showing relative differences, with relative times being shown. Charts

presented here were line, column and step line charts.

- *tracking data* is collected utilizing machine vision to gather precise spatio-temporal information about the players and equipment in real-time during the course of play. Visualization of this data type follows trajectories of player movement or density and areas that a player or a team occupied the most on the field.
- *meta-data* adds context to the events that the previous two capture. These concern rules, stadiums, physical characteristics of players, kit colors, team badges etc. Visualization can add championship logo, team colors, player nationalities, tournament brackets, predictions etc.

The role that visualization finds in sports is to be of use for analytical purposes, in order to get better understanding and gain insight about the data, and for narrative purposes, to present data in a way that is much easier to understand.

Sulsters presented a way to simulate Formula One race strategies in [2]. They built a simulation model that can determine the optimal race strategy utilizing the same lap time data source as the used for this paper. Their model uses discrete-event simulation to simulate lap times of drivers during a race, with imitation of other on-circuit events that affect lap times. The simulation model was better with predicting the end positions and worse when it came to predicting race times. Results of the simulation were presented with overall race results and total race time, and were not visualized in any way.

Interactive, calendar-based web visualization tool, that was the result of [3] has a similar idea to our visualization software with the same data source. Lamprecht et al. presented races in a heatmap depending on the fastest lap driven, and by selecting a race, user will see two charts with single color-coding approach, faster the lap or better the finishing position is, the color chosen is dark green, while on the opposite end the color chosen is red, which means that not all drivers will ever be presented with the same color. There are options available to isolate specific drivers' lap representation and zoom to a specific part of the data displayed.

Visualization of player performance in the *Overwatch* video game gave another data visualization approach in [4]. For every aspect of the game, Braun et al. arranged data clusters along the X-Axis, while the values were arranged along the Y-Axis. Every value presented on the chart, was marked with a „X“ symbol, while the average value was marked with a horizontal line that is parallel to the X-Axis, and in turn serves as a referent value. Average value of every data cluster was marked with a „+“ symbol which gives an opportunity of comparing with the total average value, and to compare data between the clusters.

Skau and Kosara, in [5] researched how people perceive and read pie and donut charts, and give

recommendations how to use these types of charts. One of their discoveries is that the central angle is not the only, nor the primary factor when reading pie charts. Also, the center can be removed and not affect the precision of the chart reading. On the other hand, area-only charts resulted as being surprisingly effective. When approaching the radius of the pie chart, changing it interfered with people's ability to read the chart, and the authors call for this to be avoided.

In [6], Ottley, Kaszowska, Crouser and Peck utilized eye tracking to see how people extract information from text and visualization. Visualization was better for information retrieval, while textual data got more attention from people. Neither visualization nor text alone was effective in aiding information retrieval and processing. One way that the authors found could be the way to bridge the gap is to employ both representations in an interactive visualization, and take advantage of what text and visualization provide.

### III. MATERIALS, METHODS AND PROPOSED SOLUTION

#### A. Data

The Formula One data is provided by an internet API, [ergast.com/mrd](https://ergast.com/mrd). It provides the user with both *box-score* and *meta-data* that consists of both drivers, teams and circuits data, and also the subject of this paper, lap times and other race result data.

API queries require a GET request with arguments that define what data will be retrieved. There is a predetermined caching period, which lasts one hour from the first retrieval of a specific dataset, but the caching time can be changed.

Ergast API cannot take more than four queries per second and 200 queries per hour, which limits data availability to some degree, but we took that into consideration when approaching the interaction and visualization options.

Lap time data, when being used for processing, is represented in seconds, while when it is being represented in textual form is shown in *mm:ss.sss* format. When presenting lap times of a specific driver, the chart is colored in the color of the driver's team.

#### B. Visualization Options

The software is made using the Windows Presentation Foundation (WPF), where we organized our software into five windows.

Starting (Main) Window offers to the user to choose the data to visualize: lap times of a driver in a single race, compare the lap times of a driver in a race in two different seasons, or compare the lap times of two drivers in a race.

The user selects the season (2019, 2020 or 2021), the race name, and driver name depending on the data he wants to visualize, and gets the resulting chart.

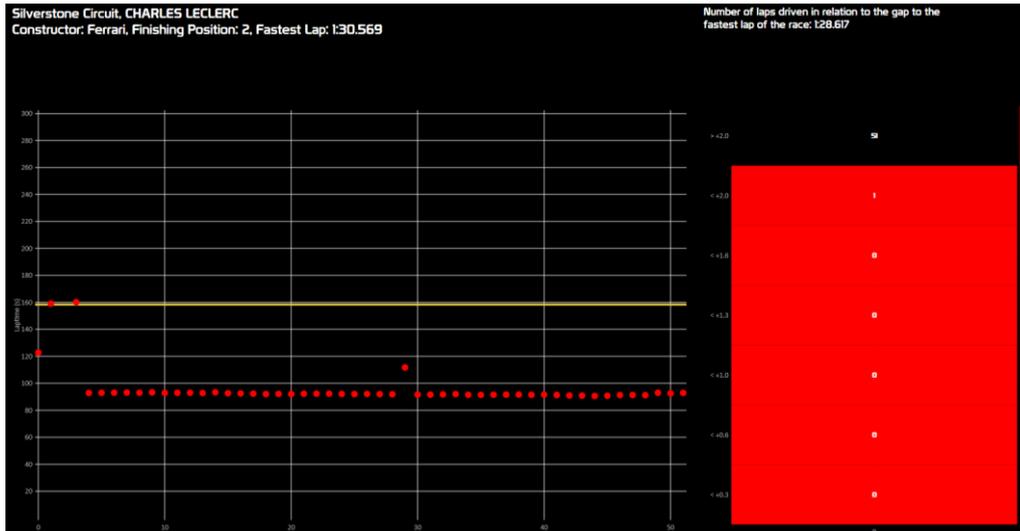


Figure 2. Visualization of lap times of a driver in a single race

Visualization of lap times of driver in a single race is represented with a point chart and a heatmap. The point chart represents all of the lap times in a race and it is shown with a vertical line that corresponds to the time of the second lap, which is the first competitively posted lap time in a race after lap one, where drivers do a standing start, so the user can get an understanding of the driver's race pace. This solution is inspired by the results presented in [4].

Circuit and driver name are also shown, along with driver's team name, finishing position in the race and the fastest lap time the driver posted during the race. The heatmap on the side shows how many laps did the chosen driver drive with a specific gap to the overall fastest time of the race, where the gaps are: up to 0.3 seconds, from 0.3 to 0.6 seconds, from 0.6 to 1 second, from 1 to 1.3 seconds, from 1.3 to 1.6 seconds, from 1.6 to 2 seconds or more than 2 seconds. This visualization solution is displayed in Figure 2.

Visualization of the comparison of the chosen driver's lap times in a race in two different seasons is represented as a column chart with two bars being shown for each lap. Since the chart is showing multiple data for a single driver, to escape using the same color twice, the second bar is shown half transparent.

The same text data is shown as in the previous visualization solution with options to zoom and move through the zoomed-in chart. Another part of this visualization solution is a pie chart variant that shows total race time for the driver in both seasons. This visualization

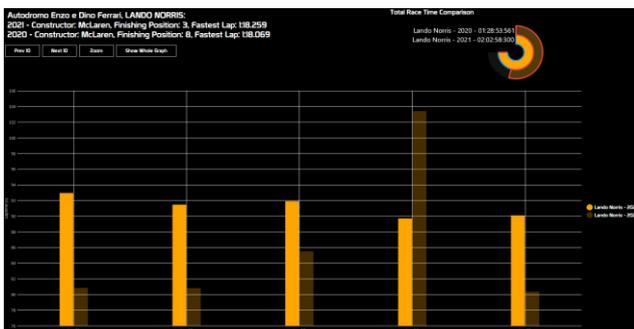


Figure 3. Visualization of lap times comparison for a chosen driver in a chosen race in two different seasons

solution is displayed in Figure 3.

Visualization of the comparison between the lap times of two drivers in the same race is shown with the race and circuit name, along with relevant race data for both drivers: their names and team colors, gap between the first and the second driver chosen, their respective fastest lap of the race, and the average speed of the fastest lap displayed with a speed-o-meter graphic and text label with the time itself, a solution that was proposed in [6].

Next to that, there is a Nightingale Rose Chart (pie chart with varying radii), which represents drivers' lap times with the slower lap time showing with the maximum radius, while the faster lap time is showing with a radius that goes up to the percentage of the gap between the laps in relation to the slower lap time.

With this type of chart not having Data labels, starting point of the chart is displayed in form of an arrow, that shows where the times for the first lap are shown. This visualization solution is displayed in Figure 4. Even though in [5], varying radii pie charts were presented as a bad solution, here it can be used to show which driver had better pace overall.

All of the described charts have tooltips with data values. When the user moves his mouse pointer over a particular chart area, they will be presented with a tooltip that shows the chart value.

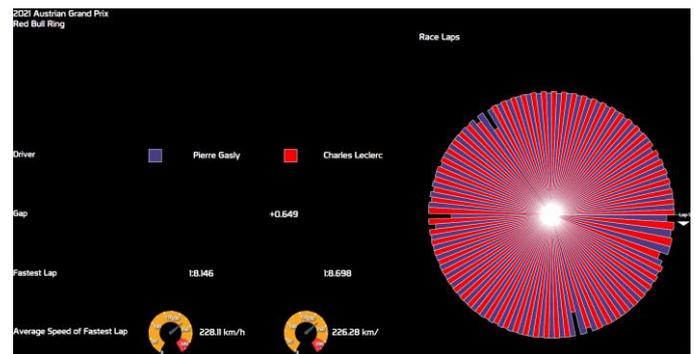


Figure 4. Visualization of lap times comparison of two drivers for a chosen race

### C. Prediction Algorithm

Lap time prediction works in a similar way to visualization, it can be chosen as an option in the Main Window, with the additional data input being the first timed lap of a driver in the race (in seconds).

Depending on the time given, the algorithm will take all of the times from the chosen race in the last two years that the chosen driver has driven. Then, depending on the gaps between the first lap and the subsequent laps in the races, it will calculate the lap times for the chosen race when looking at the relative pace of the chosen driver in that race in the past:

- If the chosen driver has driven in only one occurrence of race chosen in the past, the algorithm will take his pace relative to the first lap of that race and apply those lap time gaps to the time provided.
- If the chosen driver has driven in two occurrences of race chosen in the past, the algorithm will take 60% of the most recent race pace relative to the first lap and 40% of the other and combine them to apply lap time gaps to the time provided.
- If the chosen driver has driven in three occurrences of race chosen in the past, the algorithm will take 50% of the most recent race pace relative to the first lap, 30% of the race in the middle, and 20% of the oldest race, and combine them to apply lap time gaps to the time provided.
- If the chosen driver has driven in four occurrences of race chosen in the past, the algorithm will take 30% of the most recent race pace relative to the first lap, 25% of the two races in the middle, and 20% of the oldest race, and combine them to apply lap time gaps to the time provided.

These lap times are visualized with a line chart where the prediction times are shown in purple color, while the actual lap times of the chosen race for the driver are shown in his team's color, so the user can understand how precise the prediction was.

The prediction algorithm also finds the fastest lap out of the laps predicted and displays it next to the actual fastest lap, so that they can also be compared. This chart solution also offers the zoom option, along with the option to move through the zoomed-in chart. This is displayed in Figure 5.

### IV. PREDICTION RESULTS

There are three possible outcomes (limitations) to the prediction algorithm:

1. With any incident or stoppage during lap one, the subsequent predicted laps could be much slower than the actual laps, since the reference point is a slow lap time. Example used: Pierre Gasly, Hungarian Grand Prix, 2021

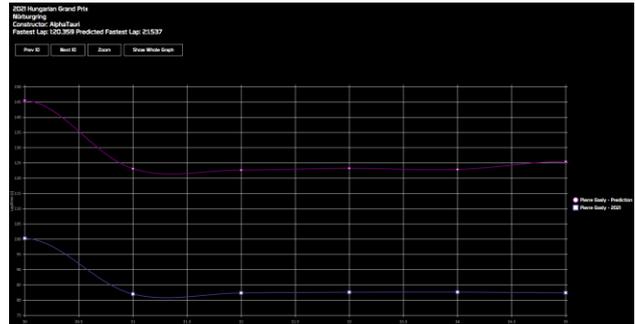


Figure 5. Visualization of actual and predicted lap times for a chosen driver in a chosen race

2. With a regular lap one, the prediction algorithm can be quite close to the actual lap times if the driver continues with good pace similar to the past seasons. Example used: Valtteri Bottas, Austrian Grand Prix, 2021
3. With a regular lap one, the driver can experience technical glitches that would hamper his pace, or if he is driving a car that does not perform as good as the cars he drove in past seasons, his actual times might be slower than those predicted. Example used: Daniel Ricciardo, British Grand Prix, 2021

These outcomes are shown in relation to the predicted fastest lap in Table 1.

### V. CONCLUSION

Result of this paper is WPF implemented software that proposes different ways to present similar datasets. It could also have been used to visualize any lap time data that comes from the Ergast API, like those simulated in [2]. The API's limitations were bypassed by getting all of the race lap times in a single query, and not having to take one lap time per query.

There are different ways to compare data, but also to inspect it more closely with options to zoom, move through the zoomed-in parts of the charts, and read the values from the tooltips provided.

When looking at future work, model of this solution could be expanded to be more similar to the one showed in [2] and offer comparison of driver pace when running on a chosen tire compound.

The prediction algorithm could be modified further, so that it takes data for two or more drivers and shows the

TABLE I. THREE OUTCOMES: ACTUAL AND PREDICTED LAP TIMES

	Actual Lap Time	Predicted Lap Time
Outcome 1	1:20.359	2:01.537 (41.178 seconds slower)
Outcome 2	1:08.375	1:08.522 (0.147 seconds slower)
Outcome 3	1.31.284	1.11.789 (almost 20 seconds faster)

prediction of their lap times on the same chart. We would also look at some other performance visualizing charts, similar to the one proposed in [4], and find ways to utilize them with these datasets. Finally, one interesting challenge would be to transfer this software into a smart phone app.

#### ACKNOWLEDGMENT

This research has been supported by the Ministry of Education, Science and Technical Development through the project no. 451-03-68/2020-14/200156: "Innovative scientific and artistic research from the FTN (activity) domain".

#### REFERENCES

- [1] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, S. Carpendale, "State of the art of sports data visualization", *Computer Graphics Forum*, 37(3): 663-686, 2018
- [2] C. Sulsters, "Simulating Formula One Race Strategies", 2018
- [3] T. Lampprecht, D. Salb, M. Mauser, H. Wetering, M. Burch, U. Kloos, "Visual Analysis of Formula One Races", 23rd International Conference Information Visualisation (IV): 94-99, 2019
- [4] P. Braun, A. Cuzzocrea, T. Keding, C. Leung, A. Padzor, D. Sayson, "Game Data Mining: Clustering and Visualization of Online Game Data in Cyber-Physical Worlds", *Procedia Computer Science*, 112: 2259-2268, 2017
- [5] D. Skau, R. Kosara, "Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts", *EuroVis16: Eurographics Conference on Visualization*, 2016
- [6] A. Ottley, A. Kaszowska, R. Jordan Crouser, E. M. Peck, "The Curious Case of Combining Text and Visualization", 2019, *EuroVis19: Eurographics Conference on Visualization*

# Energetics Aspect of Smart Buildings Based on Internet Of Things Architecture – An overview

Maja Gaborov, Dragan Ivetić, Srđan Popov

University of Novi Sad, Faculty of Technical Science in Novi Sad, Serbia  
majagaborov26@gmail.com, ivetic@uns.ac.rs, srdjanpopov@uns.ac.rs

**Abstract - The Internet of Things (IoT) is a growing area of research in the context of smart cities. The need for remotely accessible data and information / knowledge on the technical performance of buildings, at any time, from anywhere, regardless of the type of parameters, together with the need for full remote control will lead to the development of IoT for buildings. "Intelligent Buildings" (IBs) are designed with this technology to optimize the performance of their life cycle. The aim of this paper was to find architectures for application in smart buildings, predominantly architectures applied in the field of energy. Efficient management of heating, ventilation and air conditioning (HVAC) systems in smart buildings is one of the main applications of the IoT paradigm. The IoT architecture for the application of the HVAC model predictive control system (MPC) in a real environment is presented. Thanks to the proposed IoT-based architecture, the sensor, control and drive subsystems are connected to the Internet, and a remote interface with the HVAC control system is guaranteed to end users. The design and application of a presence sensor platform that can be used to accurately detect occupancy at the level of individual offices is also presented. Their presence sensor is cheap, wireless and can be gradually applied in existing buildings, and significant opportunities for energy savings have been discovered.**

**Key words: IoT, architecture, smart city, smart buildings, HVAC system, energy**

## I. INTRODUCTION

Smart cities concepts focus on the main changes in people's lifestyle from traditional to technological solutions. Technological solutions are being proposed to increase citizen's quality of life. Smart cities deal with the deployment of a network of devices (including IoT, mobile, and sensors devices). IoT data sources can connect to different data sources in the smart cities, including third-party applications and other related databases [1]. IoT is emerging as the next big wave of digital presence for billions of devices on the Internet. [2].

Smart Buildings are essential building blocks of Smart Cities. Intelligent buildings (IB) have been gaining increasing interest in recent years, as various IB technologies have been developed. The European Group on Intelligent Construction (UK) or the Institute for Intelligent Construction (USA) are focused on IB from a performance perspective. Therefore, the focus from this point of view is on user comfort, the ability to

quickly adapt to changing user needs, efficiently manage resources and minimize life cycle costs. IB should be viewed from a multi-industrial perspective, including the right combination of architecture, structure, information technology, automation, environment and energy, services and facility management, such as minimizing life cycle costs, maximizing comfort and adapting properly to cultural incentives [3]. A smart building is any structure that uses automated processes to automatically control the building's operations including heating, ventilation, air conditioning, lighting, security and other systems. A smart building uses sensors, actuators and microchips, in order to collect data and manage it according to a business' functions and services. This infrastructure helps owners, operators and facility managers improve asset reliability and performance, which reduces energy use, optimizes how space is used and minimizes the environmental impact of buildings [4].

This paper presents architectures which apply in the concept of smart buildings in field of energy.

## II. IOT AND SMART BUILDING

A smart city is defined as a city ecosystem that leverages connected technologies for the purpose of the betterment of city operations and intelligent near real-time decision making [5]. Smart cities are based on autonomous and distributed infrastructure that includes intelligent information processing and control systems heterogeneous network infrastructure, and ubiquitous sensing involving millions of information sources [6]. The Smart Cities architectures envisioned or implemented up to date deal mostly with use cases from the following categories: energy, waste disposal, environmental management, and transport [7][8]. A city usually consists of a very complex system of different types of infrastructure like ICT infrastructure, civil engineering infrastructure, social networks, financial network, etc. All of these systems require management effort like monitoring, reporting, and interaction to ensure efficient performance of all activities [9]. The IoT can be defined as a set of interconnected objects, sensors, electronic devices and services which can help a human to perform his tasks and duties from anywhere [10]. The devices are connected with the web using different technologies such as ZigBee, Bluetooth, etc. The communication between the devices and the user is carried out using a web-based infrastructure using the Mesh services [10].

According to some authors, term Iot was first used in 1998 [11]. The number of deployed IoT devices by 2021 just for smart buildings would be around 10.8 billion (2.8 billion for residential buildings) [12]. The world human population is predicted to increase in mid-2050 to 9,804 million. In addition, over half of the world human population resides in cities [13]. It needs not only network connections and extension of the Internet but also the assistance of satellite remote sensing and other positioning systems. In this way, the Internet of Things can accurately locate any object and realize remote operation [14]. IoT does not allow the use of a universal software architecture for different fields in which it is used, but needs to be adjusted according to the requirements of users [15]. The object or 'thing' in IoT can be any physical object which is equipped with a processing unit, sensors and actuators and is connected to internet via wired or wireless connectivity [16][17]. Every object in IoT has an identifier through which it is identified within the network [16]. As the city grows, new problems arise e.g., traffic congestion, waste management, pollution, parking allocation, etc. and the resources are scarce [18].

There are many possibilities for a smart city that is enabled with smart buildings and well provisioned and managed internet of things (IoT) frameworks. The benefits of such frameworks can be felt in multiple domains. On planning an integrated city, data captured from the infrastructures within a city can be analyzed and utilized for effective resource and services provisions within the city. In urban mobility, automated planning and demand responsive to public transportation integrated with self-driven vehicles. On safety, security and health of cities and their citizens, real-time surveillance and analytics can lead to better provision of emergency services, weather planning, early warnings on floods, tornadoes, man-made disasters and pandemics [19].

Smart Building (SB) is a new concept where the building benefits from the development of automation and communication technologies to create smart environments which are more efficient in the use of the available resources and much more secure [20]. Smart buildings (SB) are structures that use automated processes to control operations such as heating, ventilation, air conditioning, lighting, and security, and allowing sophisticated monitoring and control over their functions [21]. Smart buildings comprise all communication technologies that enable different sensors, objects, and functions within a building to communicate/interact with each other and also to be managed, controlled, and automated in a remote way [22]. According to a general survey, in the U.S., buildings are responsible for around 38% of the total carbon dioxide emissions, 71% of the total electrical energy consumption, 39% of the total energy usage, 12% of water consumption, and 40% of nonindustrial waste. Hence, smart buildings can create a large-scale impact on environmental conservation, financial savings, and energy preservation [22].

IoT has also influenced conventional building models and with its state-of-the-art technology has transformed common designs into smart ones. Smart buildings contain all communication technologies that allow various sensors, objects and functions in the building to communicate with each other and are also managed, controlled and automated remotely [23]. A related and cost-effective user-level IoT application is the support of IoT-enabled smart buildings. Commercial space has substantial requirements in terms of comfort, usability, security, and energy management. IoT-based systems can support these requirements in an organic manner. In particular, power over Ethernet, as part of an IoT-based solution, offers disruptive opportunities in revolutionizing the in-building connectivity of a large swath of devices. Data from the Survey on Energy Consumption in Commercial Buildings show that in 2012, there were 5.6 million commercial buildings in the United States, covering 87.4 billion square meters of floor space [24]. Globally, buildings (residential and commercial) are responsible for over 40% of total energy consumption [24].

Components of smart buildings are:

- **Occupancy Sensors:** Motion detectors can be employed to confirm the authenticity of the usage of resources. Temperature, lights, and other utilities can be adjusted according to needs and thus augment savings. These components can make a huge impact in office environments.
- **Light Sensors:** The variance of light according to the time of the day and the amount of sunshine can help in detecting the appropriate amiable light for rooms.
- **Thermal Sensors:** Weather conditions can be taken into account to set apposite temperature values. Further data concerning time zone and area can be a factor in making decisions.
- **CO2 Sensors:** Excess carbon dioxide in the air can be undesirable for well-being. With an increase in occupancy of an area, the level of CO2 also accentuates. Hence, the level of fresh air can be adjusted for alarming rates to develop a salubrious environment within the building area.
- **Actuators:** An actuator can be employed to convert electrical energy into some form of useful energy. Some examples are heating or cooling elements, speakers, lights, displays, and motors.
- **Medical Sensors:** In the healthcare field, IoT can be a really useful tool for the measurement and monitoring of medical parameters. Medical sensors can help in providing feedback and analysis to doctors or patients family at crucial times.

The benefits of smart buildings can be listed as follows:

- **Improved Efficiency:** Sensors employed in the vicinity help in the analysis of data which helps to predict the appropriate conditions and balance out the

requirements, thus by keeping a check on the consumption and overutilization of resources and energy, we can avoid losses and improve efficiency.

- **Easier Maintenance:** With complete control and active monitoring of the systems, reliability, performance of parts, and other components can be easily tracked and maintained. Detectors can also be applied to find the reasons behind malfunctioning or other issues.

- **Savings:** With sensors helping in identifying underused and overused areas and smart control systems allowing optimal climatic conditions to be maintained, a major factor of energy is saved and along with it an equally significant amount of lucre is procured.

- **Safety:** For mitigation and disaster management purposes, smart buildings can play a vital role in securing occupants and surroundings [22].

To implement the architecture, it is suggested to consider some points. First, it's very important to define the objectives of the SB, including what would be measured and/or controlled. Second, the implementation team must choose the components that they should acquire according to their needs, which involves if the building is a factory, an office, a household or other type of edification. The components include the MCUs, SBCs sensors and actuators. Third, the MCUs and SBCs must be programmed and the IoT devices must be configured all together, all in a specific way that serves the objective. After the SB has been successfully implemented, it's important to make sure that it maintains a high level of operability. Devices must receive maintenance in defined recommended periods to prevent failure and the SB must have a monitoring system able to check if one is down. There're special cases where the implementation team would have to take further actions, like covering some devices with protective cases or using dehumidifiers when the SB is near coast areas, where humidity can easily damage electronic devices [23].

### III. HEATING, VENTILATION AND AIR CONDITIONING (HVAC) SYSTEM IN SMART BUILDINGS

In recent times, the increase of energy efficiency is a pivotal goal for energy policy makers that aim at promoting a conscious, cost-effective, and sustainable energy use in management of resources and infrastructures, transportation and logistics, the production and industrial sector, and all activities related to human life. A significant part of energy consumption concerns the energy demand of buildings that in Europe amounts to 40% of the total and is mainly due to the so-called Heating, Ventilation and Air Conditioning (HVAC) systems. HVAC systems are devoted to guarantee hygrothermal comfort in building indoor environments and their automated management can largely impact the virtuous behavior of its end-users. In particular, it is of paramount importance to implement novel control architectures that on the one hand allow the energy optimization of HVAC systems without neglecting the thermal comfort of building

occupants and on the other hand offer to the policy makers and citizens (i.e., the end-users) an interactive tool for the monitoring and control of the HVAC system. This can be achieved by combining two main modules: a control algorithm and a smart physical infrastructure. The first module must be devoted to the energy and thermal comfort optimization while the second takes care of the sensing, communication, data storage, and actuation of the HVAC system on the basis of the end-users' requests. For the control algorithm module, many techniques can be considered that have been developed during the years for HVAC systems. However, the majority of them consist typically in the intuitive on/off controllers that cannot compensate for the high thermal inertia of many HVAC processes, or in simple PID controllers whose tuning is a complex activity and whose performance degrades if the system conditions vary. Only in the last two decades the more promising Model Predictive Control (MPC) approach is taking off: this control technique allows to effectively integrate issues such as disturbance rejection, constraint satisfaction, and slow-moving dynamic control together with energy efficiency strategies into the controller formulation. Furthermore, thanks to the decreasing costs of smart devices, the large availability of distributed sensors and data analytics tools, and in general the advances of Information and Communication Technology (ICT), the implementation of optimal control approaches for the energy efficiency and thermal comfort optimization is becoming more immediate and affordable. It is then evident that MPC becomes useless if it is not associated to a proper smart physical infrastructure that allows the collection/forwarding of actual data from/to the field. The Internet of Things (IoT) offers a proper solution by allowing the connection of sensors, actuators, and other objects to the Internet, and thus permitting the perception of the world, as well as the interaction with it [24].

Buildings are among the largest consumers of electricity in the US. A significant portion of this energy use in buildings can be attributed to HVAC systems used to maintain comfort for occupants. In most cases these building HVAC systems run on fixed schedules and do not employ any finegrained control based on detailed occupancy information [25]. The US Department of Energy estimates that buildings consume 70% of the electricity in the US. Recent efforts have focused on making buildings more energy efficient, including research that target specific areas such as HVAC lighting and managing IT energy consumption within buildings. The energy usage in a building can typically be divided amongst several subsystems, including plug loads, lighting, and mechanical equipment used for climate control. Mechanical equipment includes the combined heating, ventilation, and air-conditioning (HVAC) loads and constitutes a significant amount of energy consumption. Traditionally, most HVAC systems use only temperature and humidity as the primary inputs in determining cooling requirements. This limitation can often lead to inefficient energy usage. For example, a

room might be cooled to 22.9C regardless of whether there are any occupants [25].

#### IV. IOT BASED CONTROL SYSTEM ARCHITECTURE

Figure 1 presents the overall IoT based control system architecture which allows to optimize the energy efficiency and thermal comfort of the internal environment of smart buildings. “It consists of the following elements:

- a net of sensors that perceives the environmental conditions and sends measurements to a gateway;
- a set of HVAC modules;
- a net of actuators that control the HVAC modules and communicate with the gateway;
- a gateway that connects the nets of sensors and actuators to the Internet;
- an external Application programming interface (API) that provides forecasts of the weather conditions;
- an external database server that collects/forwards data from/to the field and from/to the control unit;
- a control unit that communicates with the database server and where the MPC algorithm is executed;
- an IP device that acts as end-user interface that is connected to the database server and hosts a dashboard dedicated to monitoring the state of the environment and setting the control system mode” [24].

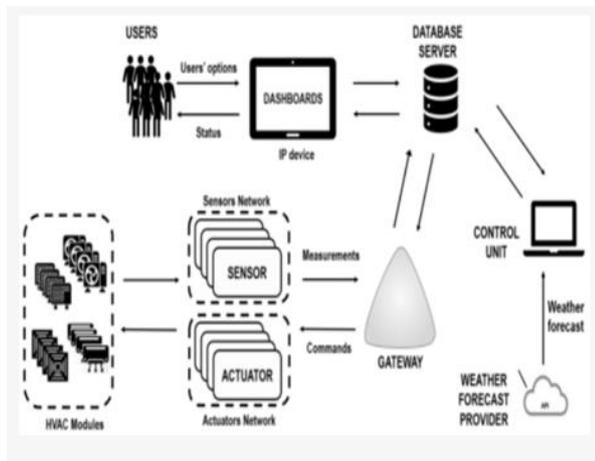


Figure 1. The high-level system diagram of the proposed IoT based architecture[24]

The environmental indoor conditions (i.e., temperature, CO2 level, number of occupants, etc.) and the energy consumption are measured by means of the net of sensors. The measurements are periodically sent to the gateway, which communicates with the database server, and then to the control unit where the MPC algorithm is deployed. This algorithm provides to the HVAC systems the control actions that ensure the best compromise in terms of energy consumption and comfort for the given comfort constraints during a particular time horizon. These control actions are sent via the gateway to the devices that actuate the HVAC

modules [24]. “The HVAC modules modify the room temperature according to the decisions taken by the MPC algorithm. The database server stores the measurements of temperature and energy consumption. These measurements are displayed at the end user by means of the dashboard on an IP device, which is connected to the database server. The dashboard allows users to interact with the control unit and select the desired temperature and the desired control mode” [24].

#### V. RELATED WORK

In one paper authors proposed an IoT based architecture for the implementation of Model Predictive Control (MPC) of HVAC systems in real environments. The considered MPC algorithm optimizes on line, in a closed-loop control fashion, both the indoor thermal comfort and the related energy consumption for a single zone environment. Thanks to the proposed IoT based architecture, the sensing, control, and actuating subsystems are all connected to the Internet, and a remote interface with the HVAC control system is guaranteed to end-users. In particular, sensors and actuators communicate with a remote database server and a control unit, which provides the control actions to be actuated in the HVAC system, users can set remotely the control mode and related set-points of the system while comfort and environmental indices are transferred via the Internet and displayed on the end-users’ interface. The proposed IoT based control architecture is implemented and tested in a campus building at the Polytechnic of Bari (Italy) in a proof of concept perspective. The effectiveness of the proposed control algorithm is assessed in the real environment evaluating both the thermal comfort results and the energy savings with respect to a classical thermostat regulation approach[24].

In one study authors proposed a self-managing architecture for multi-HVAC systems in buildings, based on the “Autonomous Cycle of Data Analysis Tasks” concept. A multi-HVAC system can be plainly seen as a set of HVAC subsystems, made up of heat pumps, chillers, cooling towers or boilers, among others. Their approach is used for improving the energy consumption, as well as to maintain the indoor comfort, and maximize the equipment performance, by means of identifying and selecting of a possible multi-HVAC system operational mode. The multi-HVAC system operational modes are the different combinations of the HVAC subsystems [26].

In some paper authors have presented the design and implementation of a low-cost and incrementally deployable occupancy detection system using battery operated wireless sensor nodes. Their evaluation across a ten room initial deployment shows that their choice of sensors and occupancy detection algorithm can detect occupancy accurately. Using this occupancy information as input to a simulation model of a building, they show that the HVAC energy consumption can in fact be reduced from 10% to 15% using their system. They also believe that the actual energy savings might even be greater with more

aggressive control algorithms that can utilize the occupancy information provided by our sensors to duty-cycle HVAC systems further [25].

Model predictive control is theoretically suitable for optimal control of the building, which provides a framework for optimizing a given cost function (e.g., energy consumption) subject to constraints (e.g., thermal comfort violations and HVAC system limitations) over the prediction horizon. However, due to the buildings' heterogeneous nature, control-oriented physical models' development may be cost and time prohibitive. Data-driven predictive control, integration of the "Internet of Things", provides an attempt to bypass the need for physical modeling. One paper presents an innovative study on a data-driven predictive control (DPC) for building energy management under the four-tier building energy Internet of Things architecture. Here, they develop a cloud-based SCADA building energy management system framework for the standardization of communication protocols and data formats, which is favorable for advanced control strategies implementation. Two DPC strategies based on building predictive models using the regression tree (RT) and the least-squares boosting (LSBoost) algorithms are presented, which are highly interpretable and easy for different stakeholders (end-user, building energy manager, and/or operator) to operate. The predictive model's complexity is reduced by efficient feature selection to decrease the variables' dimensionality and further alleviate the DPC optimization problem's complexity. The selection is dependent on the principal component analysis (PCA) and the importance of disturbance variables (IoD). The proposed strategies are demonstrated both in residential and office buildings. The results show that the DPC-LSBoost has outperformed the DPC-RT and other existing control strategies (MPC, TDNN) in performance, scalability, and robustness [27]. This paper reports an innovative study combining the data-driven predictive control strategy with a complex cloud SCADA-based building energy management platform, which attempts to standardize communication protocols and data formats and further implement advanced control strategies. The platform also provides useful data representations to different stakeholders (end-user, building energy manager, and/or operator), enabling the platform flexibility and scalability. They present two algorithms, based on RT and LSBoost, to create control-oriented models for the DPC. Moreover, an efficient feature selection method, which depends on the principal component analysis and the importance of disturbance variables, is leveraged to decrease the model's dimension and further alleviate the DPC optimization problem's complexity. They then apply the DPC to two different case studies for energy consumption in residential and office buildings. The numerical simulation shows that the DPC-LSBoost provides lower energy consumption while maintaining the required thermal comfort compared to the MPC, the TDNN, and the DPC-RT. With the same environmental comfort demand, compared with the TDNN, the peak power consumption with the DPC-LSBoost can be

reduced by 11.92% and 30.4%, even compared to the DPC-RT 4.99% and 11.54% that are achieved. These advantages make the DPC-LSBoost an attractive tool for large-scale cyber-physical energy systems to reduce energy consumption. Also, in the context of prediction time, comparing with the MPC, the prediction time of the DPC-LSBoost is reduced by 72 s and 294.7 s, respectively [27].

## VI. CONCLUSION

Smart buildings are a rising endeavor of technology and a promising amalgamation of IoT with the contemporary world.

One of main contributions to the consumption of energy in buildings is the heating, ventilation and air condition (HVAC) system. Multi-HVAC systems can be seen as an intermediate level of modeling, composed by different HVAC subsystems (e.g. boilers, chillers, heat pumps, among others), that define their operations and enabling the simulation of their behavior at a given time. HVAC systems are an essential part in buildings, since they are responsible for offering a temperature and air quality acceptable interior in buildings such as offices, hospitals, centers commercial, among others. They are also the main energy consumers in buildings. Therefore, apply adequate energy management in these systems, without compromise interior comfort, contributes greatly to the energy efficiency in buildings. A significant portion of this energy use in buildings can be attributed to HVAC systems used to maintain comfort for occupants. Buildings are known to be one of the largest consumers of electricity in world.

The literature was searched and the authors in the mentioned works use IOT-based architectures for smart buildings and it was concluded that most smart buildings are based on HVAC systems which are used to achieve energy savings and improve energy consumption.

Future research could examine whether smart buildings are the largest consumers of energy in Serbia and how electricity savings are achieved.

## REFERENCES

- [1] Sinaeepourfard, A., Krogstie, J., Petersen, S. A., & Ahlers, D. (2019). F2c2C-DM: A Fog-to-cloudlet-to-Cloud Data Management architecture in smart city. In 2019 IEEE 5th World Forum on Internet of Things (WF-IoT) (590-595).
- [2] Simmhan, Y., Ravindra, P., Chaturvedi, S., Hegde, M., & Ballamajalu, R. (2018). Towards a data-driven IoT software architecture for smart city utilities. *Software: Practice and Experience*, 48(7), 1390–1416. doi:10.1002/spe.2580
- [3] Brad, B. S., & Murar, M. M. (2014). Smart buildings using IoT technologies. *Stroitel'stvo Unikal'nyh Zdanij i Sooruzenij*, (5)
- [4] Kim, T. H., Ramos, C., & Mohammed, S. (2017). Smart city and IoT.
- [5] Abu-Matar, M., & Davies, J. (2017). Data driven reference architecture for smart city ecosystems. In 2017 IEEE Smart World, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI) (1-7).

- [6] Sharma, P. K., & Park, J. H. (2018). Blockchain based hybrid network architecture for the smart city. *Future Generation Computer Systems*, 86, 650-655.
- [7] Nitti, M., Pilloni, V., Giusto, D., & Popescu, V. (2017). IoT Architecture for a sustainable tourism application in a smart city environment. *Mobile Information Systems*, 2017
- [8] Medina, C. A., Pérez, M. R., & Trujillo, L. C. (2017). IoT paradigm into the smart city vision: a survey. In 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) (695-704).
- [9] Tanwar, S., Tyagi, S., & Kumar, S. (2018). The role of internet of things and smart grid for the development of a smart city. In *Intelligent communication and computational technologies (23-33)*. Springer, Singapore.
- [10] Khan, M., Babar, M., Ahmed, S. H., Shah, S. C., & Han, K. (2017). Smart city designing and planning based on big data analytics. *Sustainable Cities and Society*, 35, 271-279. doi:10.1016/j.scs.2017.07.012
- [11] Ahmed, S. H., & Rani, S. (2018). A hybrid approach. *Smart Street use case and future aspects for Internet of Things in smart cities*. *Future Generation Computer Systems*, 79, 941-951.
- [12] Verma, R., Brazauskas, J., Safronov, V., Danish, M., Merino, J., Xie, X., ... & Mortier, R. (2021). SenseRT: A Streaming Architecture for Smart Building Sensors
- [13] Sinaeepourfard, A., Krogstie, J., Petersen, S. A., & Gustavsen, A. (2018). A zero emission neighbourhoods data management architecture for smart city scenarios: discussions toward 6Vs challenges. In 2018 International Conference on Information and Communication Technology Convergence (ICTC) (658-663). doi.org/10.1109/ICTC.2018.8539669
- [14] Zhang, C. (2020). Design and application of fog computing and Internet of Things service platform for smart city. *Future Generation Computer Systems*, 112, 630-640.
- [15] Gavrilović, N., & Mishra, A. (2021). Software architecture of the internet of things (IoT) for smart city, healthcare and agriculture: analysis and improvement directions. *Journal of Ambient Intelligence and Humanized Computing*, 12(1), 1315-1336.
- [16] Malche, T., Maheshwary, P., & Kumar, R. (2019). Environmental monitoring system for smart city based on secure Internet of Things (IoT) architecture. *Wireless Personal Communications*, 107(4), 2143-2172.
- [17] Ajumol P A, Elizabeth Isaac (2020), Nova Genesis - An Advanced Architecture for Internet of Things, *International Journal of Recent Technology and Engineering (IJRTE)* ISSN: 2277-3878, Volume-9 Issue-1, DOI:10.35940/ijrte.A2324.059120
- [18] Camero, A., & Alba, E. (2019). Smart City and information technology: A review. *Cities*, 93, 84-94
- [19] Inibhunu, C., & McGregor, A. C. (2020). A Privacy Preserving Framework for Smart Cities utilising IoT, Smart Buildings and Big Data. In 2020 IEEE 22nd International Conference on High Performance Computing and Communications; IEEE 18th International Conference on Smart City; IEEE 6th International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (1096-1103).
- [20] Carrillo, E., Benitez, V., Mendoza, C., & Pacheco, J. (2015). IoT framework for smart buildings with cloud computing. In 2015 IEEE First International Smart Cities Conference (ISC2) (1-6).
- [21] Mazzara, M., Afanasyev, I., Sarangi, S. R., Distefano, S., Kumar, V., & Ahmad, M. (2019). A reference architecture for smart and software-defined buildings. In 2019 IEEE International Conference on Smart Computing (SMARTCOMP) (167-172).
- [22] Khanna, A., Arora, S., Chhabra, A., Bhardwaj, K. K., & Sharma, D. K. (2019). IoT architecture for preventive energy conservation of smart buildings. In *Energy Conservation for IoT Devices (179-208)*. Springer, Singapore.
- [23] L. Aguilar, S. Peralta and D. Mauricio, "Technological architecture for IoT smart buildings," 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), 2020, (1-6), doi: 10.1109/ICECCE49384.2020.9179358.
- [24] Carli, R., Cavone, G., Ben Othman, S., & Dotoli, M. (2020). Iot based architecture for model predictive control of hvac systems in smart buildings. *Sensors*, 20(3), 781.
- [25] Agarwal, Y., Balaji, B., Gupta, R., Lyles, J., Wei, M., & Weng, T. (2010). Occupancy-driven energy management for smart building automation. In *Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building (1-6)*.
- [26] Aguilar, J., Garces-Jimenez, A., Gallego-Salvador, N., De Mesa, J. A. G., Gomez-Pulido, J. M., & Garcia-Tejedor, A. J. (2019). *Autonomic Management Architecture for Multi-HVAC Systems in Smart Buildings*. *IEEE Access*, 7, 123402-123415. doi:10.1109/access.2019.2937639 1
- [27] Ke, J., Qin, Y., Wang, B., Yang, S., Wu, H., Yang, H., & Zhao, X. (2020). Data-Driven Predictive Control of Building Energy Consumption under the IoT Architecture. *Wireless Communications and Mobile Computing*, 2020.

# An Android-based Application for Reading Serbian Identity Cards – Challenges and Design Considerations

J. Jovanović\*, D. Dragan\*, D. B. Gajić\*, and V. B. Petrović\*

\* University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia

jovana.jovanovic@uns.ac.rs, dinud@uns.ac.rs, dusan.gajic@uns.ac.rs, pveljko@uns.ac.rs

**Abstract – Digitalization has taken off in Serbia and is supported in all aspect of Government. The Serbian Identity Card is one of the documents that support digital procedures. The Identity Card is used to identify people and in order to read the data from it, it is necessary to use a computer and a desktop application. This is a problem when it is necessary to identify people in the field. This paper describes an Android application that solves this problem. We describe the architecture of the application and the problems we encountered during its development. The application is published on Google Play and is available to the general public. The application has been received fairly well and its success can be seen through its excellent rating on Google Play.**

## I. INTRODUCTION

Digitization in Serbia has taken off and is supported in all aspects by the Government of Serbia [1]. One document that supports digital procedures is the Serbian Identity Card (ID) [2]. A Serbian ID can contain a contact microcontroller enabling the use of digital IDs. A majority of Serbian population is using just this type of ID. As a result, there is an official desktop application for reading Serbian IDs, named ČELIK, published by the Serbian Ministry of Internal Affairs [3]. Of course, the application requires adequate card-reader hardware. Although the data from a Serbian ID can be exclusively used in digital form, almost all of the procedures which include a person's identification require, in addition to the ID itself a supplemental document containing a printout of the data held on the chip.

There are situations when the establishment of a person's identity is needed in the field, where the use of a computer is not practical. This includes both government agencies and the private sector. For example, insurance or property agents are required to identify their customers. Thus, they need a smartphone application which will enable them to read the customer's ID if they wish to do this outside the office. Interestingly, to the best of our knowledge at the moment of writing, beside our application, there are no software solutions for reading Serbian IDs on Android smartphones of the new generation. Part of the reason is that special equipment is needed to read smart cards in general. These readers usually do not come with card reader software. And they do not come with such specialized software as the one

needed for reading Serbian IDs, because of the proprietary format of Serbian IDs. Thus, in cooperation with MHT, the official distributor of ACS mobile card readers, we developed Čelik2Go application for reading Serbian IDs [4]. Our application works with ACR39U-ND and ACR39U-NF pocket smart card readers [5], [6].

There have been several challenges in developing the Čelik2Go application. One of them was the format of the Serbian ID, which is not freely available. Another challenge was the format in which data from the ID has to be printed. For it to be recognized as a valid document, the printed document has to look the same as if it was printed from the ČELIK application. Also, as it is highly unlikely that a printer will be present in the field, the application has to support output in the PDF format and to enable sharing of this generated document for later printing. Naturally all of this must support full confidentiality.

During the previous ten years the Serbian Government has issued several versions of Serbian ID. This often led to incompatibilities between the old Serbian IDs and the software relying on them. Therefore, the third challenge was that the application had to support all the versions of the Serbian IDs currently in circulation.

The fourth and final challenge was user experience. Based on the potential market research done by MHT [7] group and their experience in selling pocket smart card readers, we had a good idea who our potential users are going to be. They range from expert and proficient users of Android smartphones to novice users with bad attitudes towards smartphones. Our users are of all ages, from early twenties to late seventies. In addition, the application's most important requirement was that it should be used outside as part of business interaction. Given the nature of the user and the nature of this requirement, we opted for a design which minimizes interaction complexity and allows the most important actions to be achieved in three clicks or less ensuring that the most important choices at any given time are clearly marked and distinguishable.

The Čelik2Go application is published and available on Google Play for a small annual fee. The application has at the moment more than 50 users and its score on Google Play is 4.6 which clearly indicates that we have been successful in overcoming all of the challenges and bring our users a quality application with good user experience.

The rest of the paper is organized as follows. In Section 2, we present the current state in the field. We describe the applications that are used for reading Serbian IDs and give a brief overview of applications used for reading smart cards in general. In Section 3 we describe the architecture of Čelik2Go application. We also describe in brief the format of Serbian IDs and the smart card readers that the Čelik2Go application is compatible with. In Section 4 we discuss how we overcome the challenges presented in the Introduction. Section 5 concludes the paper and presents our plans for future work.

## II. BACKGROUND

This chapter discusses the ČELIK desktop application and existing Android applications for the same purpose created for earlier versions of Android.

### A. ČELIK

ČELIK [3] is a desktop application developed by the Serbian Government used for reading Serbian IDs. The data read from the ID card can be printed out and used in that form to identify persons. The application is primarily intended for business entities that need to identify citizens as a part of their business. It has a simple design and is easy to use. Reading data from an ID is simple as it requires the user just to insert the ID into a smart card reader to read the data that is then displayed on the application's home screen. The data can also be printed in a paper format approved by the Serbian Government or saved as a PDF document.

### B. Android application for reading ID card

Before implementing Čelik2Go application, we searched Google Play store for similar applications to help us design the app. We found an OMNIKEY application that supports a specific type of OMNIKEY reader. The application reads ID data, and displays it on the mobile device's screen. Here we point out few application disadvantages: (1) the application usage is limited to OMNIKEY card readers which makes users dependent on one card reader vendor, (2) to save PDF of the ID data, the user has to be registered which might trigger concerns about user's personal data, and (3) the application does not support the latest ID format standard. However, the application has a simple and easy to use design. As it can be seen in Figure 1., the application consists of a single screen to display data which makes it easy to navigate and suitable for both experienced and new users.

The other applications found on Google Play are based on document scanning. The data obtained are based on an image obtained by scanning. Some of these applications are Android Id Card Reader [9], ReadID Me [10] and BlinkID [11].



Figure 1. OMNIKEY application

## III. ARCHITECTURE OF THE APPLICATION

The format of the ID, the compatibility with smart card readers and the architecture of the application will be described in this chapter.

### A. Serbian ID format

An ID is a document issued by the Serbian Ministry of Internal Affairs, which serves to identify citizens. An ID includes document data and basic citizen data. Document data includes the ID number, the date of issuing the ID, the date of ID expiration and the institution that issued the ID. Name, surname, unique master citizen number (UMCN), name of one parent, gender, residence, place of birth and date of birth are the basic information about the owner of the ID. In addition to these basic data, the ID also contains images of biometric data such as a photograph, a fingerprint, and a signature. The residence data consists of data on the residence of the owner of the ID consisting of the state, city, street, house number, and the date when the residence was last changed.

### B. ACS pocket smart card reader

Čelik2Go app is compatible with two types of reader ACR39U-ND [5] and ACR39U-NF [6]. ACR39U is a portable smart card reader suitable for various applications in any field of business. This smart card reader, no larger than a USB stick, is capable of supporting demanding smart card applications. It provides valuable and reliable functionality to meet user's security needs. It is multipurpose, as it can be used to read several types of smart cards, which include IDs, traffic licenses, as well as national health insurance cards. ACR39U supports ISO 7816 Class A, B, and C smart cards (5 V, 3 V, and 1.8 V) and microprocessor cards with T=0 and T=1 protocol. In addition, it supports a wide range of memory cards. This makes it ideal for a broad range of solutions, such as PIV, Physical and Logical Access Control, Digital Signature, and Online Banking applications.

The difference between the two mentioned readers is in the type of connector used. One reader has a newer USB C connector and the other uses a type USB B which is another thing that allows for widespread use of this reader. Figure 2. shows how the reader is connected to a mobile phone.



Figure 2. Reader connected to a mobile phone

### C. Architecture of the application

The Čelik2Go application was created for Android devices, primarily for mobile devices such as smartphones and tablets. The application is supported on Android devices start with Android 5, Lollipop, up to the latest, Android 11. Based on the market research done by MHT, we decided to focus our development solely on Android devices, because that is the type of the phone used by most of our potential users.

The application consists of two parts. One part is an Android application developed using Android's XML and the Java programming language, and the other part is the Spring application. The main part is the Android application because it reads the data from the chip, displays the data, and creates a PDF that is identical to the printed document created by the ČELIK desktop application. The application's layout is implemented using Android's XML. An EVK library for the Android operating system [12] was used to connect to the reader and read data from the chip.

To make the application available to the general population, it was published on Google Play. One of the requirements was to add licenses to the application, in order to protect the application's code and to make sure that one user account can use the application on only one device. It was necessary to use the LVL library [13] offered by Google Play to solve this request. With this library, we are able to communicate with a Google Play server that performs license verification. Additionally, it is necessary to set a policy to check the license. For that purpose, we used StrictPolicy. The Strict Policy is a strict verification policy, which means that every time an application is launched, it also checks whether the application is licensed or not. A difficulty was ensuring that only one device can be used with one license. The default behavior of Google Play is to allow a single account to use the application on multiple devices. Also, Google Play offers us an interface for the implementation of our device limiter, but due to the access modifier of a

specific interface it was not possible to implement it. To successfully solve this problem, we implemented the Spring application as a licensing server. Each time an affirmative response is received from Google Play for a license, we also contact the background server to verify that the license is correct for that device. In addition to licenses, there is a request for implementing subscriptions to the application, which can be monthly or annual. Subscriptions are resolved using the Google Play Billing library [14]. In addition to the integration using the library, it is necessary to adjust certain settings on the Google Play Console, such as the type of application (free or paid), the type of subscription and the price.

## IV. IMPLEMENTATION AND DISCUSSION

There were four important challenges concerning the application's functionality and usage we encountered during application development. These are decoding the format of the ID, matching the form of the ČELIK printed ID card report, supporting different types and editions of the ID and maintaining a consistent user experience. Each of these problems will be discussed in a separate subsection.

### A. ID format

All of the data this application needed to read was on the chip embedded in the ID card. It was necessary to read out this data. The used libraries allowed us to get the raw binary form of the data in the form of byte strings. Careful manipulation and reverse engineering were then employed to determine the offsets at which crucial data was stored. Further, the byte strings needed to be decoded into human-readable string form using correct character encodings.

### B. Format for printing reading data from ID

The PDF document our application creates must have the same format as the printed document obtained using the ČELIK application. This problem was solved through careful matching of the formatting extracting the PDF produced by the ČELIK application. Once the format was created the iText library was used to generate the first pass of the PDF from the Android application. The data was added first, and then in the second pass formatting lines.

### C. How to support multiple different ID formats

Since the biometric ID was introduced to Serbia, its formats have changed. There are two types of ID, one is of 'Apollo' type and the other is of the 'Gemalto' type. Gemalto IDs are the ones currently in use, and the slightly older ones are Apollo. For these reasons, two types of ID were created in the application. Gemalto [15] and Apollo [16] IDs differ only in the format in which they store data, so it was necessary to define formats for both types. This part of the application is implemented in a modular way and it is very simple to extend it to support new types of ID.

*D. User experience*

First of all, the goal was to create an application that will be used by people of all ages and people with different expertise in the use of mobile devices. In order to successfully solve this problem, we first started creating an user interface flow diagram (UIFD) [17] in order to immediately see all the advantages and disadvantages of the solution we created. The parts we paid special attention to were the appearance of the application itself and the number of actions needed to perform a certain functionality.

The first version of the UIFD is shown in Figure 3. When the application is launched, the user is shown the main window from which, if user launches the application for the first time, he goes to the window for entering data to check the license, and only if the license is correct the application home page opens. It is designed to have two buttons on the home page, one for reading a new ID and the other for viewing all previously read IDs. Clicking on the new ID button opens a window where the data is displayed after the reader is connected to the mobile device and the ID is inserted. If the user wants to view all processed IDs from the home page the user can access the list by clicking on the button ‘view all IDs’. The list is displayed on the screen where they can select a

particular ID from the list to delete it, share it with someone or save it on an arbitrary location in a phone’s filesystem. To read the ID and save the data, the user should perform four actions: click on the button in the home window, enter the serial number, and then click on the new ID button and save when the data is successfully read. To view all IDs, he needs 3 actions, and one more to display a certain ID.

The evolved UIFD of the final application is shown in Figure 4. . If we look at this UIFD we will see that there are some significant differences in relation to the UIFD from Figure 3. The first and main difference relates to the number of actions required to perform an operation. Since the main function of the whole application is reading the ID, then the user should be enabled to do that functionality as simply as possible. The application is connected to Google Play, so there is no need for a screen on which data on the serial number is entered and thus the license is checked. Dropping this window reduced the number of actions. Also, the home page has been changed, and the list of read IDs is now displayed on that page, so that with one click the user can see all IDs that he has read so far and saved in his mobile device as shown in Figure 5. .

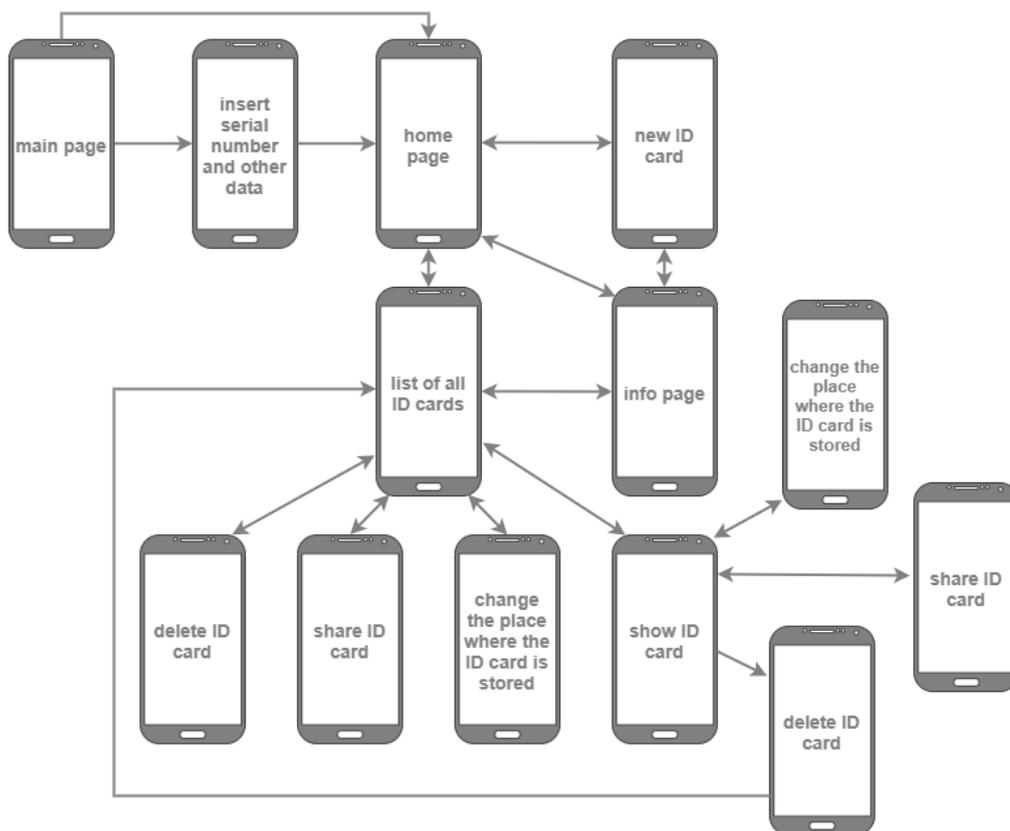


Figure 3. First version of the UIFD

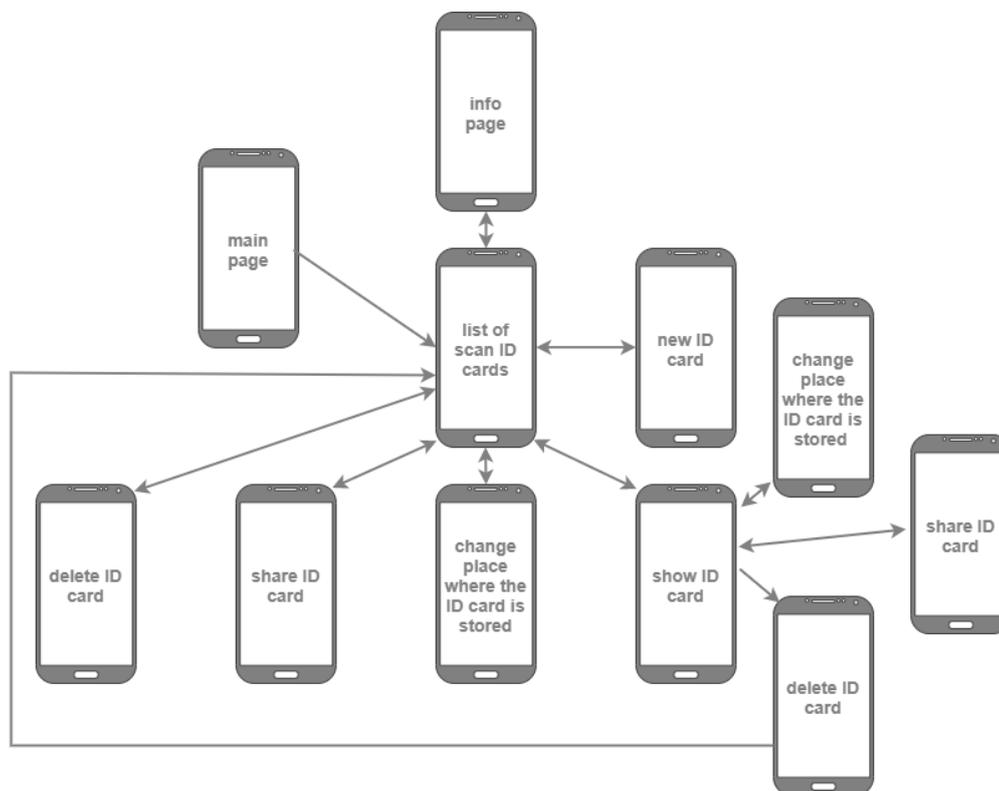


Figure 4. Final version of the UIFD

Clicking on one of the IDs displays the PDF of the selected ID, in addition to viewing with one click, we can also delete the PDF document of a previously read ID, change the place where the document is stored or share it to someone via one of the available applications on the mobile device. By clicking the ‘Scan’ button, a connection is made between the connected reader and the mobile device and the data read is displayed on the screen. Thus, we need three actions to read the ID and store the read data, which means one less than the original idea of what the application should look like.

By modifying and removing individual windows, we got an application that requires fewer steps to perform the action. In addition to the number of actions, another

important thing that affects the usability of applications is the avoidance of errors and the design of the application itself. Since the users of this application are a diverse group, we decided on a simpler design so that even beginners can use the application in an easy way. The simple design of this application also helps to reduce the mistakes that users can make. Only the necessary actions at any given time are displayed on the screen. Only available actions are ever displayed, if an action cannot be performed then it is not even displayed on the screen at that time. Also, all actions are clearly highlighted and icons that are commonly used are used to mark them.

## V. CONCLUSION

Digitalization in Serbia is in full swing and there are a lot of procedures that include reading data from smartcard IDs. A lot of these procedures take place outdoors, in the field and a smartphone solution for reading Serbian IDs is needed. Čelik2Go, as described in the paper, is one such a solution. The entire process of creating the application is described in the paper. We described the problems we encountered during the development and how we solved them. The main problems that we had to address were: reading the format of the IDs, creating PDFs which match the printed of the ID reading, compatibility with different types of IDs, and adaptability to different user groups.

The application was published on Google Play. Based on the statistics and the reviews so far, we can say that we have successfully solved all the challenges and requirements that we encountered during the development of the application. Currently, the application has over 500 downloads and is rated 4.6.

Although we have managed to solve the present problems, we are expecting that new problems will arise



Figure 5. Home page of Čelik2Go

in the future. Serbian Government has so far changed the ID format and card type several times, so it is safe to assume they will do so again. Also, it is possible that there will be changes in Serbian law that will lead to changes in Serbian IDs. For this reason, we designed the application so that it can be easily modified and expanded to meet any new demands resulting from these changes.

As part of our future work, we will continue to maintain the application so as to support latest developments in the Android platform, and to continue to provide support for any changes in Serbian IDs that may arise. Also, we plan to support other documents types in addition to the personal ID cards, such as driver and vehicle licenses.

#### REFERENCES

- [1] Pitic, Goran & Savic, Nebojsa & Verbić, Srđan., "Digital Transformation and Serbia." *Ekonomika preduzeća*. LXVI.10.5937/EKOPRE1802107P, 2018.
- [2] Identity card law, Published in "Official Gazette of the Republic of Serbia" No. 62/06., <http://arhiva.mup.gov.rs/domino/zakoni.nsf/Identity%20Card%20Law.pdf>, Belgrade, 2008
- [3] Aplikacija citac elektornske licne karte, Ministarstvo unutrasnjih poslova, <http://arhiva.mup.gov.rs/domino/dokumenta.nsf/aplikacijalkl>, Retrieved 8 August 2021
- [4] Čelik2Go, <https://celik2go.com/>, Retrived 20 August 2021.
- [5] ACR39U-ND, <https://www.acs.com.hk/en/products/449/acr39u-nd-pocketmate-ii-smart-card-reader-micro-usb/>, Retrieved 8 August 2021
- [6] ACR39U-NF, <https://www.acs.com.hk/en/products/425/acr39u-nf-pocketmate-ii-smart-card-reader-usb-type-c/>, Retrieved 8 August 2021
- [7] Citac pametnih kartica, <https://citaci.kartica.rs/>, Retrieved 8 August 2021.
- [8] OMNIKEY aplikacija, Google Play, <https://play.google.com/store/apps/details?id=com.freelancewatermelon.licnakarta&hl=en&gl=US>, Retrieved 15 August 2021
- [9] Android ID Card Reader, <https://www.nfcread.com/en/android-id-card-reader>, Retrieved 15 August 2021
- [10] ReadID Me, <https://play.google.com/store/apps/details?id=nl.innovalor.nfciddo.cshowcase&hl=en&gl=US>, Retrieved 15 August 2021
- [11] BlinkID – ID card and passport scanner, <https://play.google.com/store/apps/details?id=com.microblink.blinkidapp&hl=en&gl=US>, Retrieved 15 August 2021
- [12] EVK library for Android, <https://www.acs.com.hk/en/mobile-support/>, Retrieved 20 August 2021
- [13] Licensing Verification Library, <https://developer.android.com/google/play/licensing/overview#LV L>, Retrieved 9 August 2021
- [14] Integrate the Google Play Billing Libray into your app, <https://developer.android.com/google/play/billing>, Retrieved 9 August 2021
- [15] Gemalto, <https://www.esmartsys.com/en/partners/gemalto/Pages/gemalto.aspx>, Retrieved 12 August 2021
- [16] Apollo, [http://www.apolloelectronics.com/en/products\\_list\\_3\\_22.html](http://www.apolloelectronics.com/en/products_list_3_22.html), Retrieved 12 August 2021
- [17] User Interface Flow Diagrams (UI Storyboards): An Agile Introduction, <http://agilemodeling.com/artifacts/uiFlowDiagram.htm>, Retrieved August 2021

# The Influence of Audio vs. Multimedia Classroom Instruction on Critical Thinking of EFL Learners

Lela Ivanovska

Faculty of Information and Communication Technologies, Bitola, North Macedonia  
lela.andreevska@uklo.edu.mk

**Abstract - The popularity of information technology (IT) over the past decade has brought about the innovative use of the Internet in English language learning and teaching process. Using computers and Internet can improve EFL learners' language abilities, and they also reach real learning experiences. They gain the opportunity to access useful language resources and communicate with native English speakers through computers and Internet. Critical thinking is a higher-order cognitive skill that is indispensable to students, readying them to respond to a variety of complex problems that are sure to arise in their personal and professional lives. The aim of the paper is to explore the relationship between critical thinking ability of EFL learners and the use of different media such as audio vs. visual texts in English classes. The participants of this study are 80 first year university students from the Faculty of Information and Communication Technologies, Bitola, North Macedonia. Watson-Glaser Critical Thinking Questionnaire is used as an instrument to evaluate the learners' critical thinking. For these purposes MANOVA was used. In this paper we specifically focus on the use of audio and multimedia files in order to extend the students critical thinking skills through interactive activities that force them to diverse opinions, analyze and synthesize course content. The results revealed that video materials with text performed better than audio group.**

## I. INTRODUCTION

The popularity of information technology (IT) over the past decade has brought about the innovative use of the Internet in English language learning and teaching process as Evans (Evans, 1983, according to Ilievski, 1999: 192) in his study "Computer Challenge" will say: "Instead of increasing muscle strength, we aim towards increasing brain strength". Using computers and Internet can improve EFL learners' language abilities, and they also reach real learning experiences. In addition, computers contribute beneficially to learners' learning and development. They gain the opportunity to access useful language resources and communicate with native English speakers through computers and Internet. Computers are important in language learning because they help students to think critically in their learning process and make them have active and stable knowledge. That is, they are provided with more creative activities to analyze and assess through using computers. Critical thinking is a higher-order cognitive skill that is indispensable to students, readying them to respond to a variety of complex

problems that are sure to arise in their personal and professional lives. Critical thinking has been recently introduced and gained a high position in foreign language teaching (FLT) settings so that nowadays enhancing critical thinking in learners is considered one of the foreign language teachers' tasks. The relevance of critical thinking with technology is to prepare the students in order to be the best solver and make the better decision. It is important for the students to be independent thinker since there are many jobs needed skillful workers which have critical thinking ability. Meanwhile, the students should know how to be creative and to develop reflective and logical thinking in order to decide some problems.

Based on the above thought, this study investigates the use of audio and video with texts recording lessons in stimulating critical thinking among EFL students from a public university in North Macedonia. More precisely, the lessons contain a wide range of topics about technology that encourage students to answer the questions through interactive activities that help students advance their thinking.

## II. THE INTERPLAY OF TECHNOLOGY AND CRITICAL THINKING

There is a thin line between language and thought in relation to interaction (Bowerman and Levinson, 2001; Chomsky, 1975; Vygotsky, 1978; Whorf, 1956). It means they are interrelated, and teachers must train students in order to develop their linguistic and cognitive skills by contrasting the target language with their own language, hypothesizing the grammatical rules of the target language. The use of technologies as cognitive tools engage students in thinking while they learning by visualizing with technology, learning by reflecting, and learning by exploring that it is not to do without technology support. When students use audio and video activities they can examine question and reflect on what they learn. In other words, with these media tools students have opportunities to learn through visual formats, mental models and problem-oriented activities that provide them with higher degree of CT (Carmichael and Farrell, 2012). The correlation between critical thinking skills and a technology-rich environment provides an improvement of students' motivation, scaffolding, and feedback. The Fig.1 below represents the cognitive skills at the foundation of critical thinking (Facione, 2011).

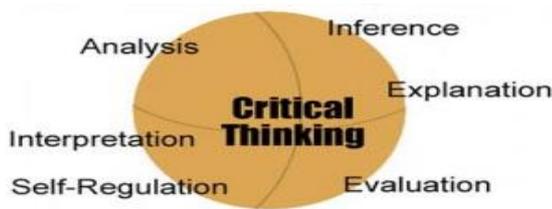


Fig.1 Core critical thinking skills

The image shows when students think critically, they engage in the following process: analysis, synthesis, evaluation, problem-solving, communication and reflection. It means they can develop both lower-level (remembering, understanding) and higher level (applying, analyzing, evaluating and creating) critical thinking skills (Facione, 2011). Huang et al. (2012) mention that learning environments that incorporate technological tools into classrooms enable students to develop arguments supported by making thinking processes ‘visible’, and eventually foster enhanced critical thinking skills. However the utilization of audio and video materials could improve students’ critical thinking ability and language teaching. Audio tends to use one channel, that include only audio information which is spoke language information. Now, combinations of media as input could be selected and the learners have a possibility to be exposed to authentic language. This kind of media provides “methods that reflect the cognitive processes necessary to successfully perform a given learning tasks”(Clark and Paivio 1991). Multimedia means a combination between sight and auditory and can improve learning process and achieve high critical skills. Digital learning helps students see the importance of what they are learning and make a connection between theoretical ideas and critical thinking (Mayer, 1997). Mayer pointed out when learners use multimedia to stimulate thinking they are better in applying, coherence, modality redundancy and individual differences. Specifically, when learners use auditory-verbal and visual-pictorial channel, verbal thinking is involved. Besides, this learning media was one of dominant aspects after learning method which could improve learning process and achieve high learning result (Sudjana and Rivai, 2007, p.2). Technology makes students to be self-dependent thinkers (Burgess, 2009). According to Abreu (2010) media literacy can enhance students’ critical thinking skills, as a big challenge for the 21st century. The application of learning technological tools and critical thinking skills provides an opportunity for interested students to achieve higher levels of knowing and to practice critical thinking skills (Carmichael and Farrell, 2012). Li (2010) claims that through activities that include processes by providing Internet resources, designed tasks, and interaction with others, students can develop critical thinking skills more effectively. Critical thinking is the ability to connect new knowledge to previous knowledge, to construct and evaluate arguments, and solve problems systematically. Students are better able to grasp complex concepts when tasks are explained using a wide array of modalities (verbal, visual, graphical, and symbolic) and instructional formats (audio files, video lectures, graphic displays, and simulations). Digital learning environments foster critical thinking and increase the accessibility of content by offering learners more

options for applying knowledge and skills. The capacity to, monitor, evaluate and control thinking while completing new tasks—helps support critical thinking and transfer of knowledge. By providing, coaching, extensive modeling, scaffolding, and problem solving, technology offers learners opportunities to build metacognitive skills. (Yang and Wu, 2012).

In relation to cognitive research the extensive student practice is a vital component of learning. Technology provide more opportunities for students to practice skills and concepts. It helps foster critical thinking by transferring knowledge from short-term to long-term memory as a significant process that helps students apply and remember information to new settings (Mandernach, 2006).

### III. METHOD

Using both qualitative and quantitative methods, this experimental study seeks answers to the following research question: Is there a significant difference between critical thinking ability of EFL learners and the use of different media such as audio vs. visual texts in English classes?

#### A. Participants

Participants of the study are 80 first year university students from the Faculty of Information and Communication Technologies, Bitola, North Macedonia who study English as a foreign language. For the purpose of this study two intact groups consisting of 40 in each are included.

#### B. Instruments

To carry out the research investigation, 3 different instruments were employed in the present study:

*Straightforward Quick Placement & Diagnostic test* – a language proficiency test,

*Audio and video with text recording lessons/interactive activities*

Seven topics which are in accordance with the English subject in the first year study program same for both groups are selected as audio and visual materials. The participants are asked to respond to the materials in a personal way.

#### *Watson-Glaser Critical Thinking Questionnaire*

In order to evaluate the students’ critical thinking, Watson-Glaser Critical Thinking Questionnaire was used. It is consisted of five subtests: a) Inference b) Recognizing Unstated Assumptions c) Deduction d) Interpretation e) Evaluation of Arguments.

#### C. 3.3 Procedures

For the purpose of this study the following procedures were followed:

First, a general proficiency test was administered in order to make sure of the proficiency level of the students (intermediate). Then, 80 students were divided into two

groups (audio and video materials with text) of 40. Second, one of the groups (called audio) had a task only to listen the materials, while the second one (called video materials with text) to hear, see and read the same materials. The audio/video files took up 2-3 minutes. The students were required to complete the questions through interactive activities in stimulating critical thinking in relation to the level of numeration and explanation, explanation of functions and characteristics and explanation of differences and similarities about technology during one month. Third, the students were asked to complete Watson-Glaser Critical Thinking Questionnaire.

#### IV. RESULTS AND METHODS

Based on the results we can conclude that students are active with particular interest on higher thinking skills in their audio and video activities. It means that every student manages their learning process in order to achieve the relational thinking. They can share their experience and knowledge actively. This implies that knowledge and understanding are constructed when students are engaged actively. In this case students prefer a certain single different learning style. They have different ability to manage the input through technology tools. In fact, technology has an impact on students involvement in learning activities. Indirectly, that activities have been mixed with critical thinking abilities.

Table 1 shows that the sample is 40 observations (students). In this sample, the average of critical thinking level in audio group is 36.45. 24 students are above that average (with an average score of 41.21) and 16 students are below that average (with an average score of 29.18).

Table 2 shows that the sample is 40 observations (students). In this sample, the average of critical thinking level in audio group is 62.7. 21 students are above that average (with an average score of 69.95) and 19 students are below that average (with an average score of 54.68). Related to the above results students are motivated to discovery critical thinking. Providing words with images, pictures or other graphics enhances critical thinking skills instead of materials that include words. Results show that students from 'audio group' have "lack of focus", they are less critically oriented then the students from the group video materials with text, and not enough able to receive, evaluate and respond to a message. They are not much engaged in activities that require them to develop critical thinking skills and evaluate and analyze course content. The students from the second group 'video materials with text' are more inspired to think critically and deeply and to seek out different solutions.

TABLE I: Audio materials

Dummy variables for critical thinking	Critical thinking level	Observations
0	29.1875	16
1	41.29167	24
Total	<b>36.45</b>	<b>40</b>

TABLE II: Video materials with text

Dummy variables for critical thinking	Critical thinking level	Observations
0	54.68	19
1	69.95	21
Total	<b>62.7</b>	<b>40</b>

On this basis multimedia emphasizes critical thinking rather than rote learning. It helps students advance their knowledge and engage them in multiple levels of critical thinking through reflective activities. Students can construct hypotheses and make decisions. It promotes the development of students' critical thinking skills and self-regulation. Students are the active creator of their knowledge. This means that students must be creative and active in all their activities. Critical thinking and multimedia increase the development of students' expressions and different learning style as well as the ability to manage and monitor the input. The visual and auditory nature of video stimulates critical thinking, it provides opportunities for interacting materials. This kind of information forms can cause heavier cognitive load, it provides broader context and visual details that can help develop critical thinking skills. Results indicate that multimedia is closely a superior tool for visually demonstrating "how-to" through the presentation of authentic information. Multimedia aids critical thinking in a number of ways. These include stimulating interest, increasing knowledge and assisting in comprehension. Multimedia design enable students to make their sense of what and how they think. It is based on action-oriented. Results display the fact that when students are presented with visual aids such a printed text or visual information they create their own critical thinking through experience, they think actively, create schemes, etc. They use both "top-down" and "bottom-up" learning methodology while thinking process. In this thinking, the students get the details of the given idea and explain opinions by making thinking processes 'visible'. This positive indicators of usage of multimedia help students develop their critical skills, make connections between arguments and information and meet their individual needs.

#### V. CONCLUSION

The findings of the study found a significant difference between the critical thinking ability of EFL learners and the use of different media such as audio vs. visual texts in English classes. The study suggests that presenting information with image, sound and text gives better results versus presenting information only with sound and positively reflects on students' knowledge. Multimedia creates an effective learning environment that motivates students to develop critical thinking skills. It also provides effective tools for students to share the content and construct knowledge through learning activities. Technology has an impact on the intellectual learner and on the development of both micro and macro world technology. It can be used as an effective tool to facilitate critical thinking skills within a higher education setting or

in the content of audio and multimedia learning. Critical thinking also extends to the field of linguistics and learning process in particular.

Technology can extend language classes so critical thinking can open new avenues for foreign language teachers. However, multimedia plays a visible role in enhancing critical thinking and language pedagogy. Bringing together critical thinking skills and learning technological tools may be beneficial in that it provides an additional opportunity for students to practice critical thinking skills, problem solving and to express feelings and thoughts.

#### REFERENCES

- [1] Bowerman, M., & Levinson, S.C. (2001). Introduction. In M. Bowerman and S. C. Levinson (eds.), *Language acquisition and conceptual development*, 1-16. Cambridge: Cambridge University Press.
- [2] Burgess, M. (2009) Using WebCT as a supplemental tool to enhance critical thinking and engagement among developmental reading students. *Journal of College Reading and Learning*, 39 (2) ,10-33
- [3] Carmichael, E., & Farrell, H. (2012). *Evaluation of the effectiveness of online resources in developing student critical thinking: Review of literature and case study of a critical thinking online site. Journal of University Teaching and Learning Practice*, 9(1).
- [4] Carmichael, E., & Farrell, H. (2012). Evaluation of the Effectiveness of Online Resources in Developing Student Critical Thinking: Review of Literature and Case Study of a Critical Thinking Online Site. *Journal of University Teaching and Learning Practice*, 9(1), 4.
- [5] Chomsky, N. (1975). *Reflection on Language*. New York: Pantheon.
- [6] Clark, J. M., and Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149-210.
- [7] De Abreu, B. (2010). Changing Technology = Empowering students through media literacy education. *New Horizons in Education*, 58(3), 26-33.
- [8] Facione, Peter A, *Think Critically*, Pearson Education: Englewood Cliffs, NJ, 2011.
- [9] Halpern, D. F. (1996). *Thought and knowledge: An introduction to critical thinking* (3<sup>rd</sup>ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc
- [10] Huang, K. H., Hung, K. C., & Cheng, C. C. (2012). *Enhancing interactivity in geography class: Fostering critical thinking skills through technology. Problems of Education in the 21<sup>st</sup> Century*, 50.
- [11] Ilievski A. (1999). *Teoretskite osnovi na nastavata*. Biagl-Bitola
- [12] Kabilan, M.K. (2000). Creative and critical thinking in language classrooms. *The Internet TESL Journal*, 6 (6). Retrieved November 21, 2005 from <http://itseli.org/Techniques/KabilanCriticalThinking.html>
- [13] Li, H. C. (2010). Using podcasts for learning English: perceptions of Hong Kong Secondary 6 ESL students. *Début: The Undergraduate Journal of Languages, Linguistics and Area Studies*, 1(2), 78-90.
- [14] Mandernach, B. J. (2006). Thinking critically about critical thinking: Integrating online tools to Promote Critical Thinking. *Insight: A collection of faculty scholarship*, 1, 41-50.
- [15] Mayer, R. E. (2001). *Multimedia learning*. Cambridge, England: Cambridge University Press
- [16] Sudjana, N. & A. Rivai. (2007). *Media Pengajaran*. Bandung: Sinar Baru Algesindo.
- [17] Vigotsky, L. S. (1978). Interaction between learning and development. In Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (Eds.), *Mind in Society: The development of higher psychological processes*. Cambridge, MA: Harverd University Press.
- [18] Worf, B. L. (1956). *Language, Thought and Reality: selected writings of Benjamin Lee Whorf*. J. Carroll (Ed.). Cambridge, MA: MIT Press.
- [19] Yang, Y. T. C., & Wu, W. C. I. (2012). Digital storytelling for enhancing student academic achievement, critical thinking, and learning motivation: A year-long experimental study. *Computers & Education*, 59(2), 339-352.

# Modelling Business Intelligence Systems for Effective Decision Making

Mihalj Bakator\*, Dragan Čočkaló\*, Dejan Đorđević\*, Melita Čočkaló-Hronjec \*\*, Dragana Milosavljev\*

\* University of Novi Sad, Technical Faculty "Mihajlo Pupin", Đure Đakovića bb, 23000, Zrenjanin, Republic of Serbia

\*\* High school "Laza Kostic", Laze Lazarevica 1, 21102, Novi Sad, Republic of Serbia

mihalj.bakator@uns.ac.rs, dragan.cockalo@tfzr.rs, dejan.djordjevic@tfzr.rs, melita.cockalo@gmail.com, milosavljev@yahoo.com

**Abstract** - The modern business environment is characterized by the globalization of markets and the digitalization of socio-economic dimensions. Such an environment brings challenges to enterprises when it comes to achieving a competitive position on the market maintaining adequate business performance. Managers of enterprises have to include a wide variety of tools, techniques and methods when it comes to effective and efficient decision making. Decision support systems have the potential to significantly improve business performance through enhancing the decision making process. In this paper, the application of business intelligence system for effective decision making is discussed. The main goal is to present a theoretical model that can be used for future research in this domain. Additionally, suggestions and guidelines regarding the application of business intelligence systems in enterprises are noted.

## INTRODUCTION

The modern business environment, which is characterized by the globalization of markets, the rapid development and distribution of modern information-communication technologies (ICTs), and the post-pandemic business paradigm, brings challenges for domestic enterprises, and for enterprises in developed countries as well. Due to the globalization of markets, obtaining and maintaining competitiveness is difficult, as small and medium-sized enterprises are competing against big corporations and vice-versa [1]. In order to improve competitive ability, enterprises have to address the constant changes on the international market, and to adapt flexible business strategies that are sustainable in the long-term. This requires the continuous evaluation and analysis of internal and external business indicators and metrics. Managers have to be able to overview every business process and the majority of impactful market trends. This further indicates that effective and timely decision-making is an imperative. The decision making process is a crucial part of effectively and efficiently managing business processes. Managers have to address and take into consideration every possible ICT solution that can help them make short- and long-term decisions.

Business intelligence (BI) is often used to describe large decision support systems that are used in enterprises. Business intelligence represents a key infrastructural investment and a great deal of attention is given to it by chief-information-officers (CIO) in large enterprises, and by managers in SMEs [2]. Business intelligence possesses

various analytical capabilities including central data storage (data warehouse), data analysis, trend analysis, and business process optimization. Business intelligence systems possess value creation capabilities through effective and data-driven support in the domain of various business process including, but not limited to, manufacturing, marketing, sales, accounting, inventory, storage and other internal operations [3]. Business intelligence systems present an immense opportunity for enterprises as they have a big potential to increase productivity, improve risk management and increase overall profitability [4]. Business intelligence systems as a form of decision support systems have the ability to provide infrastructural support for knowledge management, which also contributes to effective and efficient decision making processes [5].

In this paper, the application of business intelligence systems for decision making in enterprises is analyzed. The main goal is to develop a generic theoretical model that can be used for future research and analysis in the domain of business intelligence systems application in various enterprises and in their various sectors. The paper consists of four main sections (excluding the Introduction and Conclusion sections). The first section provides an overview on business intelligence and business intelligence systems. The second section addresses the application of business intelligence systems in decision making processes. The third section presents the theoretical model for decision making that is based on business intelligence systems. The fourth section discusses suggestions and guidelines for enterprises when it comes to the implementation and application of business intelligence systems and similar solutions.

## I. BUSINESS INTELLIGENCE SYSTEMS

The fourth industrial revolution - Industry 4.0 brought the digitalization of business processes through data-based and leveraged value creation for the customer [6]. Business intelligence is not functional on its own, but rather it presents the core of data collection, storage, and analysis, and these processes are often delegated to other systems or even third parties. Business intelligence is a very effective tool to extract large amounts of information from big datasets (Big Data), thus it includes analytical abilities and operations, data integration, content analysis for quality, and overall it has a strategic importance in key functional areas of conducting business [7].

Business intelligence systems can significantly contribute to enterprises in their goal of achieving competitiveness. Some of the cornerstone applications of ICTs in business that characterize business intelligence systems in various capacities are data visualization, data reporting, report generating, enterprise-wide data warehouses, process monitoring and conducting analytical tasks [8].

Business intelligence systems include activities, processes, tools and techniques that can be used in commerce, manufacturing industries, law, government, and other technological areas and non-manufacturing industries. In addition, communication is the main mediator and facilitator of the technical side and application side of business intelligence systems. Intelligence and the system concept indicate the ability to learn and extract information from existing datasets and to apply that knowledge in an enterprise that will support specific business processes and activities [9].

It is important to note that business intelligence systems that are used in decision making can be addressed also as decision support systems. These are important and crucial tools that include a wide array of statistical analysis, artificial intelligence, and effective big data analysis. Business intelligence systems can improve real-time reporting and forecasting of various internal and external trends (profitability, market share, work overload, bottlenecks, revenue, sales etc.), customer relations and overall it can significantly contribute to faster decision making [10]. Business intelligence systems can handle a large amount of data and simplifies the whole data gathering, storage and analysis process on different organization levels. This further indicates that such systems can provide a detailed overview on the external and internal business environment of the enterprise. This overview can significantly contribute to the decision making process and overall it can improve business performance.

Furthermore, it can be noted that business intelligence systems have to be aligned with business process management. Through this alignment, business intelligence systems can integrate additional value into business process management, which further translates into value for the customer and for the enterprise (stakeholders) [11]. The application of business intelligence systems in enterprises derives from the expansion of ICT use in business. The digital transformation that affects both markets and enterprises has brought changes not only how business interact between each other on the market, and with customers, but also how managers apply modern ICTs within the enterprise.

## II. APPLICATION OF BUSINESS INTELLIGENCE SYSTEMS IN DECISION MAKING

The process of decision making in a business environment requires experience, leadership and knowledge as key factors of efficient and effective decisions. Business intelligence tends to focus on the large amounts of data for extracting significant information regarding the internal and external business environment.

Well-established business intelligence systems can be a strong basis for expanding business and achieving competitiveness [12]. Business intelligence solutions can be applied in various business sectors and for various business processes. The implementation and application of business intelligence solutions depends on the needs, capabilities, and strategies of the enterprise. The main principle of a business intelligence solution is real-time data, which is, in addition, actionable and contains key information for effective and efficient decision making [13].

Further, combining cloud-computing technologies with business intelligence solutions creates a whole new level of data and information management with the focus on decision making. Cloud-based business intelligence solutions can be used through the Internet. The data and information can be accessed from almost any device that has Internet connection capabilities and minimum required hardware and software. Authorization is also an imperative, as business intelligence systems contain exactly that - "intelligence" or information on the business/enterprise. Cloud-based business intelligence solutions provide on-ground support and they are flexible, require less resources, require less infrastructure, and the majority of technical/non-business oriented factors are operated by the cloud platform provider. Some of the main platforms for cloud business intelligence are SaaS Business Intelligence solutions, Business Intelligence Platform as a Service, and data integration tools [14]. However, cloud business intelligence platforms bring challenges of their own. For example, there are security issues of sensitive data; latency of real-time data, which is slower compared to on-site business intelligence solutions; lack of expertise as cloud technologies advance; data integration, which can be tedious if there are no adequate mechanisms in place for distributing and sharing data; and control over data, as the provider is the main governance body of the cloud business intelligence system, there is a risk of losing the whole system alongside with data if the provider is not in-sync with the enterprise's goals and strategies [14].

Business intelligence systems are driven by the fast growth of technology and its diffusion in the business world. Such systems integrate complete sets of techniques, tools (software solutions) and methodologies that are applied in enterprises with the goal to obtain, store and analyze data through which effective decision making can be conducted [15]. It was also noted that the implementation process is just the beginning when it comes to applying business intelligence systems. More precisely, adoption by the employees and managers also plays a crucial role, as resistance can annul the potential positive effects of such implemented systems.

When implementing a business intelligence solution it is necessary to evaluate the situation within the enterprise as there is a big chance for inadequate implementation, which further leads to failure or produce sub-par results. More precisely, over 70% of business intelligence system implementations fail and this has an extremely negative effect on the enterprise as the implementation process is resource draining [14].

Applying business intelligence systems in a business environment enables the overview of specific business indicators (external and internal) as the obtained data is transformed into a perspective that helps managers to make effective and efficient decisions. Business intelligence systems aim at turning the obtained data into structured information that will help the enterprise to improve its business performance [16]. The basis of the business intelligence systems includes the analysis of data stored in operational databases, forecasting and trend prediction on the market, strategic indicator evaluation, and combining intellectual capital with the decision making process.

Manufacturing enterprises can implement and apply specific business intelligence solutions that can help improve and optimize manufacturing processes with the goal to increase efficiency and to secure real-time data collection and analysis. Through this real-time data, managers can obtain significant insight into the manufacturing process and arrange changes accordingly. Changes can be in the form of load capacity optimization, additional employee hiring, process speed reduction, etc. [17]. In this process of real-time data presentation, the business intelligence system involves algorithmic calculation and evaluation of material flow, inventory capacity and status, supplier status, manufacturing schedules, orders, and stock levels.

The advantage of business intelligence systems when it comes to decision making lies in its capability of fast-paced selection of irrelevant information and can present complex relations between various indicators in accordance with the enterprise's goals and set projections [18]. These systems can improve decision-support and it can manage to obtain and analyze large datasets and transforms them into potential actions that can enhance organizational capabilities (optimization and adjusting resources, schedules, supply chains etc.). It uses the extracted knowledge from data as drivers for innovation, creativity, productivity, and sustainable expansion of business activities. Through these activities, enterprises can significantly increase their chances when it comes to achieving a competitive position on the globalized international market [19].

### III. DECISION MAKING MODEL BASED ON BUSINESS INTELLIGENCE

Based on the conducted overview of business intelligence system application in enterprises and its specific use in decision making, a theoretical model is developed that provides a concise overview on the mechanisms that drive the concept of effective decision making that is supported by business intelligence. The model is generic in nature and for it to be applicable in enterprises it has to be defined in more detail with specifications, parameters, and characteristics of the enterprise that would implement it and apply it. This further indicates that the model is not applicable in the practical sense, but rather presents the main modules and relations that are part of an integrated whole, which is decision making through business intelligence systems. The modules/elements on the model are numbered and

additional details are noted. The model is presented on Figure 1.

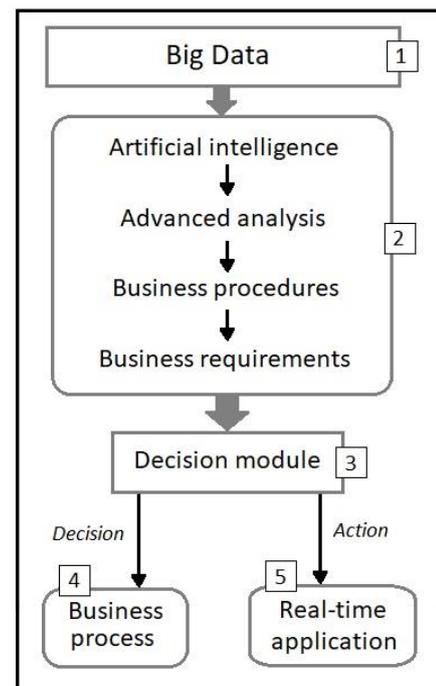


Figure 1. Decision making model based on business intelligence

The model presented on Figure 1. has five main modules/elements. The first element (1) Big Data refers to the large volume, large variety and velocity of data that is collected from various networks, but mainly from the Internet. Such data in itself is not sufficient for application and decision making, as it contains a lot of "noise". Hence, in the second element of the model (2) artificial intelligence-based data analysis and advanced statistical analysis with prediction scenarios is conducted. The analyses are conducted in accordance with the existing business procedure and business requirements. This way, the extracted information from the large amounts of data can be applied for specific business processes and decision making can be specific and can contribute to business performance.

The decision making process is conducted in the third element (3) of the model. Here, the managers and even authorized employees are supported by valuable information extracted from databases. From the decision module, there are two main results. The first is a decision that is integrated into business processes (4) in order to improve them and see the potential expected results. The second result of the decision module is an action that is applied in real-time (5). These actions are applied in managerial strategic and operational processes. Compared to decisions in business processes (4), real-time application (5) is conducted in stand-by business processes, which require timely action. This approach increases efficiency as time-sensitive issues and processes are categorized and decisions are "tailored" in a timely manner, while decisions for non-urgent business processes are handled differently as they don't require prioritized action. This option is crucial especially when there are

multiple actions/decisions required for different processes, and prioritization is imperative.

#### IV. GUIDELINES AND SUGGESTIONS

Based on the analyzed literature in the domain of business intelligence systems, big data and their application in enterprises, the following suggestions and guidelines for improving business performance through decision making are presented:

- domestic enterprises should evaluate their business processes and to identify which processes and functions can be supported through business information systems.
- when choosing a business intelligence solution, managers have to aware of the enterprise' financial and intellectual capabilities as well as the enterprises requirements.
- the business intelligence system should be configured in manner that includes a wide array of business processes and the analysis part of the system should include modularity to some degree, which would allow for narrow and specific decision making.
- if the enterprise doesn't have enough financial resources for internal business intelligence systems, managers could consider such systems provided from third parties. These solutions are mainly cloud-based and can be appropriate for enterprises of any size.
- data security has to be taken into consideration depending on from where the data is obtained. Regulation has to be followed and privacy laws should be taken into consideration when extracting and applying data.
- before considering any type of business intelligence system, managers have to evaluate the situation in which the enterprise is. Namely, business intelligent system can't solve structural issues and strategic issues within an enterprise. Similarly, if the enterprise can't achieving even moderate competitiveness on the market, then the potential of a business intelligence system to improve business performance is low.
- business intelligence systems have the main role as decision support systems, and they should be treated as such. This means that these systems are not a solution but rather a type of enhancement. However, in rare occasions, if the main issue is decision making, then a business intelligence system model for decision making would, in fact significantly contribute to business performance improvement.

Overall, enterprises have to consider the possibilities of modern technologies. Especially advanced systems (in

this case business intelligence systems) that have the potential to improve the various aspects of conducting business. It is also important to note, that conducting business in the modern, digitalized economic environment, requires complex and flexible approaches and business strategies that can adapt quickly to dynamic changes on the market.

#### V. CONCLUSION

The digitalization of socio-economic dimensions, the globalization of markets, and the rapid development of technologies, has led to new business dynamics, where enterprises face challenges and issues regarding competitive ability and sustainable business development. In this paper, the application of business intelligence systems for decision making was addressed. It can be concluded that such systems possess tremendous potential if applied correctly. The presented generic and theoretical model provides an overview on how big data from various sources transform into a decision that can help improve business processes.

The main limitation of this paper is the lack of meta-analysis and the lack of empirical research. However, the paper is structured as a review paper with discussion on suggestion and guidelines. The paper provides a solid basis for future research, and contributes to the existing body of literature as it analyzed the modern concept of business intelligence application. For future research it is recommended to conduct surveys in enterprises to investigate indicators (decision making time, communication, revenue, reaction to market changes etc.) pre- and post- business intelligence system implementation.

#### REFERENCES

- [1] M. Bakator, D. Đorđević, D. Čočkalović, M. Nikolić, and M. Vorkapić, "Lean startups with industry 4.0 technologies: Overcoming the challenges of youth entrepreneurship in Serbia," *Journal of Engineering Management and Competitiveness (JEMC)*, vol. 8, no. 2, 89-p.101, 2018.
- [2] D. Arnott, F. Lizama, and Y. Song, "Patterns of business intelligence systems use in organizations," *Decision Support Systems*, vol. 97, pp. 58–68, 2017. doi:10.1016/j.dss.2017.03.005
- [3] A. Popović, B. Puklavec, and T. Oliveira, "Justifying business intelligence systems adoption in SMEs." *Industrial Management & Data Systems*, 2018. doi:10.1108/imds-02-2018-0085
- [4] L. M. D. Yiu, A. C. L. Yeung, and T. C. E. Cheng, T. C. E. "The impact of business intelligence systems on profitability and risks of firms," *International Journal of Production Research*, pp. 1–24, 2020. doi:10.1080/00207543.2020.1756506
- [5] Á. Rocha, H. Adeli, L. P. Reis, S. Costanzo, I. Orovic, and F. Moreira, "Trends and Innovations in Information Systems and Technologies," *Advances in Intelligent Systems and Computing*, 2020. doi:10.1007/978-3-030-45688-7
- [6] F. E. Bordeleau, E. Mosconi, and L. A. de Santa-Eulalia, L. A. "Business intelligence and analytics value creation in Industry 4.0: a multiple case study in manufacturing medium enterprises," *Production Planning & Control*, pp. 1–13, 2019. doi:10.1080/09537287.2019.1631458
- [7] Y. Y. Lim, and A. P. Teoh, "Realizing the strategic impact of business intelligence utilization," *Strategic Direction*, vol. 36, no. 4, pp. 7–9, 2020. doi:10.1108/sd-09-2019-0184

- [8] M. J. Sousa, and I. Dias, I. "Business Intelligence for Human Capital Management," *International Journal of Business Intelligence Research*, vol. 11, no. 1, pp. 38–49, 2020. doi:10.4018/ijbir.2020010103
- [9] J. A. Becerra-Godinez, J. L. Serralde-Coloapa, M. S. Ulloa-Marquez, A. Gordillo-Mejia, and E. Acosta-Gonzaga, E. "Identifying the main factors involved in business intelligence implementation in SMEs," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 1, pp. 304-310, 2020.
- [10] M. Alasiri, Mohanad, and A. A. Salameh, "The Impact of Business Intelligence (BI) and Decision Support Systems (DSS): Exploratory Study," *International Journal of Management*, vol. 11, no. 5, pp. 1001-1016, 2020. Available at SSRN: <https://ssrn.com/abstract=3631747>
- [11] D. Suša Vucec, V. Bosilj Vukšić, M. Pejić Bach, J. Jaklič, and M. Indihar Štemberger, M. "Business intelligence and organizational performance. *Business Process Management Journal*," ahead-of-print(ahead-of-print), 2020. doi:10.1108/bpmj-08-2019-0342
- [12] A. George, K. Schmitz, and V. C. Storey, "A Framework for Building Mature Business Intelligence and Analytics in Organizations," *Journal of Database Management*, vol. 31, no. 3, pp. 14–39, 2020. doi:10.4018/jdm.2020070102
- [13] A. Tripathi, and T. Bagga, "Leading Business Intelligence (BI) Solutions and Market Trends, " *Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020*. Available at SSRN: <https://ssrn.com/abstract=3568414> or <http://dx.doi.org/10.2139/ssrn.3568414>
- [14] S. S. Patil, and R. Chavan, "Cloud business intelligence: an empirical study," *Stud. Indian Place Names UGC Care J*, vol. 27, 747-754, 2020.
- [15] S. Ahmad, S. Miskon, T. A. Alkanhal, and I. Tlili, "Modeling of Business Intelligence Systems Using the Potential Determinants and Theories with the Lens of Individual, Technological, Organizational, and Environmental Contexts-A Systematic Literature Review," *Applied Sciences*, vol. 10, no. 9, p. 3208, 2020. <https://doi.org/10.3390/app10093208>
- [16] A. Grigorescu, D. Baiasu, and R. I. Chitescu, "Business Intelligence, the New Managerial Tool: Opportunities and Limits," *Ovidius University Annals, Economic Sciences Series*, vol. 20, no. 1, pp. 651-657, 2020.
- [17] W. Danilczuk, and A. Gola, "Computer-aided material demand planning using ERP systems and business intelligence technology, " *Applied Computer Science*, vol. 16, no. 3, 2020.
- [18] Y. Chen, Yand Z. Lin, "Business Intelligence Capabilities and Firm Performance: A Study in China," *International Journal of Information Management*, 102232, 2020. doi:10.1016/j.ijinfomgt.2020.1022
- [19] M. Bakator, D. Đorđević, and D. Čočkaló, "Developing a model for improving business and competitiveness of domestic enterprises," *Journal of Engineering Management and Competitiveness (JEMC)*, vol. 9, no. 2, pp. 87-96, 2019.

# Improving CRM with Internet of Things and Big Data

Mihalj Bakator\*, Dragica Radosav\*, Mila Kavalić\*\*, Dragana Milosavljev\*, Edit Terek Stojanović\*

\* University of Novi Sad, Technical Faculty "Mihajlo Pupin", Đure Đakovića bb, 23000, Zrenjanin, Republic of Serbia

\*\* University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21102, Novi Sad, Republic of Serbia

mihalj.bakator@uns.ac.rs, dradosav@tfzr.rs, milakavalic@gmail.com, dragana.milosavljev@tfzr.rs, edit.terek@tfzr.rs

**Abstract - Due to the dynamic changes on the international and digitalized market, enterprises have to adjust their business strategies in order to obtain and maintain a competitive position on the market. Finding use in Internet of Things and Big Data for enterprise opens doors for development in various business functions. The large volume, velocity and variety of data presents a "goldmine" when it comes to obtaining significant information that has the potential to increase business performance. In this paper, the improvement of CRM systems through Internet of Things and Big Data is addressed. The paper presents a generic theoretical model of a CRM system that is based on Internet of Things and Big Data. This way an overview is provided when it comes to the development of CRM systems and their application in domestic enterprises. Modern CRM systems present an important part of conducting business in dynamic markets where there is a "battle" for every customer and retaining existing customers is an imperative. Overall, this paper provides a solid basis for future research in this domain.**

## INTRODUCTION

The modern business environment is characterized by the fast-paced distribution information and data from customer to customer, from business to business, and from customer to business. In such an environment customer relationship management (CRM) systems are becoming an imperative when it comes to achieving and maintaining a competitive position on the international and globalized market [1]. CRM systems, if implemented and applied correctly, can positively affect the development of appropriate relationships with key customers and market segments [1]. CRM in the context of the modern business environment focuses on activities and processes that define revenue streams and through demand can be optimized [2]. The strategic importance of CRM resonates through business performance indicators, thus enterprises have to find the most appropriate solutions when it comes to technology application within CRM. Internet of Things (IoT) and Big Data Analytics (BDA) possess the necessary framework for obtaining data, and for effectively processing the data for good relationship building with customers. The noted technologies have capabilities for ensuring benefits and positive outcomes in the customer-enterprise relationship [2].

Big Data Analytics can be viewed as the analysis of multidimensional data that is accessed through various forms of digital media. The majority of data comes from

the Internet of Things. As the development of hardware and software progresses at an immense rate, the available data that is being hyper-distributed across Internet platforms, requires larger and more sophisticated and high-powered infrastructures that can manage the analysis of such robust datasets [3]. More precisely, the application of Big Data Analytics requires the parallel advancement of the implement infrastructure alongside with the advancement of data that is distributed and obtained.

Furthermore, as noted earlier, the Internet of Things, which includes the wide use of devices and software that have the capability to record and capture information, as well as to distribute it in an automated-way [4]. It can be defined as interrelations between computing devices and physical objects and it involves a wide range of technologies that emerged in the digital revolution. The communication transpires between two or multiple devices on the Internet [4]. Internet of Things in the context of data driven CRM systems, can be viewed as part of Big Data analytics. More precisely, Internet of Things represents one of the main sources of data that is analyzed with the goal to develop and improve the customer-enterprise relationship.

This paper consists of four main sections (excluding the Introduction and Conclusion sections). The first section addresses Internet of Things and CRM. The second section Big Data Analytics and CRM. The third section presents the theoretical model of a CRM system that is based on Internet of Things and Big Data. The fourth section notes guidelines and suggestion regarding CRM system application with Big Data and Internet of Things.

## I. INTERNET OF THINGS AND CRM

The Internet of Things can be viewed as complex, intelligent and innovative network of interconnected devices and services where the main constructs are information-communication technologies (ICTs) [5]. This complex network of smart devices and internet-based services provide extensive and comprehensive information on customer experience, product information during and post- exploitation, and other type of metrics. Marketers can apply this information for developing better products and services with the goal to improve the relationships with customers [5].

Internet of Things is based on four key areas. These are sense (collecting large datasets from sensors that are implemented in various scenarios); communicate (refers to data communication and how the pre-existing platforms changed and transformed and that smart devices provide a new channel through which data is collected); analyze (includes the analysis of data collected from devices and from Internet platforms in real-time); and act (this refers to the action and closing a cycle, which is based on the collected, stored and analyzed data) [6].

Product development through Internet of Things can differ in type, and there are several main types such as new-to-the-customer, new-to-the-enterprise, broadening of existing product lines and brands, product modifications, innovations and improvements, remarketing of products and reducing costs [5].

As the globalization of markets has brought challenges to enterprises across industries, achieving a competitive edge on the international market is difficult and for the majority of enterprises practically impossible [7]. CRM systems in such a dynamic business environment represent an important technology segment that help organizations manage business processes in a manner that will result in better relationships with customers. Modern ICTs have made it possible for enterprises to apply sophisticated solutions at lower costs compared to a previous period when advanced ICTs and devices with large computing powers were exclusive to enterprises that could afford them due to their high sale price [6]. Today, the wide use of modern ICTs, opens doors for enterprises when it comes to CRM. More precisely, modern CRM systems, which are based on advanced ICTs, have the potential to collect significant data on about customers. From here, the CRM system is involved with large number of business processes including lead generation, data analytics, campaign management, solution customization, content management, and enterprise marketing management [6].

CRM systems based on the Internet of Things can be viewed as e-CRM that includes several key indicators and dimensions such as product and service customization, customer care and support, online community integration, convenience of product and service use, personalization of values, interactivity with customers. It is important to note that a lot of CRM initiatives are not successful, and just a handful of CRM strategies bring adequate results. Customer data is collected, stored and analyzed for achieving and improving sustainable profits, creating value for the enterprise and for the customer, and to improve customer satisfaction [8]. In addition, the current business environment and current digitalized markets require that enterprises adapt to the behavior of digital consumers. A CRM model that is based on online customer and consumer data can be applied in combination with consumer-related software and hardware and through this the management process can be improved which further translates into improved business performance [8].

Internet of Thing has changed the dynamics and relations between customer and enterprises. Customer networks are intensified through hyper-dissemination of information about products, services, enterprises, and

other customers' experiences. These digitalized customer networks are cornerstone of data collecting in the modern CRM systems that are based on the Internet of Things and other Internet platforms. [9]. An important segment of such CRM systems is the development of trust between customers and enterprises, as often sensitive data is collected. Overall, the Internet of Things in the function of CRM systems represents a complex network of customer data which is accessed. The collection and analysis of robust datasets resides under the domain of Big Data Analytics.

## II. BIG DATA ANALYTICS AND CRM

Big Data Analytics can be viewed as an integrated whole of data mining tools, techniques and technologies that are implemented into enterprises with the goal to extract valuable information about customer and consumer behavior with the goal to generate business reports. These reports include information on trending and future trends as well as insight on other metrics, with the goal to improve business performance [10]. The business performance improvement comes from an effective CRM system that is based on the noted data analysis. Big Data Analytics refers to the analysis of robust datasets, while the Big Data phrasing refers to the dataset itself. Such data is considered to be a crucial factor for innovation, technological improvement, strategy development, and business process improvement.

Big Data tools make massive analyses of large datasets possible. Big Data can provide significant and statistically acquired information on the existing relationships between internal and external business metrics. In this paper the main focus of Big Data is its application in CRM systems. Data collected with the purpose to support CRM systems is often categorical and this can be a challenge for extracting information for decision support. However, smaller and categorical datasets can be useful for determining specificities. Namely, larger datasets can annul some relations between indicators if they are not numerous. This can be an issue if the enterprise aims at these low probability events and occurrences. Therefore, Big Data Analytics has to be optimized and synchronized with the objectives of enterprise and the implemented CRM system. [11].

Modern CRM systems rely on Big Data for acquiring data that is high in volume, variety, veracity, velocity and provides value for the enterprise. Enterprises, both public and private, move towards big data solutions with the goal to extract value from digital sources such as social networks, smart devices, voice recordings, image processing, government data, consumer and customer activity online, and search engine data [12]. The market for this data is so vast, that there are numerous enterprises that specialize in collecting, storing and processing data from the noted sources. They package this data and sell it to enterprises who themselves don't have internal Big Data Analytics systems. After purchasing the data is applied for improving CRM systems. These CRM systems are the frontline of enterprises and require accurate data in order

to ensure efficient and effective support for developing better customer relationships.

Furthermore, Big Data is not without limitations and flaws. Some of the key noted and experienced flaws of such large datasets are vagueness (lack of volume threshold definition, and the lack of understanding what kind of data should be collected and what should be looked for in order to extract value); lack of managerial focus (there is no certainty when it comes to big data analytics use and its influence on financial metrics, and this may be due to data sources which are not, by design, suitable for direct analysis regarding its impact on business indicators); trivial significance (with the advancement of ICTs it was somewhat expected that the volume, variety and velocity of data will dramatically increase, however, with such data, traditional statistical approaches are not appropriate and statistical significance based on these approaches is not applicable with big datasets). [13].

### III. IMPROVING CRM AND BUSINESS PERFORMANCE

Enterprises implement and apply big data with the goal to gain advantage, technology resources, technological and innovative competence, support from top management, and regulatory environment. Customer big data has the potential to significantly improve and CRM performance and operation management efficiency. In addition, big data analytics can improve supply chain logistics and financial performance [14]. However, in the same study it was noted that even though due to the advancement of modern ICTs, collecting and storing customer data is not challenging as it used to be, this data does not contribute to the CRM system without proper analysis and processing. The process of data mining involves techniques, tools and methodologies to extract insightful and valuable information on the specific areas of interest. The high volume and high velocity data has to be effectively and efficiently analyzed and managed in order to extract operable value creation and to enable big data related strategies.

Improving CRM practices and overall business performance is mainly based on derived value from large datasets through the process of data mining. Customer behavior data can be categorized and graded in order to develop internal specific customer segmentation reports. Based on these reports, the enterprise can optimize and modify its strategy models and operational functions so it results higher customer satisfaction. [15].

Improving CRM and business performance through Internet of Things and Big Data requires different models of data mining and data analysis that depend on the set goals that the enterprise wants to achieve. These different models, or better say, different frameworks depend on the volume, type, velocity, variety of data that is being collected. Data can be collected online and offline as well. The analysis in this case has to include a comparative approach, too. CRM performance and overall business performance involves the level of how extracted value is applied in business processes and in CRM practices. Utilization of data depends not only on the dataset and the sophistication that involves its analysis, but also on

managerial processes that decide on specific actions that are to be conducted within business functions, including the CRM system. [16].

Big Data can additionally enhance business operations that further positively affects business performance through increasing support in decision making. This is particularly true for supply chain management. Data from supply chains can be collected and analyzed in manner that provide insight on how to optimize the distribution and logistics. This approach has a tremendous impact on lowering costs and increased efficiency of product and service delivery, which positively affects customer relationships and contributes to the effectiveness of CRM practices [17].

Business performance and CRM performance and directly and indirectly affected by the application of Big Data Analytics. The indirect improvements come the application of data mining in other business functions including marketing (location-based marketing; micro-segmentation of markets; sentiment and trend analysis; and multichannel performance improvement); merchandising (pricing optimization, design optimization, placement optimization, assortment optimization); supply chain (distribution optimization, supplier negotiation, logistics improvement, inventory management); operations (transparency of conducted actions, labor input management); new business models (web-based markets and market segmentation; price comparison services [18]. A theoretical model of a CRM system that is based on Internet of Things and Big Data is presented on Figure 1.

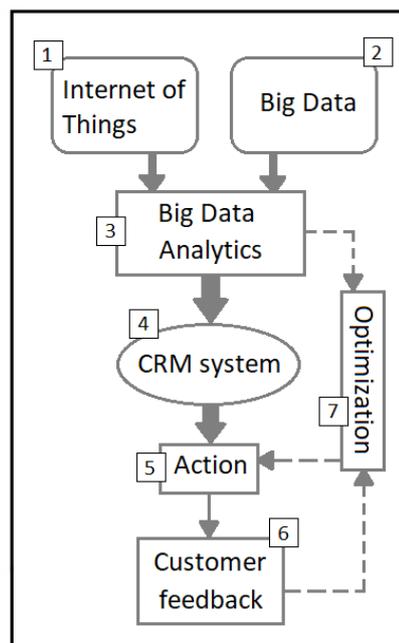


Figure 1. CRM system based on Big Data and Internet of Things

The model presented on Figure 1. has seven main elements. The first two elements are Internet of Things (1) and Big Data (2). As noted earlier both of these include large amount of networks and devices from which data can be collected. After this data is collected it is not yet in an adequate form to be applied in the enterprise. Therefore, the data has to be analyzed, categorized, and

extracted in a manner that it provides valuable information on the subject of interest. This part of data analysis and information extraction is conducted in the third element of the model - Big Data Analytics (3). Afterwards, the obtained valuable information is applied within the existing CRM system (4). The application can be in the form of decisions and actions. The action part (5) refers to the enterprises operational and strategic actions that aim at improving the relationship with customers. Next, the results of specific actions are evaluated through customer feedback (6). Based on the feedback, comparisons are made with the existing actions, plans and strategies, and optimization is conducted where necessary and applicable. This way a continuous loop of optimization and improvement can assured, which can significantly affect the outcome of business processes.

#### IV. SUGGESTION AND GUIDELINES

Based on the noted literature sources in the domain of Internet of Things, Big Data, and CRM systems, and in accordance with the developed theoretical model, the following suggestions and guidelines for domestic enterprises are noted:

- Existing CRM systems should be evaluated and weaknesses have to be identified. From here, the potential of Big Data and Internet of Things application can be taken into consideration as solutions or means of improvement.
- Managers have to define and identify existing relationships with customer as well in what percentages are these relationships the direct result of CRM practices. Based on the obtained information, managers can decide if CRM practices should be improved and modern ICTs implemented or a whole new CRM system approach is needed from bottom-up.
- Internet of Things and Big Data applications in CRM systems can be expensive depending on the needs of the enterprise. Thus, managers have to determine if the existing CRM system is sufficient, or the risk of implementing new technologies has positive expected value.
- The application of Internet of Things and Big Data in CRM systems can be through a third party service provider. These providers possess the necessary infrastructure and can extract information for clients (enterprises). However, for achieving maximum efficiency and efficacy, internal CRM systems with the noted technologies are preferable.
- An important aspect of CRM systems is data security. Enterprises have to conduct Big Data Analytics and use customer information in accordance with regulatory policies.

In sum, the Internet of Things and Big Data as two important sources of customer data can significantly

contribute to existing CRM systems. It is important to note that, like with any other modern technology, without an existing strong basis (in this case a well established CRM system), the implementation of a new technology doesn't guarantee success, rather it should be viewed as an enhancing mechanism.

#### V. CONCLUSION

Internet of Things and Big Data present tremendous sources of data that can help an enterprise to achieve its business goals. The theoretical model presented in this paper provides a concise insight on how Big Data and Internet of things can support CRM systems. This can further increase the intensity of customer relationship development and increase the percentage of retaining customers. It can be concluded that regardless of enterprise size and industry in which it operates, a CRM systems based on advanced technologies is imperative, especially in the new hyper-dynamic and information rich digitalized market. It is important to note, that such systems have to take into consideration privacy laws, and data security as inadequate handling of sensitive data can severely affect the enterprise.

The main limitation of the paper is lack of meta-analysis and comparison of studies conducted in this domain. Therefore, for future research, it is recommended to address similar studies in this domain and to conduct empirical analysis in enterprises that apply Internet of Things and Big Data Analytics in their CRM systems and CRM practices.

#### REFERENCES

- [1] S. Chatterjee, R. Chaudhuri, D. Vrontis, A. Thrassou, and S. K. Ghosh, "ICT-enabled CRM system adoption: a dual Indian qualitative case study and conceptual framework development", *Journal of Asia Business Studies*, vol. 15 no. 2, pp. 257-277, 2021. <https://doi.org/10.1108/JABS-05-2020-0198>
- [2] M. A. Ghazaleh, and A. M. Zabadi, "Promoting a revamped CRM through Internet of Things and Big Data: an AHP-based evaluation", *International Journal of Organizational Analysis*, vol. 28, no. 1, pp. 66-91, 2020. DOI 10.1108/IJOA-12-2018-1602
- [3] R. Math, "Big Data Analytics: Recent and Emerging Application in Services Industry.", *Big Data Analytics*, pp. 211-219, 2017. doi:10.1007/978-981-10-6620-7\_21
- [4] D. T. N. Hashem, "The reality of Internet of Things (IoT) in creating a data-driven marketing opportunity: Mediating role of customer relationship management (CRM)", *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 2, 2021.
- [5] B. Tariq, S. Taimoor, H. Najam, R. Law, W. Hassan, and H. Han, "Generating Marketing Outcomes through Internet of Things (IoT) Technologies", *Sustainability*, vol.12, no.22, p. 9670, 2020. <https://doi.org/10.3390/su12229670>
- [6] S. Lokuge, D. Sedera, T. Ariyachandra, S. Kumar, and V. Ravi, "The next wave of CRM innovation: Implications for research, teaching, and practice", *Communications of the Association for Information Systems*, vol. 46, no.1, p. 23, 2020.
- [7] M. Bakator, D. Đorđević, and D. Čočkalović, "Developing a model for improving business and competitiveness of domestic enterprises", *Journal of Engineering Management and Competitiveness (JEMC)*, vol. 9, no. 2, pp. 87-96, 2019.
- [8] A. Farmania, R. D. Elsyah, and M. A. Tuori, "Transformation of CRM Activities into e-CRM: The

- Generating e-Loyalty and Open Innovation", *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 27, no. 2, p. 109, 2019. <https://doi.org/10.3390/joitmc7020109>
- [9] Y. Yan, C. Huang, Q. Wang, and B. Hu, "Data mining of customer choice behavior in internet of things within relationship network", *International Journal of Information Management*, 2018. doi:10.1016/j.ijinfomgt.2018.11.013
- [10] S. Sharma, "Big Data Analytics for Customer Relationship Management: A Systematic Review and Research Agenda", In *International Conference on Advances in Computing and Data Sciences*, pp. 430-438, 2020. Springer, Singapore.
- [11] L. González-Serrano, P. Talón-Ballesteros, S. Muñoz-Romero, C. Soguero-Ruiz, and J. L. Rojo-Álvarez, "A Big Data Approach to Customer Relationship Management Strategy in Hospitality Using Multiple Correspondence Domain Description", *Applied Sciences*, vol. 11, no.1, p. 256, 2021. <https://doi.org/10.3390/app11010256>
- [12] M. Anshari, M. N. Almunawar, S. A. Lim, and A. Al-Mudimigh "Customer relationship management and big data enabled: Personalization & customization of services", *Applied Computing and Informatics*, 2018. doi:10.1016/j.aci.2018.05.004
- [13] P. Zerbino, D. Aloini, R. Dulmin, and V. Mininno, "Big Data-enabled Customer Relationship Management: A holistic approach", *Information Processing & Management*, vol. 54, no. 5, pp. 818–846, 2018. doi:10.1016/j.ipm.2017.10.005
- [14] C. Zhang, X. Wang, A. P. Cui, and S. Han, "Linking big data analytical intelligence to customer relationship management performance", *Industrial Marketing Management*, vol. 91, pp. 483–494, 2020. doi:10.1016/j.indmarman.2020.10.012
- [15] W. N. Wassouf, R. Alkhatib, K. Salloum, & S. Balloul, "Predictive analytics using big data for increased customer loyalty: Syriatel Telecom Company case study", *Journal of Big Data*, vol. 7, no. 2, pp. 1-24, 2020.
- [16] W. K. R. Perera, K. A. Dilini, and T. Kulawansa, "A Review of Big Data Analytics for Customer Relationship Management", 2018 3rd International Conference on Information Technology Research (ICITR), 2018. doi:10.1109/icitr.2018.8736131
- [17] Q. Li, and A. Liu, "Big Data Driven Supply Chain Management", *Procedia CIRP*, vol. 81, pp. 1089–1094, 2019. doi:10.1016/j.procir.2019.03.258
- [18] L. T. N. Dinh, G. Karmakar, and J. A. Kamruzzaman, "Survey on context awareness in big data analytics for business applications", *Knowl Inf Syst*, vol. 62, pp. 3387–3415, 2020. <https://doi.org/10.1007/s10115-020-01462-3>

# Achieving Sustainable Development Through Information Systems

Mihalj Bakator\*, Dragica Radosav\*, Nataša Đalić \*\*, Mila Kavalić \*\*\*, Dragana Milosavljev\*

\* University of Novi Sad, Technical Faculty "Mihajlo Pupin", Đure Đakovića bb, 23000, Zrenjanin, Republic of Serbia

\*\* University of East Sarajevo, Faculty of Transport and Traffic Engineering, Vojvode Mišića 52, 74000, Dobo  
Republika Srpska

\*\*\* University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21102, Novi Sad, Republic of Serbia

mihalj.bakator@uns.ac.rs, dradosav@tfzr.rs, natasa.djalic@sf.ues.rs.ba, milakavalic@gmail.com,  
dragana.milosavljev@tfzr.rs

**Abstract - Sustainable development and conducting business while taking into consideration the sustainable development goals (SDGs) presents a new reality of the modern business environment where socio-economic dimensions are affected by ecological dimensions due to the increasing negative effects of resource overexploitation. As the majority of enterprises focus on increasing profits, the negative impact on the natural environment takes its toll in various forms (climate change, fauna and flora destruction, pollution etc.). Therefore, enterprises should consider changing and adapting their business model to fit within the context of sustainability, and to take into consideration the SDGs. Sustainability and sustainable development indicators are important metrics that shape future business performance and competitiveness. In this paper, the application of information systems in the context of sustainability is analyzed. The main goal of the paper is to provide an overview on information system application in the context of sustainability. In addition, based on the literature review, suggestions and guidelines regarding information system application and sustainable business development are proposed. The paper provides a solid base for future research in the domain of information systems and sustainability.**

## INTRODUCTION

Information systems can be viewed as an integration of hardware and software that are implemented and applied in enterprises. The main goal of information systems is the collection, storage, evaluation, analysis, creation and distribution of data. The obtained information from such analyses can be applied for improving management, control, coordination and optimization of business processes. This type of data dissemination can significantly help and support both employees at lower hierarchical levels in enterprises as well as employees in middle to upper management in making decisions or improving overall business performance. Information systems can be viewed from the aspect of its components and from the aspect of its roles in the enterprise [1].

Sustainable development can be presented as the means of development that meets human needs with available resources, without endangering natural systems and the environment, thus ensuring the long-term existence of human society. The concept of sustainable

development represents a new strategy and philosophy of social development. Unfortunately, the majority of enterprises are focused on increasing profits, productivity and overall enhanced business performance. This often leads to sharp rises in resource overexploitation rates on micro (local communities, regions) and macro (national, international) levels [2]. Sustainable development goals (SDGs) that address seventeen dimensions of sustainability require enterprises to address the concept of environmental efficiency as well as impact indicators that show the volume and intensity of negative or positive impact of business processes on the environment. The application of modern information systems can improve how enterprises conduct their business from the aspect of sustainability and resilience. This resilience refers to the resilience of enterprises to the continuous and dynamic changes on the market, as well as on the resilience of the environment that is affected by business processes. Resilience in both cases can be viewed as the ability to cope with sudden changes (exploitation shocks, sharp trend changes on the market). Information systems can increase resilience through dampening the intensity of changes as enterprises can prepare for them. Similarly, if there is enough data on environmental effects of a new business strategy, the potential harm can be predicted, the negative effects can be reduced, effectively increasing the resilience of the environment [3].

In this paper, the potential of information system application in the context of sustainability is analyzed. The paper consists of four main sections (excluding the Introduction and Conclusion sections). The first section provides an overview on sustainable development. The second section addresses information systems, while the third section notes the application of information systems in the context of sustainable development. The fourth section notes suggestions and guidelines when it comes to information system application for sustainable development on a national level and on the level of enterprises.

## I. SUSTAINABLE DEVELOPMENT

The first association to sustainable development of people is usually environmental protection, ecology, recycling and renewable energy. Some of the biggest

problems that sustainable development addresses are global warming, ozone depletion, the "greenhouse effect", the disappearance of forests, the conversion of fertile land into deserts, the appearance of acid rain, the extinction of animal and plant species. Additionally to the ecological and environmental aspect of sustainable development, it also implies the balance between resource consumption and the ability to regenerate natural systems. It is also important to note that the goals of sustainable development are interconnected, so that when achieving one goal, there is a positive influence on the other goals of sustainable development. Similarly, if there is a lack of improvement in one of the sustainable development goals, this can negatively affect other sustainable development goals. There are a total of 17 sustainable development goals and these are: 1) A world without poverty; 2) A world without hunger; 3) Good health; 4) Quality education; 5) Gender equality; 6) Clean water and sanitary conditions; 7) Available and renewable energy; 8) Decent work and economic growth; 9) Industry, innovation and infrastructure; 10) Reduction of inequality; 11) Sustainable cities and communities; 12) Responsible consumption and production; 13) Climate action; 14) Life under water; 15) Life on earth; 16) Peace, justice and strong institutions; 17) Partnership to goals [4].

The way enterprises manage waste and how they treat the natural environment affects how natural resources are managed in the long-term. Through sustainable development economic growth is optimized in order to be inline with environment principles. There are 4 key principles of sustainable development. These are maintaining the same volume of production; maintenance of unchanged stocks of natural resources; setting a safe minimum standard; application of operational environmental principles [5]. The first principle mostly refers to the consumerist society. Society is often not guided by needs but by desires and impulses. People buy not only what they need but also everything they want and as much as their possibilities allow. This leads to overproduction and mass production of surpluses, which often end up unused in some form of waste. This can further have a negative impact on sustainable development. It also directly affects another principle that deals with the management of natural resources. It often happens that resources are wasted without any worries or thoughts about the future. Society needs to understand that of certain resources there is only a limited amount and that some do not belong to the group of renewable resources. An alternative should be sought for such resources and their current consumption reduced [6].

Sustainable development as an imperative for enterprises that conduct business in a globalized market, requires exceptional efforts from management when it comes to improving business to the extent that it is in line with standards that have the function of maintaining and protecting the natural environment [7]. As management information systems are based on information systems and computer infrastructure, they have the ability to report to management on business results, resources expended, and even on employee satisfaction. In this way, management can make adequate operational and strategic business decisions that will further affect the competitiveness of the

enterprise. If a superior solution in the form of information system management is developed and implemented, then there are higher chances of achieving a competitive market position.

## II. INFORMATION SYSTEMS

Information systems specialized for decision making and enterprise management belong to the concept of information system management. These systems collect, store and process large datasets from the internal and external business environment. Based on the processed data, market trends and other indicators that can affect business performance are identified and defined [8]. By identifying and defining the situation in which the enterprise is in relation to the market, managers can make decisions aimed at improving business performance and achieving a more competitive position in the market.

The components of information systems represent the core structure of these systems and as such are an integral part that cannot be left out. The five components of information systems are hardware, software, data, human resources, or more simply, people and processes. Hardware, software and data fall under technology. In addition to the above five components, due to the expansion of the application of ICT, communication as an exchange concept can also be considered a component of information systems [9].

The first type are Executive Support Systems (ESS), and are used predominantly by executives in the strategic decision-making process [9]. The next type are Management Information Systems (MIS). These information systems focus on the internal business environment where data is collected on business processes that are difficult to improve [10]. Managers can also apply management information systems to strategic planning.

The third type of information systems are Decision Support Systems (DSS). These systems focus exclusively on data collection, storage and processing in order to identify and define alternatives and make effective and efficient decisions for certain business processes [11]. Furthermore, the fourth type of information systems are Knowledge Information Systems. These systems are implemented and applied in order to create and exchange information that creates and develops intellectual capital within the enterprise [11].

The fifth type of information systems are Transaction Processing Systems (TPS). These systems are designed to increase the efficiency and effectiveness of the implementation of transactions that are by nature repeated several times within business processes. They can be used to automatically issue invoices, pay bills, pay taxes, create reports [12]

The sixth type of information systems are Office Automation Systems (OAS). These systems aim to increase and improve the efficiency and effectiveness of employees who process data and information. These information systems are mostly in the form of various software applications and integrated solutions [12].

### III. INFORMATION SYSTEMS AND SUSTAINABLE DEVELOPMENT

From the perspective of sustainable development, information systems often focus on "green IT" rather than direct on sustainable development. However, information systems possess the capacity to address multiple dimensions of conducting business. These are economic dimensions (increasing profits, reducing costs, growing revenue), social dimensions (wellbeing of local communities and engaging in corporate social responsibility), and environment dimensions (preservation of the environment, reducing and annulling overexploitation of resources, and reducing the degradation of the environment that sustains society) [13].

The new technological revolution and the digitalization of social and economic dimensions has led to big changes that affect the environment. The application of information systems can provide an interdisciplinary approach to resolving environmental issues through sustainable practices. Information systems can be integrated into organizational and economic mechanisms that embrace sustainable development [14]. Information systems can be implemented into various organizations with a wide variety of functions. For example, sustainable living was addressed through information system-based village indexing where parameters such as bank facilities, transportation, water distribution, electricity distribution, road infrastructure and other services were analyzed. Through such system, allows a more effective and efficient way of resolving issues, thus improving several aspects of the sustainable development goals [15].

Another approach of achieving and improving several sustainable development goals at once includes the application of information systems in supply chain management. This approach includes the integration of raw material selection, product design, delivery of raw materials, delivery of end products to the market, manufacturing processes, end-of-life product management, and overall it includes the integration of capital flows, material flow and information distribution [16]. On a broader scale, information systems along with advanced ICTs can significantly improve agriculture in a sustainable manner [17]. Furthermore, improving education can be achieved in the domain of computer science with the addition of sustainable development goals. Integrating information systems could increase the learning experience and improve education quality [18].

Finally, information systems application with modern ICTs are the cornerstone of a sustainable information society. The concept of information society is aimed at by numerous countries as such societies have higher social and economic growth compared to countries that didn't advance to higher levels of ICT application, or better say, are not fully operating in within the context of information society [19].

### IV. GUIDELINES AND SUGGESTIONS

Based on the analyzed literature in the domain of information system application in the context of sustainability, and taken into consideration the sustainable

development goals, the following suggestions and guidelines for domestic enterprises and the domestic economy are proposed:

- Enterprises have to evaluate their business processes and business strategies in accordance with the main sustainable development goals. After identifying critical issues, potential solutions should be addressed.
- Modern information systems should be considered for specific business functions in order to make it sustainable and reduce overexploitation.
- National incentive programs could be introduced for enterprises that modify and conduct their business in accordance with sustainable development.
- Achieving sustainability should not be only on paper. More precisely, organic, long-term sustainable strategies are an imperative, as short-term solutions are not adequate.
- Information systems can be implemented into various socio-economic and ecological dimensions. On a national level governance services should be centralized, and main governance bodies should be connected with local communities and their governance bodies.
- Information systems can be applied for national competitiveness indicator analysis and segment competitive ability into smaller parts that all the way to single enterprises.
- The implementation of information systems on a national scale should be conducted simultaneously across all urban and rural areas. Integration into a functioning information system on such a large scale requires adequate implementation percentages across the country.
- Enterprises have to identify what negative effects are they causing in which sustainable development category and modify their business process accordingly.
- One of the main problems of sustainable development is the lack of awareness among people. It is a consequence of omission in education. Raising the general educational level of the population can contribute to solving future environmental problems and provide a stronger basis for sustainable development.
- Sustainable development as a new model of society development and its interaction with nature is difficult to imagine without its reliance on science and education. In fact, the idea of the priority of education is the core of

the transformation according to the model of sustainable development of society.

Overall, information systems have the potential to improve the process of achieving sustainable development in various areas and enterprises along with governance bodies on a national level should evaluate the possibilities of information system implementation in the context of sustainability.

## V. CONCLUSION

Sustainable development and sustainable development goals require strategic approaches from enterprises and effective actions on a national level. Information systems have the potential to increase the process of achieving sustainability across the seventeen goals of sustainable development. It can be concluded that information systems, if applied correctly, can improve and enhance business process and public services in a sustainable manner across industries, including but not limited to SMEs, manufacturing enterprises, law, governance, environment preservation etc. The necessity for sustainable development and its role in socio-economic and ecological dimensions will only grow. National strategies have to take place in order to optimize economic growth and environmental preservation. Enterprises have to consider sustainable business models in order to become and stay competitive on the international market.

The main limitation of this paper is the lack of enterprise data analysis. However, the main aim of the paper was to review sustainable development and information system application along with noting suggestions and guidelines in this domain. For future research it is recommended to address sustainability indicators in enterprises. In addition, a structured meta-analysis can be conducted across industries.

## REFERENCES

- [1] K. Stendal, D. Thapa, and A. Lanamäki "Analyzing the concept of affordances in information systems" In 2016 49th Hawaii International Conference on System Sciences (HICSS) pp. 5270-5277. IEEE. 2016
- [2] N. Skorobogatova, "Sustainable Development of an Enterprise Under Industry 4.0 Conditions," 2019 International Conference on Creative Business for Smart and Sustainable Growth (CREBUS), 2019, pp. 1-5, doi: 10.1109/CREBUS.2019.8840049.
- [3] R. Heeks, and A. V. Ospina, "Conceptualising the link between information systems and resilience: A developing country field study," *Info. Sys. J.*, 2018 doi:10.1111/isj.12177
- [4] J. D. Sachs, G. Schmidt-Traub, M. Mazzucato, D. Messner, N. Nakicenovic, and J. Rockström, "Six transformations to achieve the sustainable development goals," *Nature Sustainability*, vol. 2, no. 9, pp. 805-814, 2019.
- [5] W. Leal Filho, L. L. Brandli, A. Lange Salvia, L. Rayman-Bacchus, and J. Platje, "COVID-19 and the UN sustainable development goals: threat to solidarity or an opportunity?," *Sustainability*, vol. 12, no. 13, p. 5343, 2020.
- [6] J. Hickel, "The contradiction of the sustainable development goals: Growth versus ecology on a finite planet," *Sustainable Development*, vol. 27, no. 5, pp. 873-884, 2019.
- [7] D. C. Duran, L. M. Gogan, A. Artene, and V. Duran, "The components of sustainable development-a possible approach" *Procedia Economics and Finance*, vol. 26, pp. 806-811, 2015. doi: 10.1016/S2212-5671(15)00849-7
- [8] S. Nowduri, "Management information systems and its support to sustainable small and medium enterprises," *International Journal of Business and Management*, vol. 7, no. 19, 125, 2012. doi: 10.5539/IJBM.V7N19P125
- [9] J. Valacich, and C. Schneider, "Information Systems Today: Managing the Digital World," 2017. Pearson
- [10] K. C. Laudon, and J. P. Laudon, "Essentials of management information systems," 2011. Upper Saddle River: Pearson.
- [11] S. Matayong, and K. A. Mahmood, "The review of approaches to knowledge management system studies," *Journal of Knowledge Management*, vol. 17, no. 3, pp. 472-490, 2013. doi: 10.1108/JKM-10-2012-0316.
- [12] V. Pathak, S. Wankhede, and K. S. Chandwani, K. S. "Office automation system. "International Research Journal of Engineering and Technology (IRJET), vol. 5, no. 3, 2446-2448, 2018.
- [13] C. Marnewick, "Information system project's sustainability capability levels" *International Journal of Project Management*, vol. 3, no. 6, pp. 1151-1166, 2017. doi:10.1016/j.ijproman.2017.02.014
- [14] S. A., Dyatlov, N. I. Didenko, O.S. Lobanov, and S. V. Kulik, "Digital transformation and convergence effect as factors of achieving sustainable development," In *IOP Conference Series: Earth and Environmental Science* vol. 302, no. 1, p. 012102, 2019. IOP Publishing.
- [15] S. Cholke, "Sustainable Development of Villages using Village Information System: A Geospatial Approach," *International Journal of Research*, vol 7, no. 5, 2018.
- [16] P. De Camargo Fiorini, and C. J. C. Jabbour, C. J. "Information systems and sustainable supply chain management towards a more sustainable society: Where we are and where we are going," *International Journal of Information Management*, vol. 37, no. 4, pp. 241-249, 2017. doi:10.1016/j.ijinfomgt.2016.12.004
- [17] B. Chaudhuri, and L. Kendall, "Collaboration without consensus: Building resilience in sustainable agriculture through ICTs," *The Information Society*, 2020. doi: 10.1080/01972243.2020.1844828
- [18] W. Nwankwo, and C. Njoku, "Sustainable Development in Developing Societies: The Place of ICT-driven Computer Education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 12, 290-297, 2020. Kassel, Germany: International Journal of Emerging Technology in Learning. Retrieved August 15, 2021 from <https://www.learntechlib.org/p/217538/>.
- [19] E. Ziemba, "The Contribution of ICT Adoption to the Sustainable Information Society," *Journal of Computer Information Systems*, pp. 1-11, 2017. doi:10.1080/08874417.2017.1312635

# Integration of Heterogeneous Data into Electronic Patient Records

Snezana Savoska\*, Blagoj Ristevski\*, Natasha Blazheska-Tabakovska\*, Ilija Jolevski\*,  
Andrijana Bocevaska\*, Vladimir Trajkovik\*\*

\* Faculty of Information and Communication Technologies – Bitola, University “St. Kliment Ohridski” – Bitola,  
North Macedonia

\*\* Faculty of Computer Science and Engineering, “Ss. Cyril and Methodius” University in Skopje  
North Macedonia

{ snezana.savoska, blagoj.ristevski, natasa.tabakovska, ilija.jolevski, andrijana.bocevaska }@uklo.edu.mk,  
vladimir.trajkovik@finki.ukim.mk

**Abstract** - The deficiency of data for patients with chronic diseases and other diseases and previous medical treatments shows a significant weakness with many patients. Typically, due to the healthcare system insufficiently, patients with comorbidities might not survive the diseases, especially when the disease is novel. The lack of information on genetic disorders in patients they are not aware of also contributes to increased patient deaths. This leads to the need to integrate medical and health data with various biological and other data, especially in pandemic circumstances and the increasing number of patients with chronic diseases. Patients' health data issues are evident, but they are stored in various hospital and public health systems such as electronic health records (EHRs), healthcare institutions, and laboratories. Furthermore, biological data are often not integrated and cannot be used by patients, physicians, and specialists to treat particular diseases. Although the urgent need for healthcare and medical data integration is apparent, personal data protection laws are rigorous. They do not allow much progress in the field without implementing patient healthcare data security and privacy standards. One solution for this issue is establishing a personal health record (PHR) as an integrative system for the patient. Many ontological frameworks have been proposed to unify the record formats, but none of them is accepted as healthcare standards. The efforts towards approving the HL7 standards and the well-known medical coding systems promise future data integrations. Also, some attempts are made to associate particular diseases with data obtained from external environmental sensors that measure disease-related data. Using these data, called exposure data or exposome, one can clarify the increasing symptoms of particular diseases influenced by external factors. This paper proposes a cloud-based model for integrating healthcare and medical data from different sources as EHR, health information systems, and measurement sensors into PHR as the first stage towards integrating patient health data. The medical data, PHR, numerous biological and exposome data, and data obtained from wearables are considered and stored on the cloud following the required data security and privacy standards.

**Keywords**- Electronic Health Records, Personal Health Records, Internet of Medical Things, Data Integration.

## I. INTRODUCTION

The streams of healthcare and medical data nowadays are stored in diverse hospital information systems (HIS) providing protected data for patients in the form of EHR, medical prescription, diseases diagnoses and treatments. These data are owned by many healthcare providers and usually are not accessible to the patients. Taking into account the trends of the population ageing as well as the need for medical care that increases all the time the patients use some wearables according to the concept of IoT and its part Internet of Medical Things (IoMT). They use these wearables for health purposes, according to the concept of Ambient Assisted Living (AAL) that produce many different types of data for the patients' live conditions and some health parameters. These wearables (sensors) usually are connected with the measurement of some healthcare parameters and collect data for human behavior and health parameters and conditions. If we consider also the exposome concept and all data collected from the environmental pollutions that influence healthcare, we can realize that there are wide amounts of data connected with human health. Recent research shows the increasing interest in the usage of personal patient's data in the residence where the patient resides at the moment, providing accurate data for healthcare staff that can provide healthcare at the moment when the patient needs medical help and care. The development of artificial intelligence (AI) techniques also can provide a mechanism for healthcare risk assessment when the environmental and chemical exposure data, known as exposome, and other omics and medical data are available. This fact and the usage of this kind of data lead to the need for healthcare and medical data integration.

The proposed solution for this very complex issue is connected with the implementation of strong security standards for healthcare and medical data is the creation of a patient-centric system with Personal Health Record (PHR). This concept considers the responsibility of the patients to secure their PHR and the possibility to share with a selected medical person, temporarily. This concept demands a complex cloud-based architecture as well as the education of the medical staff and patients for

increasing digital healthcare literacy. This kind of integrated system also has to support the usage of many wearables connected with mobile applications for measuring and monitoring the vital signs of life, connected with wireless sensors. In addition, the concept has to support input of unstructured data types as healthcare history, medical images, lab results as well as some biological data or data from HIS, EHR, data owned by healthcare providers.

The second challenge nowadays is the patient's healthcare risk assessment. This risk assessment demands the usage of PHR, omics and exposome data [3] that have to be accessible for boots (software agents) for risk assessment. Such complex digital structure is created in a cloud environment, taking into account the security and privacy data protection issues, differ from one country to other. All these different types of collected data, structured, semi-structured and unstructured, have to be saved and acquired according to well-known standards in healthcare as HL7, FHIR, openEHR [16] and to use coding systems as ICD10, standard key terms, such as Medical Subject Headings and SNOMED.

The paper proposes a model of integration of healthcare and medical data in a patient-centric electronic PHR cloud-based system, connected with previously mentioned biological and medical data, omics and exposome data (Fig.1). The purpose is to provide the possibility of combining PHR data with others data types. The paper considers some related works connected with the areas of interest – the attempts for healthcare personal data integration. The third section explains the proposed model of integration of healthcare and medical data and the possibilities for semantic integration of medical, clinical, patient-oriented data, as well as sensors' and exposome data. The final section validates the model and gives some recommendations for future works.

## II. RELATED WORKS

Some many attempts and papers that describe some efforts to integrate healthcare, HIS, clinical and medical data to provide healthcare data analysis intended for healthcare decision-makers. Nevertheless, all these heterogeneous data are stored in many different places, formats and heterogeneous platforms and their integration is a very challenging and demanding task. We can mention Silvestri et al. in [10], in which they proposed a big data architecture for big data analysis considering EHR and PHR data from structured and unstructured HIS document, considering a model for decision makers' support according to the national law, using Spark, MongoDB and DL-bases AI module for NLP. An open data integration platform for patient, clinical, medical and historical data, stored across multiple HISs, is proposed in [13], in which they address patient-centred healthcare and clinical decision support requirements. In this model, they integrate further heterogeneous data sources such as data streams generated by wearable IoT devices. The distribution of scanned documents at one health institution and the design and evaluation of a system to categorize documents into clinically relevant and non-clinically relevant categories as well as further sub-classifications were described in [14]. Serbanati presents

a method for digitizing the concept of health by processing the existing information in EHRs with the help of several dedicated services [20]. Precision Medicine includes the discovery of a patient-specific pattern of disease progression and a determination of the precise therapy for that pattern, and the corresponding personalized delivery of care [7]. An IoMT platform for pervasive healthcare that ensures interoperability, quality of the detection process, and scalability in a machine-to-machine-based architecture and provides functionalities for the processing of high volumes of data, knowledge extraction, and common healthcare services, was proposed in [15].

In [8], the authors had demonstrated the feasibility of a scalable, accurate, and efficient approach for medical device surveillance using EHRs, presenting that implant manufacturer and model, implant-related complications, as well as mentions of post-implant pain can be reliably identified from clinical notes in the EHR. Liang et al. had identified three threats from real cloud-based eHealth systems, i.e., privacy leakage, frequency analysis, and identical data inference [6].

Koren, Jurčević and Huljenić aimed at investigating wireless sensors in the IoT context in contemplation of model solutions in the field of eHealth [8]. Shah and Khan outlined various secondary uses of EHR to give an idea of how effectively EHR data can be used in different domains such as clinical research, public health surveillance and clinical audits to provide effective, timely and quality healthcare facilities to the patients [5]. Gamal, Barakat and Rezk discussed different database models' appropriateness for integrating different EHRs functions with different database specifications and workload scenarios [17]. An application of an unsupervised machine learning approach in discovering latent disease clusters and patient subgroups using EHR data was described in [18]. A knowledge-driven framework able to transform disparate data into knowledge from which actions can be taken to help clinicians and data practitioners in the complex tasks of extracting valuable knowledge from heterogeneous datasets is described in [19].

Saripalle et al. have described a tethered PHR that seeks to achieve interoperability by using open-source standards and their implementation [9]. Warner and Levy had debated several emerging paradigms for integration including non-standardized efforts between individual institutions and genomic testing laboratories, "middleware" products that portray genomic information, albeit outside of the clinical workflow; and application programming interfaces that have the potential to work within clinical workflow [11].

The conclusion is that the mentioned models indeed try to connect EHR data and provide data integration for decision-makers or some national-wide integration intended to high-level decision-makers in healthcare and medicine. The security issues are considered from the aspect of the patient and the patient's country of living [2]. First, the model is a cloud-based cross-border healthcare system based on the PHR concept with an e-health strategy. The key point is that data collection can

be made out of hospitals and HIS and it is not obligatory to be connected with EHR and country of living. This concept demands also an increase in e-health and health digital literacy and can support the national and local medical and healthcare authorities [1] [4]. In addition, data integration should be wider and has to provide data integration not only for data analysis and decision-making but to provide wider integration of a patient-centric electronic PHR system with the possibility to connect them with biological and medical data. The purpose is to provide the possibility of combining PHR data with omics data as well as with open access data from exposome data that affect human health. This concept demands solving the privacy issues of the level of PHRs.

### III. MODEL OF HEALTHCARE DATA INTEGRATION

The model of a PHR-centric integrated healthcare system is built according to HL7 standards with included security and privacy concerns, especially in the part of PHR, e-prescription and the e-referral system [4]. Taking into account that this concept includes the usage of sensors for measuring the vital signs of life, according to the IoT concept, connected with mobile applications for patients and medical practitioners, many security and privacy issues have to be considered. There have to be a couple of scenarios of data integration with the roles of physicians, patients and pharmacists. Patients who have their PHR in the system can have their data stored according to HL7 and FHIR standards and provided privacy and security from cloud systems. The patient temporarily can grant access to their data to medical staff who is also registered in the system. Patients also can use PHR and mobile applications for citizens to collect healthcare data in their PHR.

The doctors can use mobile applications connected with the measurement of a patient's vital signs of life with sensors for professionals to provide data and to collect them in the PHR. All sensitive data have to be protected and secure in a cloud environment [1]. Users' data also contain sensitive data such as personal information, health family history, medical and healthcare data, protected properly according to national regulation regarding the personal data protection law in the intended country. Some scanned unstructured data can be entered in the PHR also and they can be accessible to the medical staff. Labs and biometrics reports also can be input into PHR as unstructured documents.

In the next stage, some medical and omics data can be connected with PHR and related to some diseases. These data can be analyzed by clinicians in order to provide some genotype, phenotype and metabolic data related data with some diseases. In addition, some soft data related to healthcare can be provided and integrated into PHR by a patient using environmental, social media and other data known as exposome data.

A suitable structure has to be provided to have available data for healthcare risk assessment for disease taken from PHR and environmental data, connected with location as well as some social media and stress connected data. This extremely complex task has to

include data analysis and complex algorithms, AI and medical knowledge as well as risk factor analysis.

When the integration of healthcare and medical data has to be considered, security issues are very important. Security issues can be classified into two main categories: information security and system security. Information security usually means data encryption, data integration and authentication. System security is connected with some administrative, technical and physical security levels. We can use the Authentication, Authorization and Accounting server (AAA) that have to be integrated with attribute authority (providing Role and Routing information). The encryption also has to be used as well as access control to protect and enable PHR security and privacy for the patients' country of origin. Prevention of unauthorized access to sensitive patient data at rest is provided by data encryption. AAA layer of the system uses the Keycloak Server as an Open Source Identity and Access Management server. Authenticated users (doctors and pharmacists), according to the level of digital identity assurance and patient consent, will have access to this information.

Standardization is provided to prevent malicious system misusing and has to enable security access protocols, intrusion detection and prevention techniques, providing SIEM systems, with audit logs of the users and administrator activities [2]. The proposed model is presented in Fig.1. This model relies on high-level security and privacy and provides adequate access to data for the appropriate user. In the proposed model, the first step to proper user orientation to the appropriate resource is the AAA. It uses the Keycloak server to check type, credential and the affiliation of user access. The first check for secure access is verifying the authentication - username and password to check if the user has the right to access. If the user (patient or doctor) is authenticated, the authorization check is performed, the role of the user is determined. The last step in the AAA framework is user accounting that measures the resources the user consumes during access.

In this first level of security control, the user is redirected to the appropriate control server in the appropriate domain in the country of origin or affiliation. Distribution is transparent to users, only one unique and integral location for the API URL is used by applications and end-user integrations, regardless of the origin of the request. Upon completion of the authentication and authorization procedure, the client receives an authenticated token that can be used to access the API endpoints and through them access the EHR data. Because user access data is disaggregated based on affiliation, specifically on the user's country of origin, this user identification and authorization data is stored on the federal (shared) server in the respective country and is used. The user can be assigned the appropriate role: a patient who owns the PHR data, the physician who can access and generate additional PHR data, and the pharmacist who can access only parts for e-prescription services. Role-based access control for accessing some part or while PHR data is defined in user roles. Subsystems that allow routing/redirection to appropriate



the concept of IoT, IoMT and sensors or wearables. The sensitive personal data are protected and secured properly. In the proposed model, specific medical and omics data are connected with PHR and related to diseases.

The other important advantage is the possibility for healthcare risk assessment connected with the usage of exposome data from environmental databases [12], social media, and other public data that can be connected with location and used by the software agents for healthcare risk assessment [21]. It can be very useful for personal healthcare risk assessment. All these depersonalized medical and healthcare data can be analyzed using suitable big data analytics tools for medical purposes, using disease data for disease group risk assessment.

#### ACKNOWLEDGEMENT

Part of the work presented in this paper has been carried out in the framework of the project "Cross-border initiative for integrated health and social services promoting safe ageing, early prevention and independent living for all (Cross4all)", which is implemented in the context of the INTERREG IPA Cross Border Cooperation Programme CCI 2014 TC 16 I5CB 009 and co-funded by the European Union and national funds of the participating countries.

#### REFERENCES

- [1] Snezana S. et al. (2020) Cloud Based Personal Health Records Data Exchange in the Age of IoT: The Cross4all Project. In: Dimitrova V., Dimitrovski I. (eds) ICT Innovations 2020. Machine Learning and Applications. ICT Innovations 2020. Communications in Computer and Information Science, vol 1316. Springer, Cham.
- [2] Snezana Savoska, Ilija Jolevski, Blagoj Ristevski, Natasha Blazeska-Tabakovska, Andrijana Bocevska, Boro Jakimovski, Ivan Chorbev, Vassilis Kilintzis, Design of Cross Border Healthcare Integrated System and its Privacy and Security Issues, In Proceedings of Computer and Communications Engineering, Workshop on Information Security 2019, 9th Balkan Conference in Informatics, Volume 13, 2/2019, pp.58-64
- [3] Savoska, Snezana and Ristevski, Blagoj and Blazeska-Tabakovska, Natasha and Jolevski, Ilija (2019) Towards Integration Exposome Data and Personal Health Records in the Age of IoT. 11th ICT Innovations Conference 2019, 17-19 October, Ohrid, Republic of Macedonia.pp.237-246.
- [4] Snezana Savoska, Ilija Jolevski, Architectural Model of e-health PHR to Support the Integrated Cross-border Services, In proceedings of ISGT conference 2018, pp. 42-49, Sofia, 2019.
- [5] Shahid Munir Shah , Rizwan Ahmed Khan, Secondary Use of Electronic Health Record: Opportunities and Challenges, arXiv:2001.09479v1 [cs.CY] 26 Jan 2020.
- [6] Jinwen Liang et al., Privacy-preserving range query over multi-source electronic health records in public clouds, *Journal of Parallel and Distributed Computing*, Volume 135, January 2020, Pages 127-139.
- [7] Sitapati A. et al, Integrated Precision Medicine: The Role of Electronic Health Records in Delivering Personalized Treatment, *Wiley Interdiscip Rev Syst Biol Med*. 2017 May ; 9(3).
- [8] A. Koren, M. Jurčević and D. Huljenić, "Requirements and Challenges in Integration of Aggregated Personal Health Data for Inclusion into Formal Electronic Health Records (EHR),"2019 2nd International Colloquium on Smart Grid Metrology (SMAGRIMET), Split, Croatia, 2019, pp. 1-5.
- [9] Saripalle R, Runyan C, Russell M. Using HL7 FHIR to achieve interoperability in patient health record. *J Biomed Inform*. 2019 Jun;94:103188.
- [10] Silvestri S. et al., A Big Data Architecture for the Extraction and Analysis of EHR Data, 978-1-7281-3851-0/19/
- [11] Warner J. L., Jain S.K., Levy M.A., Integrating cancer genomic data into electronic health records, Warner et al. *Genome Medicine* (2016) 8:113.
- [12] Barouki R. and all, Integration of human exposome with the human genome to advance medicine, *Biochimie* 152 (2018) pp.155-158.
- [13] Jayaratne, M., et al.. A data integration platform for patient-centered e-healthcare and clinical decision support. *Future Generation Computer Systems*, 92, 996-1008.
- [14] Goodrum, H., Roberts, K., & Bernstam, E. V. (2020). Automatic classification of scanned electronic health record documents. *International Journal of Medical Informatics*, 144, 104302.
- [15] S Rubí, J. N., & L Gondim, P. R. (2019). Iomt platform for pervasive healthcare data aggregation, processing, and sharing based on onem2m and openehr. *Sensors*, 19(19), 4283.
- [16] Ristevski, B., & Savoska, S., Healthcare and medical Big Data analytics. Applications of Big Data in Healthcare: Theory and Practice, 85, 2021.
- [17] Gamal, A., Barakat, S., & Rezk, A. (2020). Standardized Electronic Health Record Data Modeling and Persistence: A Comparative Review. *Journal of biomedical informatics*, 103670.
- [18] Wang, Y., Zhao, Y., Therneau, T. M., Atkinson, E. J., Tafti, A. P., Zhang, N., ... & Liu, H. (2020). Unsupervised machine learning for the discovery of latent disease clusters and patient subgroups using electronic health records. *Journal of biomedical informatics*, 102, 103364.
- [19] Vidal, M. E., Endris, K. M., Jazashoori, S., Sakor, A., & Rivas, A. (2019). Transforming Heterogeneous Data into Knowledge for Personalized Treatments—A Use Case. *Datenbank-Spektrum*, 19(2), 95-106.
- [20] Serbanati, L. D. (2020). Health digital state and Smart EHR systems. *Informatics in Medicine Unlocked*, 21, 100494.
- [21] Savoska, S., Fdez-Arroyabe, P., Cifra, M. et al. Toward the creation of an ontology for the coupling of atmospheric electricity with biological systems. *Int J Biometeorol* 65, 31–44 (2021).

# Derivation, Analysis and Simulation of Outage Performance of MIMO Multi-branch SC Diversity System in $\alpha$ - $\mu$ Fading and Co-Channel Interference Environment

D. Krstic\*, S. Suljovic\*\*, N. Petrovic\*, Z. Popovic\*\*\*, and S. Minic\*\*\*\*

\* University of Nis, Faculty of Electronic Engineering, Nis, Serbia

\*\* Technical College of Applied Studies, Pozarevac, Serbia

\*\*\* Higher Technical School of Professional Studies, Zvecan, Serbia

\*\*\*\* Department of Informatics and Mathematics, Faculty of Teacher Education in Prizren – Leposavic, University of Pristina, Kosovska Mitrovica, Serbia

suaadsara@gmail.com, dragana.krstic@elfak.ni.ac.rs, nenad.petrovic@elfak.ni.ac.rs, z.popovix@live.com, sinisa.minic@pr.ac.rs

**Abstract** – In this article, an analysis of outage performance of selection combining (SC) receiver exposed to  $\alpha$ - $\mu$  fading and  $\alpha$ - $\mu$  co-channel interference (CCI) is done. The expressions for probability density function (PDF) and outage probability ( $P_{out}$ ) are derived and presented in some graphs. Fading and CCI parameters impact is reviewed. Then, simulation software environment for modelling and planning of MIMO system with multi-branch SC receiver working under influence of  $\alpha$ - $\mu$  fading and CCI, as described in previous part of the paper, is described. The aim of this procedure is to minimize the transmission costs and provide the best possible Quality of Service (QoS) for defined data transmission scenario.

## I. INTRODUCTION

Multiple input, multiple output (MIMO) is an efficient antenna technology for wireless communication systems where multiple antennas are used at both - transmitter and receiver [1, 2]. The antennas at each end of the system are combined to minimize errors, enlarge data rate and improve the channel capacity. This technique allows to radio signals to transmit over many different paths at the same time. This manner of transmission with multiple versions of the same signal provides possibilities for signals to arrive at receiver without fading and co-channel interference influence. This technique increases the signal-to-noise ratio and transmission quality, what creates more stable connections [3].

A few diversity combining schemes are used in the receiver to choose the best received signal. They are: maximal ratio combining (MRC), equal gain combining (EGC) and selection combining (SC). SC is not technique with the best performance, but is the least complicated and the cheapest, and therefore often used in wireless communication systems. Because of that is analyzed in many papers [4]-[7].

In last few year many general fading distributions are analyzed in available literature for describing fading in

wireless communications. For example, in [4], the first-order statistical characteristics of dual-branch SC diversity combiner, employed to reduce the influence of  $\eta$ - $\mu$  fading to the wireless telecommunications systems performance were analyzed. The authors determined probability density function (PDF), cumulative distribution function (CDF), signal moments, variance, characteristic function and channel capacity (CC) for defined system model.

The paper [8] introduces a general  $\kappa$ - $\mu$  fading distribution. This distribution includes the Rician, the Nakagami- $m$ , one-sided Gaussian and the Rayleigh distributions. Also, the log-normal distribution may be well-approximated by the  $\kappa$ - $\mu$  distribution. This distribution shows a very good fitting to experimental data. Many researchers know deal with  $\kappa$ - $\mu$  distribution. So in the letter [9], a novel analytical, closed form expression for the outage probability ( $P_{out}$ ), when useful and interference signal experience  $\kappa$ - $\mu$  fading in the presence of Gaussian noise is calculated.

The generalized  $\alpha$ - $\kappa$ - $\mu$  fading channel is analyzed in [7]. Closed-form expressions for probability density function (PDF) at the output of SC and MRC diversity systems are given. Also, closed-form expressions for Shannon capacity for cases of SC and MRC diversity systems, and for the case of no diversity, were determined.

In newer time, the  $\alpha$ - $\mu$  distribution was proposed for modeling the small scale fading variations. The  $\alpha$ - $\mu$  distribution is also a general PDF characterizing the signal envelope variation in multipath fading [10]-[18]. The property of generality of the  $\alpha$ - $\mu$  distribution motivates researchers to consider this distribution to model the small scale multipath fading environments, but a little complicated expression for PDF restricts the performance evaluation of this generalized channel model.

So, in [11], the expressions for the amount of fading (AoF) and the average CC for  $\alpha$ - $\mu$  fading channel were derived. These expressions can be reduced to the

expressions for other channel models, obtained from  $\alpha$ - $\mu$  fading as special cases [10], and that: Rayleigh distribution by setting  $\alpha=2, \mu=1$ , Nakagami- $m$  ( $\mu$  becomes a fading parameter  $m$  by putting  $\alpha=2$ ), Weibull ( $\alpha/2$  as a fading parameter if  $\mu=1$ ). The level crossing rate (LCR), the average fade duration (AFD), and the joint statistics of the correlated  $\alpha$ - $\mu$  random variables were analyzed in [10]. In [12], the multivariate distribution of the  $\alpha$ - $\mu$  distribution is derived. Further, the  $P_{out}$ , the moment generating function (MGF), and the bit error probability (BEP) evaluation were obtained in [13].

The mathematical expressions and simulation of the  $\alpha$ - $\mu$  channel model is described in detail in [2]. In [14], the problem of energy detection of an unknown deterministic signal over fading channels is observed. A closed-form expression is derived for the average probability of detection of the energy detector over  $\alpha$ - $\mu$  generalized fading channels with SC diversity reception.

In [15], a model for the complex  $\alpha$ - $\mu$  fading channel is presented. An efficient, simple, and general method to generate complex  $\alpha$ - $\mu$  samples is proposed and the bit error rate (BER) performance of an orthogonal frequency-division multiplexing (OFDM) system is derived. Channel parameters (nonlinearity, clustering, and power imbalance of components in-phase and quadrature) are analyzed.

An enriched  $\alpha$ - $\mu$  distribution as feasible model for fading in wireless communications systems is investigated in [16]. Some performance measures of wireless communications systems over this enriched  $\alpha$ - $\mu$  fading are determined in terms of the Meijer G-function and analysis of the system reliability investigated.

In [17], the authors analyzed joint influence of non-linearity, described by  $\alpha$ - $\mu$  distribution, and shadowing with inverse gamma distribution, to the wireless channel. An approximate expression for PDF was derived over  $\alpha$ - $\mu$ /inverse gamma fading channel by using Gauss-Laguerre quadrature polynomial. Further, this formula is used for derivation of CDF, moment generating function (MGF) and  $n$ th moment. After, the average symbol error probability (SEP) with MRC and EGC diversity is obtained.

The paper [18] investigates the secrecy performance of the Wyner's wiretap model, where the main channel and eavesdropper channel suffer correlated  $\alpha$ - $\mu$  fading. New expressions for the average secrecy capacity and secrecy  $P_{out}$  are derived for the considered scenario and the effect of correlation on the secrecy performance has been analyzed.

The average bit error rate of generalized spatial modulation (GSM)-Spatial multiplexing (SM) is calculated for fading channels described by  $k$ - $\mu$ ,  $\eta$ - $\mu$  and  $\alpha$ - $\mu$  distributions and compared with other existing schemes (GSM and spatial modulation) in [19]. It is proved that proposed scheme gives the best result.

Our paper is organized over five sections. In Introduction, we described papers from the area. In the second section,  $P_{out}$  of the signal to interference ratio (SIR) at the output of multi-branch SC receiver operating in  $\alpha$ - $\mu$  environment was derived. In the third one, we showed numerical results and analyzed parameters'

influence. Planning and simulation environment are presented in the fourth section. Finally, in Section V we interpret findings and conclude the paper with possibilities for future work.

## II. CALCULATING THE OUTAGE PROBABILITY OF SIGNAL TO INTERFERENCE RATIO AT THE OUTPUT OF THE MULTI-BRANCH SC RECEIVER

In this section, outage probability of the MIMO system which uses multi-branch SC receiver, is observed. The transmitted desired signal experiences  $\alpha$ - $\mu$  fading as well as co-channel interference. The model of the SC receiver is shown in Fig. 1.

The input signal envelopes are denoted as  $x_1, x_2, \dots, x_n$ . SC receiver has  $L$  input branches ( $L \geq 2$ ). SC receiver chooses the signal from the antenna with the highest SIR, and conducts it to the user. Output signal is denoted by  $x$ . The envelopes of co-channel interference appearing at the receiver input are  $y_1, y_2, \dots, y_n$  at each of the  $L$  SC receivers' inputs. A corresponding output CCI signal is  $y$ .

The PDF of the signal envelopes follow the  $\alpha$ - $\mu$  distribution given by [10]:

$$p_{x_i}(x_i) = \frac{\alpha \mu^\mu x_i^{\alpha\mu-1}}{\Omega_i^\mu \Gamma(\mu)} e^{-\mu \frac{x_i^\alpha}{\Omega_i}}, \quad (1)$$

where  $\alpha$  is a positive arbitrary parameter related to the nonlinearity of the wireless medium,  $\mu \geq 1/2$  is the inverse of the normalized variance,  $\Gamma(\cdot)$  is a Gamma function. In a described channel, CCI also follows  $\alpha$ - $\mu$  distribution:

$$p_{y_i}(y_i) = \frac{\alpha \mu^\mu y_i^{\alpha\mu-1}}{s_i^\mu \Gamma(\mu)} e^{-\mu \frac{y_i^\alpha}{s_i}}. \quad (2)$$

The mean square values of the signal envelopes  $x_i \geq 0, i=1,2, \dots, L$ , are denoted by  $\Omega_i$ , while the mean square values of the CCI envelopes  $y_i \geq 0, i=1,2, \dots, L$  are denoted by  $s_i$ :

$$\Omega_i(x_i) = \bar{x}_i^2, s_i = \bar{y}_i^2. \quad (3)$$

The ratio of the desired signal envelope and the CCI at the  $i^{th}$  input branch of the SC receiver is:

$$z_i = \frac{x_i}{y_i}, x_i = z_i y_i. \quad (4)$$

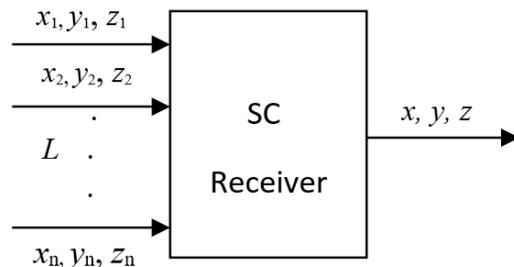


Figure 1. The model of the multi-branch SC receiver

The PDF of the desired signal to CCI ratio, SIR  $z_i$ , is given by [20]:

$$p_{z_i}(z_i) = \int_0^\infty dy_i y_i p_{x_i}(z_i y_i) p_{y_i}(y_i) = \frac{\alpha z_i^{\alpha\mu-1} (\Omega_i s_i)^\mu \Gamma(2\mu)}{\Gamma^2(\mu) (\Omega_i + s_i z_i^\alpha)^{2\mu}}. \quad (5)$$

The CDF of  $z_i$  is [21]:

$$F_{z_i}(z_i) = \int_0^{z_i} dt p_{z_i}(t) = \frac{\alpha (\Omega_i s_i)^\mu \Gamma(2\mu)}{\Gamma^2(\mu)} \int_0^{z_i} dt \frac{t^{\alpha\mu-1}}{(\Omega_i + s_i t^\alpha)^{2\mu}}. \quad (6)$$

The integral in expression (6) can be solved employing Beta function [20]:

$$\int_0^\lambda \frac{x^m}{(a+bx^n)^p} dx = \frac{a^{-p}}{n} \left(\frac{a}{b}\right)^{\frac{m+1}{n}} B_z\left(\frac{m+1}{n}, p - \frac{m+1}{n}\right),$$

$$z = \frac{b\lambda^n}{a + b\lambda^n}, a > 0, b > 0, n > 0, 0 < \frac{m+1}{n} < p. \quad (7)$$

By combining the expressions (6) and (7), we can rewrite the expression for CDF of  $z_i$ . It becomes:

$$F_{z_i}(z_i) = \int_0^{z_i} dt p_{z_i}(t) = \frac{\Gamma(2\mu)}{\Gamma^2(\mu)} B_{\frac{s_i z_i^\alpha}{\Omega_i + s_i z_i^\alpha}}(\mu, \mu), \quad (8)$$

where  $B_z(a, b)$  is the incomplete Beta function defined in [22; 8.39].

At the SC receiver output, the SIR for  $i=1,2, \dots, L$  will be:

$$z = \max(z_1, z_2, \dots, z_i, \dots, z_L). \quad (9)$$

Mathematically,  $P_{out}$  is equal to CDF. So, if use equation (8), where  $i=1,2, \dots, L$ , we can calculate  $P_{out}$ . Consequently,  $P_{out}$  of  $z$  at the SC output is obtained in the form [8]:

$$P_{out} = F_z(z) = (F_{z_i}(z_i))^L = \left( \frac{\Gamma(2\mu)}{\Gamma^2(\mu)} B_{\frac{s_i z_i^\alpha}{\Omega_i + s_i z_i^\alpha}}(\mu, \mu) \right)^L. \quad (10)$$

PDF of the SIR at the output of the  $L$ -branch SC receiver can be calculated by using expressions (5) and (9) [7]:

$$p_z(z) = L p_{z_i}(z_i) (F_{z_i}(z_i))^{L-1} = L \frac{\alpha z_i^{\alpha\mu-1}}{\Gamma^2(\mu)} \cdot \left( \frac{(\Omega_i s_i)^\mu \Gamma(2\mu)}{(\Omega_i + s_i z_i^\alpha)^{2\mu}} \left( \frac{\Gamma(2\mu)}{\Gamma^2(\mu)} B_{\frac{s_i z_i^\alpha}{\Omega_i + s_i z_i^\alpha}}(\mu, \mu) \right)^{L-1} \right). \quad (11)$$

Graphic view of system performance at output of multi-branch SC receiver is given in Figs 2 and 3. Some fading parameters change, while some are taken to be equal for all curves. The graphics are plotted using Software Mathematica and Origin.

### III. REVIEW AND ANALYSIS OF NUMERICAL RESULTS

In order to analyze the influence of fading and interference severity on the concerned performance, numerically obtained results are presented graphically. Figs. 2 and 3 show PDF and CDF of signal-to-interference ratio at the output of the SC receiver versus crossing threshold. Obtained results presented in the figures are calculated in the expressions (10) and (11).

In both figure is taken that  $\Omega_1=\Omega_2=\dots\Omega_n=1$ , and  $s_1=s_2=\dots s_n=1$ . The cases of two or more SC combiner inputs are observed.

The PDF of output SIR is shown in Fig. 2. One can be seen that the performance of the system is improved with increasing number of branches  $L$  because PDF decreases. Also, the influence of nonlinearity parameter  $\alpha$  and parameter  $\mu$  on PDF can be visible from this figure.

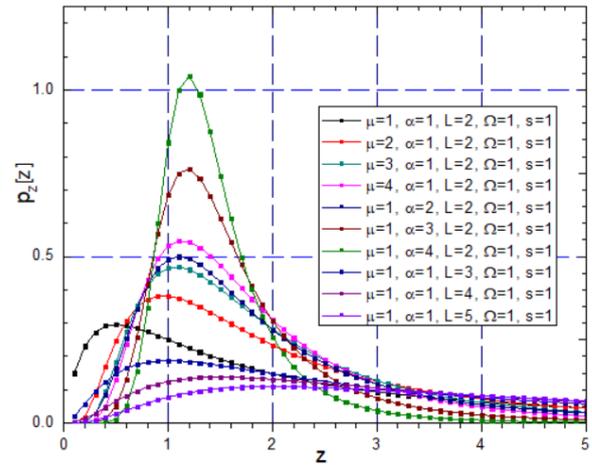


Figure 2. PDF of MIMO system with multi-branch SC combiner versus SIR, for different values of parameters  $\mu$ ,  $\alpha$  and  $L$

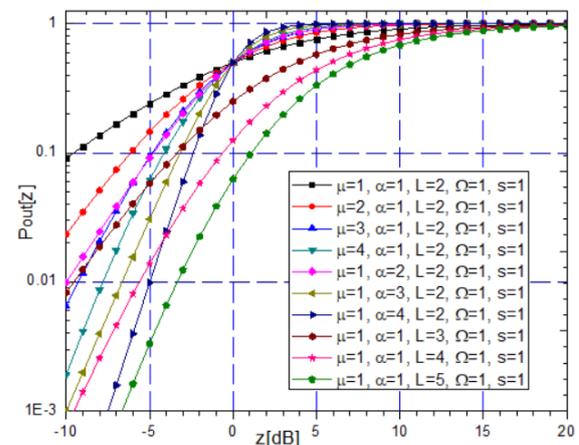


Figure 3.  $P_{out}$  of observed system with multi-branch SC receiver, for different values of parameters  $\mu$ ,  $\alpha$  and  $L$

It can be noticed from Fig. 3 that as the parameters  $\mu$ ,  $\alpha$  and  $L$  increase,  $P_{out}$  decreases and the system has better performance and a lower probability of failure.

#### IV. PLANNING AND SIMULATION ENVIRONMENT

This section introduces a software environment for providing modelling and planning of MIMO multi-branch SC diversity system that can be accessed using web browser, built upon [23, 24]. In Fig. 4, the working principle of this tool is depicted [5; Fig. 4].

In the first step, it is necessary to draw a diagram representing network model. The model takes into account the following factors when it comes network planning [23, 24]: base stations, service consumers (surveillance systems, drones, autonomous vehicles, robots), environmental configuration (obstacles) and characteristics of communication channel.

Furthermore, the value of outage probability is calculated NVIDIA CUDA [25], leveraging the Graphics Processing Unit (GPU) as described in [23]. This way, the time required for loop-based calculations is reduced, as this approach provides parallelization in this case. Due to fact that loops are constituent part of  $P_{out}$  computation expression, the speed up of up to 59 times was achieved, if compared against the equivalent implementation using Mathematica, running on CPU.

Moreover, we make use of GPU hardware to predict the demand of a service for given location using deep learning regression model implemented in PyTorch [26]. We consider the following as input variables: location, day within the week and average daily temperature. The implemented deep neural network used for this purpose has three hidden layers with 50 nodes performing ReLU activation. For training, the mean squared error (MSE) cost function was used and learning rate 0.01. The model achieved relative error around 11%. Fig. 5 shows an excerpt of Python code that illustrates the architecture of the described neural network.

In the last step, when we have both the  $P_{out}$  and service demand prediction values, linear optimization procedure is executed in order to find the optimal base station assignment (from set BS) for the observed locations within smart city (from set L). AMPL [27] was used for optimization model definition. The aim of objective function is to maximize the performance by minimizing the value of  $P_{out}$ , while keeping the costs as small as possible:

$$\text{minimize } \sum_{bs \in BS, l \in L} assign[bs, l] P_{out}[bs, l] cost[bs, l]. \quad (12)$$

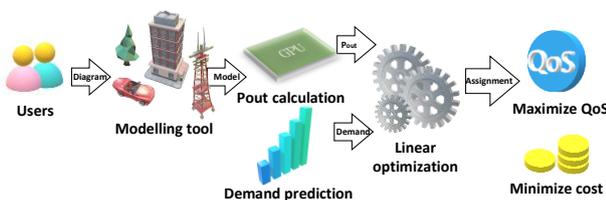


Figure 4. Network planning software environment workflow

```
class DemandModel(torch.nn.Module):
    def __init__(self, input_length):
        super(MyNeuralNetwork, self).__init__()
        self.input_size = input_length

        self.l1 = torch.nn.Linear(input_length, 50)
        self.l2 = torch.nn.Linear(50, 50)
        self.l2 = torch.nn.Linear(50, 50)
        self.l3 = torch.nn.Linear(50, 1)

    def forward(self, x):
        predicted = F.relu(self.l1(x))
        predicted = F.relu(self.l2(predicted))
        predicted = F.relu(self.l3(predicted))
        out = self.l4(predicted)
        return out

demand_predictor=DemandModel(5)
```

Figure 5. Regression neural network Python code excerpt

TABLE I. RESULTS FOR VARYING NETWORK MODEL SIZE

Base station $s$ [num]	Location $s$ [num]	$P_{out}$ [s]	Pred. [s]	Opt. [min]	Cost red. [%]
3	1	0.37	0.08	0.16	61
6	3	0.88	0.19	0.29	58
9	4	2.04	0.25	1.05	72

On the other side there is a constraint that for each location  $l$ , the value of channel capacity [28] for base station  $s$  at given location  $cap[bs, l]$  has to sufficient regarding the previously predicted demand value  $demand[l]$ :

$$\sum_{bs \in BS} assign[bs, l] cap[bs, l] \geq demand[l], l \in L. \quad (13)$$

Decision variable  $assign[bs, l]$  will take value 1 if base station  $bs$  will be placed on location  $l$ , but it is 0 in the other case.

In context of approach evaluation, the performance improvement and cost reduction percentage, together with times of execution for varying sizes of models are presented in Table I. According the results, it can be concluded that more time is required in case of larger models, while the achieved performance improvement and cost reduction depend on the specific instance of model that is considered.

#### V. CONCLUSION

In this paper, we analyzed a MIMO system with a multi-branch SC receiver in the presence of generalized  $\alpha$ - $\mu$  fading and CCI with the same distribution. For such scenario we derived an expression for outage probability and presented it graphically. The impact of fading and CCI parameters is observed and some conclusion are made. Namely, a system performance is more favorable in case the  $P_{out}$  is smaller, what is achieved with lower values of  $\mu$  and bigger  $\alpha$ , for lower SIR. Likewise, enlarging the number of input branches  $L$ , what gives possibilities to combiner to choose the branch with the best SIR, leads to better system performance.

Employing obtained expression for  $P_{out}$  at the output of the multi-branch SC combiner, we have proposed the

software environment for simulation of  $P_{out}$ . The approach with GPU for calculation the  $P_{out}$  considerably decreases the response time of mobile network simulation tool. Also, implementation of linear optimization shows promising results in network planning as it reduces the costs of system, staying QoS at the same value.

#### ACKNOWLEDGMENT

This paper is under projects TR33035, III44006 and III47003 of the Ministry of Education, Science and Technological Development, Republic of Serbia.

#### REFERENCES

- [1] R. S. Kshetrimayum, *Fundamentals of MIMO Wireless Communications*, Cambridge University Press, 2017.
- [2] B. Kumbhani and R. S. Kshetrimayum, *MIMO wireless communications over generalized fading channels*, Boca Raton, FL: CRC Press, 2017.
- [3] M. K. Simon and M. S. Alouini, *Digital Communication Over Fading Channels*. 2nd Ed., New Jersey: Wiley-Interscience; 2005.
- [4] S. Z. Milosavljevic, D. N. Milic, and P. C. Spalevic, "Performance of SC diversity receiver in fading environments with  $\eta$ - $\mu$  distribution", 26th Telecommunications Forum (TELFOR), 2018. doi:10.1109/telfor.2018.8612145
- [5] S. Suljovic, D. Krstic, and N. Petrovic, "Derivation and simulation of outage probability for 5G wireless system with L-branch SC receiver influenced by Rician fading and Nakagami-m co-channel interference", The 63rd International Symposium on Electronics in Marine ELMAR-2021, September 13-15, Zadar, Croatia
- [6] D. B. Djolic, D. M. Stefanovic, and C. M. Stefanovic, "Level crossing rate of macro-diversity system with two micro-diversity SC receivers over correlated Gamma shadowed  $\alpha$ - $\mu$  multipath fading channels", IETE Journal of Research, 2015. DOI: 10.1080/03772063.2015.1075913
- [7] M. Savic, M. Smilic, and B. Jaksic, "Analysis of Shannon capacity for SC and MRC diversity system in  $\alpha$ - $k$ - $\mu$  fading channel", University Thought, Publication in Natural Sciences, Vol.8, No.2, pp. 61-66, 2018.
- [8] M. D. Yacoub, "The  $\kappa$ - $\mu$  distribution: a general fading distribution", IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No.01CH37211), Atlantic City, NJ, USA, pp. 1427-1431 vol.3, 2001. doi:10.1109/vtc.2001.956432
- [9] N. Bhargav, C. R. N da Silva, Y. J., Chun, S. L. Cotton, and M. Yacoub, "Co-channel interference and background noise in kappa-mu Fading Channels", IEEE Communications Letters, 21(5), 1215-1218, 2017. <https://doi.org/10.1109/LCOMM.2017.2664806>
- [10] M. D. Yacoub, "The  $\alpha$ - $\mu$  distribution: a physical fading model for the Stacy distribution", IEEE Transactions on Vehicular Technology, vol. 56, no. 1, pp. 27-34, January 2007.
- [11] A. M. Magableh and M. M. Matalgah, "Channel characteristics of the generalized alpha-mu multipath fading model", 7th International Wireless Communications and Mobile Computing Conference, 2011. doi:10.1109/iwcmc.2011.5982766
- [12] R. A. De Souza and M. D. Yacoub, "The multivariate  $\alpha$ - $\mu$  distribution with arbitrary correlation and fading parameters," IEEE International Conference on Communications, ICC 2008, pp. 4456-4460, 19-23 May 2008.
- [13] A. M. Magableh and M. M. Matalgah, "Moment generating function of the generalized  $\alpha$ - $\mu$  distribution with applications," IEEE Communications Letters, vol. 13, no. 6, June 2009, pp. 411-413. doi:10.1109/lcomm.2009.090339
- [14] H. Y. Darawsheh and A. Jamoos, "Performance analysis of energy detector over  $\alpha$ - $\mu$  fading channels with selection combining", Wireless Pers. Commun., 2014, ID 287333979, DOI 10.1007/s11277-013-1595-9
- [15] W. H. M. Freitas, R. C. D. V. Bomfin, R. A. A. de Souza, and M. D. Yacoub, "The complex  $\alpha$ - $\mu$  fading channel with OFDM application", International Journal of Antennas and Propagation, vol. 2017, Article ID 2143541, 7 pages, 2017. <https://doi.org/10.1155/2017/2143541>
- [16] J. T. Ferreira, A. Bekker, F. Marques, and M. Laidlaw, "An enriched  $\alpha$ - $\mu$  model as fading candidate", Mathematical Problems in Engineering, vol. 2020, Article ID 5879413, 9 pages, 2020. <https://doi.org/10.1155/2020/5879413>
- [17] V. K. Upadhyay, S. K. Soni, P. S. Chauhan, "An approximate statistical analysis of wireless channel over  $\alpha$ - $\mu$  shadowed fading channel", International Journal of Communication Systems, Vol. 34, Issue13, First published: 16 June 2021 <https://doi.org/10.1002/dac.4884>
- [18] A. Mathur, Member, Y. Ai, M. Cheffena, and G. Kaddoum, "Secrecy Performance of Correlated  $\alpha$ - $\mu$  Fading Channels", IEEE Communications Letters, Volume: 23 Issue: 8, Aug. 2019, DOI: 10.1109/LCOMM.2019.2920347
- [19] H. Bitra and P. Ponnusamy, "Analysis of GSM-SM over  $\kappa$ - $\mu$ ,  $\eta$ - $\mu$  and  $\alpha$ - $\mu$  fading channels", Wireless Personal Communications, 2019. doi:10.1007/s11277-019-06931-9
- [20] S. Suljović, D. Krstić, D. Bandjur, S. Veljković, and M. Stefanović, "Level crossing rate of macro-diversity system in the presence of fading and co-channel interference", Revue Roumaine des Sciences Techniques, Publisher: Romanian Academy, vol. 64, pp. 63-68, 2019.
- [21] D. Krstić, S. Suljović, D. Milić, S. Panić, and M. Stefanović, "Outage probability of macro-diversity reception in the presence of Gamma long-term fading, Rayleigh short-term fading and Rician co-channel interference", Annals of Telecommunications, vol. 73, Issue 5-6, pp. 329-339, June 2018. doi: 10.1007/s12243-017-0593-4.
- [22] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products Academic*. New York: 1980.
- [23] N. Petrović, S. Koničanin, D. Milić, S. Suljović, and S. Panić, "GPU-enabled framework for modelling, simulation and planning of mobile networks in smart cities", ZINC 2020, pp. 1-6. <https://doi.org/10.1109/ZINC50678.2020.9161773>
- [24] N. Petrović, S. Vasić, D. Milić, S. Suljović, and S. Koničanin: "GPU-supported simulation for ABEP and QoS analysis of a combined macro diversity system in a Gamma-shadowed  $k$ - $\mu$  fading channel", Facta Universitatis: Electronics and Energetics Vol. 34, No 1, 2021, pp. 89-104. <https://doi.org/10.2298/FUEE2101089P>
- [25] J. Sanders and E. Kandort, *CUDA by example: an introduction to general-purpose GPU programming*, Addison-Wesley, 2011.
- [26] PyTorch [online], available on: <https://pytorch.org/>. Last accessed: 21/06/2021.
- [27] R. Fourer, D. Gay, and B. Kernighan: "AMPL: A Modeling Language for Mathematical Programming", Duxbury Press, 2003.
- [28] D. Krstić, S. Suljović, N. Petrović, Z. Popović, S. Minić: "Channel capacity of macrodiversity system in Gamma shadowed  $k$ - $\mu$  fading environment", AIIT 2020 International conference on Applied Internet and Information Technologies, pp. 125-130, 2020.

# The impact of air pollution on bacteriological flora

D. Jovanovski\*, E.M. Jovanovska\*, K. Popovska\*\*, A. Naumoski\*

\* UKIM-Faculty of Informatics and Computer Engineering/Laboratory for Ecoinformatics, Skopje, N. Macedonia

\*\* UKIM-Medical Faculty/Department for Microbiology, Skopje, N. Macedonia

damjan0003@gmail.com, elenamitreska@gmail.com, popovska.katja@gmail.com, andreja.naumoski@finki.ukim.mk

**Abstract - Our atmosphere may not seem like a habitat suitable for bacteria to many people, but with some recent research, we get the opposite. They indicate that it contains a diverse range of bacteriological flora. Does, and to what extent, polluted air affect the presence of these bacteria in our atmosphere, and does it catalyze their reproduction? We will answer the answers to these questions with regular measurements of air pollution at several outdoor and indoor locations, the difference between outdoor and indoor pollution, and bacteriological measurements of indoor air at the same locations.**

## I. INTRODUCTION

With each passing year, more and more air pollution is imposed on our daily lives as an unacceptable daily routine that is very difficult to control and avoid. World Health Organization (WHO) reports [1] are becoming increasingly alarming with the number of lives being taken by air pollution. The World Health Organization claims that as much as 7.6% of global mortality is a product of air pollution or more than 7 million per year. Of which over 3 million deaths are the product of ambient air pollution and another 4.3 million deaths are the product of household pollution in more than 3,000 cities worldwide [2]. This problem raises another question, is air pollution a catalyst for bacteria and viruses in the atmosphere? Our atmosphere may not seem like a habitat suitable for bacteria to many people, but with some recent research, we get the opposite. Bacteria that are considered normal indoor air flora live with us and in our atmosphere every day of our lives, and it is usually composed of: Staphylococcus, Corynebacterium, Bacillus, and gram-negative fermenting bacteria. Studies show that the atmosphere contains a diverse range of bacteriological branches that are not recognized as normal, some of which are: Firmicutes, Actinobacteria, Proteobacteria, and Bacterioids that can form active bacteriological communities [3]. Their complexity is similar to that of bacteria found in terrestrial and marine environments. Today we have a very limited number of studies on this issue of the metabolic activity of microbes in the air. Microorganisms in the atmosphere can come into contact with two types of pollutants: in PM (particulate matter) particles, or microorganisms can be part, or embedded in them and also gaseous pollutants, such are: ozone (O<sub>3</sub>), nitrogen oxides (NO<sub>x</sub>), and volatile organic compounds (VOC) PM and tropospheric ozone are considered to be the most problematic pollutants in Europe. Especially the respiratory part of PM particles, PM<sub>2.5</sub> (diameter below 2.5µm), is associated with heart and respiratory diseases. Heavy metals are claimed to be

one of the major toxic components of PM<sub>2.5</sub> because they can be toxic at low concentrations. We aim to find the direct impact of air pollution on the existence and their behavior by regularly measuring air pollution over the annual seasons indoors and outdoors, as well as the bacteriological presence in the air by measuring it with appropriate laboratory equipment [7].

## II. MATERIALS AND METHODOLOGY OF MEASUREMENT (METHOD AND MATERIALS)

### II.1. Sampling for air quality control

We realized the external collection of aerial samples at two locations, one external and one internal. The external measurement was performed with the device (BAM-1020 MetOne Instruments Inc.) which, in addition to accurate PM measurements, also gives readings for O<sub>3</sub>, NO<sub>2</sub>, SO<sub>2</sub>, and CO. Internal aerial measurement is performed with an apparatus (Aerocet 831 MetOne Instruments, Inc.) that provides accurate readings for PM<sub>1</sub>, PM<sub>2.5</sub>, PM<sub>4</sub>, PM<sub>10</sub> & TSP particles.

### II.2. Collecting the microbiological samples from the air

We applied the internal collection of aero-microbiological samples in two microbiological laboratories through impactor type sampler. Hard or adhesive substrates (55mm x 90mm) are used with for the Impactor samplers. The time required to take a sample with such a sampler is between 3-5 minutes. The most commonly used TSA (Trypticase soy agar) agar substrates with blood, MacConkey agar. This method is a quantitative method because it involves programmed collection of various volumes of air (10, 20, 50, 100, 200, 250, 500, 750 and 1000 liters). The reading is directly from the modern analytical agar system, Chromocult coliform agar (Merck).

### II.3. Air quality measurement locations

Air quality measurement takes place at two locations: one outdoors and one indoors.

**External measuring station** – The location of the city measuring station in the center of the City of Skopje (41.992436,21.423548) [5] was chosen as the location for the external measuring station [5] in order to obtain information on air pollution in one of the busiest places in the city, also near the Medical Faculty Skopje where the internal measuring station is located.

**Internal measuring station** - The Medical Faculty Skopje - Microbiology and Parasitology was chosen as the location for the indoor measuring station, as the ideal location for monitoring the correlation of the indoor ambient air with the outdoor air on the influence of bacteria. We started the measurement in the winter period of the year when the pollution is expected to be the highest, and it will be measured one year annually, in all 4 yearly seasons on average twice a month. From the internal measurement, we get results from PM<sub>1</sub>, PM<sub>2.5</sub>, PM<sub>4</sub>, PM<sub>10</sub> & TSP particles, we get the results from the readings every minute and we take an average value of 60 minutes of measurement as a result of a certain parameter. In our study, we place the greatest emphasis on PM<sub>2.5</sub> and PM<sub>10</sub> particles.

TABLE 1: EUROPIAN AIR QUALITY STANDARDS.

	SO <sub>2</sub>	NO <sub>2</sub>	PM <sub>10</sub>	PM <sub>2.5</sub>	O <sub>3</sub>	CO
<i>Very High</i>	>500	>400	>180	>110	>240	>20
<i>High</i>	350-500	200-400	90-180	55-110	180-240	10-20
<i>Mild</i>	100-350	100-200	50-90	30-55	120-180	7.5-10
<i>Low</i>	50-100	50-100	25-50	15-30	60-120	5-7.5
<i>Very low</i>	0-50	0-50	0-25	0-15	0-60	0-5

II.4. Criteria and significance of the values of the microbiological measurement of the bacteriological air pollution

The upper limit for the hygienic correctness of the air is 150-200 colonies of saprophytes, when the number of saprophytic colonies is greater than or equal to 200 we say that the air quality does not meet the required microbiological hygiene (reference values are according to ISO 14698-1 / 2 International Standard to control contamination in cleanrooms and other controlled environment).

TABLE 2: ISO 14698-1 / 2 International standard for bio contamination control in clean rooms and other controlled environments

Saprophyte colonies / value	Criterion
<i>0-149</i>	<i>Satisfies the hygienic correctness of the air</i>
<i>150-200</i>	<i>Upper limit of tolerance for hygienic air quality</i>
<i>≥200</i>	<i>Does not satisfy the hygienic correctness of the air</i>

## II.5. Machine learning methodology

Machine learning algorithms have been used very recently in many areas, as well as in environmental studies, to gain new knowledge or improve the understanding of the existing body of knowledge regarding the inner-working elements of a given ecosystem. In this direction, in this paper we apply a particular machine learning algorithm of decision trees, notably Predictive Clustering Trees (PCT) [8], that are well known for their easy interpretation of the final model for the biological experts. The PCT algorithm operates on the principles by partitioning the data into smaller chunks of data by using a heuristic function computed on the input data. The final model depicts the relationship between the target attribute and the set of descriptive attributes. In experiments, the target attribute is the Saprophyte colonies value, while the descriptive attributes are the PM<sub>2.5</sub> and PM<sub>10</sub> measured data inside of the lab as well as the outside of the lab, combined with the information of the season that are measured (Winter or Summer). From this experimental setup, we obtained four models: two models regarding the measurements inside of the laboratory measurements (LabPM) and outside of the laboratory (OutPM) with PM<sub>2.5</sub> and another two models with PM<sub>10</sub>. The final four PCT models are selected from several models that are obtained by applying certain constraints in the algorithm. Practically, in our experimental setup we applied constraints regarding the depth of the tree model that can maximally grow and the number of leaves that each tree model can have. From these models, based on the descriptive (training) performance, and more important based on the 10-fold cross validation procedure, the predictive (test) performance of the models, four models were finally selected. These four models are presented and discussed in the next section.

III. RESULTS AND DISCUSSION

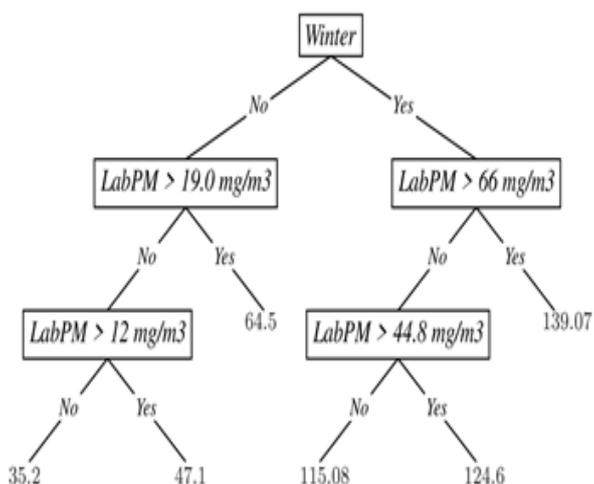


Figure 1. Measurements of total indoor air pollution by seasons associated with microbiological pollution measurements

The obtained model refers to the total measurements of PM<sub>10</sub> and PM<sub>2.5</sub> particles in indoor air compared to the degree of microbiological contamination of air expressed by the number of CFU / 1 L air (CFU = colony forming units), on nutrient blood agar substrate. The model shows comparatively the differences of the two measurements depending on the seasonal influence, i.e. winter and summer months.

The total air pollution of the "indoor" air in the winter months is at least 3.5 times higher > 66 µg/m<sup>3</sup> compared to the air pollution in the summer months > 19 mg / m<sup>3</sup>. The winter period that shows measurements above the obtained average is accompanied by higher microbiological contamination of the air 139.07 CFU / l, compared to air pollution that ranges in values around the legal norm, ie between 44.8 – 66 µg/m<sup>3</sup> and subsequently 115.08. CFU / L for values closer to the legal standard for the Republic of Macedonia and then increases to 124.6 CFU / L in parallel with the growth of taste air pollution with PM particles above average.

The total air pollution of the "indoor" air in the summer months 19mg / m<sup>3</sup> which is at least 3.5 times lower than in the winter period is followed by 2.5 times lower microbiological air pollution, ie 64.5 CFU / L. Analogous to the winter measurements and summer measurements show a decrease in microbiological pollution with a decrease in average PM air pollution (47.1 CFU / L followed by 35.2 CFU / L).

**3.1 Conclusion** The winter period is accompanied by a minimum of 3.5 time's higher indoor air pollution, which directly affects the increase in the degree of microbiological contamination of the same air.

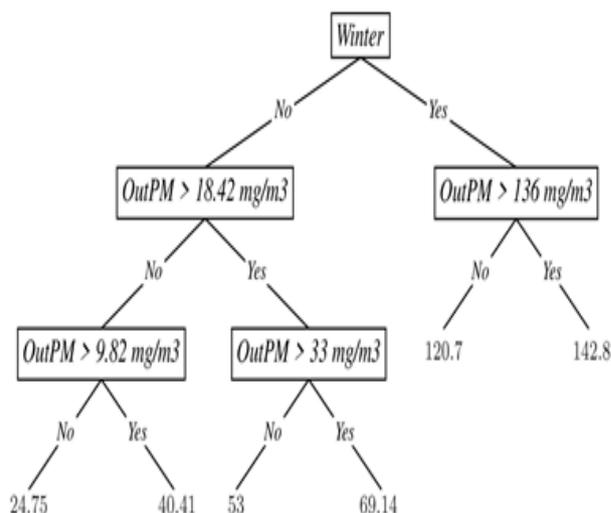


Figure 2. Measurements of total outdoor air pollution by seasons associated with microbiological indoor air pollution measurements

The algorithm refers to the total measurements of PM<sub>10</sub> and PM<sub>2.5</sub> particles in the outside air compared to the degree of microbiological contamination of the air in the laboratory air expressed by the number of CFU / L air (CFU = colony forming units), on nutrient blood agar substrate. The measurements of air pollution with total PM particles are from the measuring station of the municipality of Centar.

The algorithm shows comparatively the differences of the two measurements depending on the seasonal influence, i.e. winter and summer months.

The total air pollution of the "outdoor" air in the winter months 136 µg/m<sup>3</sup> is at least 7 times higher than the average air pollution of the outside air in the summer months. From the measurements in the winter months it can be seen that if the air pollution with total PM particles above the average 1361 µg/m<sup>3</sup> is followed by 142.8 CFU / L air in the laboratory space which is a higher value than the microbiological contamination at lower load with PM or 120.7 CFU / L.

Summer seasonal measurements, which are 7 times lower than winter, showed microbiological contamination at 53 CFU / L. and 69.14 CFU / L, if PM air pollution ranges between > 18.2 µg/m<sup>3</sup> and > 33 µg/m<sup>3</sup>, it decreases significantly to 40.41 CFU / L if total PM is around the average measured > 18.2 µg/m<sup>3</sup> and almost decreases by 2-3 times at total PM > 9.82 µg/m<sup>3</sup>.

**3.2 Conclusion** The winter period is accompanied by at least 7 times higher air pollution in the outside air or "outdoor" air, which directly affects the increase of the degree of microbiological contamination of the indoor "indoor" air, by the same at least 7 (24.75 CFU / L - 142.8 CFU / L) times between the period of lowest summer PM (> 9.82 µg/m<sup>3</sup>) air pollution and the period of highest winter air pollution (136 µg/m<sup>3</sup>)

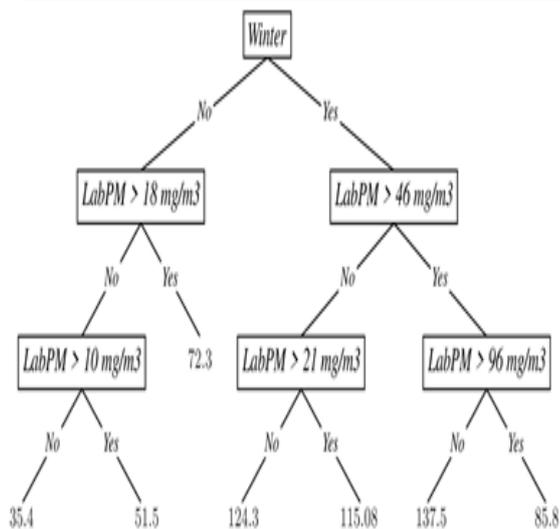


Figure 3. Laboratory measurements of pollution with PM 2.5 and microbiological contamination and relationship with seasonal variation

The algorithm refers to measurements of PM 2.5 particles in indoor air compared to the degree of microbiological contamination of air expressed by the number of CFU / 1 L air (cfu = colony forming units), on nutrient blood agar substrate.

The algorithm shows comparatively the differences of the two measurements depending on the seasonal influence, ie. winter and summer months.

Air pollution of air with PM 2.5 in the winter months is at least 2.5 times higher > 46  $\mu\text{g}/\text{m}^3$  compared to air pollution in the summer months > 18  $\mu\text{g}/\text{m}^3$ .

Winter measurements of PM 2.5 in the laboratory are on average around the legally allowed standard, ie 46  $\mu\text{g}/\text{m}^3$ . If their value increases > 96  $\mu\text{g}/\text{m}^3$ . The microbiological pollution of 85.8 CFU / L does not follow the trend of their increase, contrary to the values of PM 2.5 which is above the standard but are between 46  $\mu\text{g}/\text{m}^3$  and 96  $\mu\text{g}/\text{m}^3$  when the microbiological pollution is significantly higher, ie. 137.5 CFU / L. In some winter months where PM 2.5 particles were twice lower than the permissible standard (21  $\mu\text{g}/\text{m}^3$ ) the microbiological contamination does not decrease analogously but is around the values (115.08 cfu / L. And 124.3 CFU / L. ) which are also obtained for PM2.5 values between > 46  $\mu\text{g}/\text{m}^3$  and > 96  $\mu\text{g}/\text{m}^3$ . The summer results of PM particles of 2.5, with values > 10  $\mu\text{g}/\text{m}^3$  to > 18  $\mu\text{g}/\text{m}^3$  are sent with approximately 2 times lower values 72.3 CFU / L, compared to their numbers at > 21  $\mu\text{g}/\text{m}^3$  or 96  $\mu\text{g}/\text{m}^3$  which were measured in some winter months.

When the values of PM 2.5 fall > 10  $\mu\text{g}/\text{m}^3$  but not > 18  $\mu\text{g}/\text{m}^3$  the values of microbiological contamination are the lowest (35.4 CFU / L and 51.5 CFU / L).

**3.3 Conclusion** PM 2.5 Particle Values > 46  $\mu\text{g}/\text{m}^3$  - 96  $\mu\text{g}/\text{m}^3$  inversely affect the amount of microbiological contamination.

Only their values >  $\mu\text{g}/\text{m}^3$  but not > 18  $\mu\text{g}/\text{m}^3$  are sent to a greater reduction of microbiological contamination.

TABLE 3. Microorganisms that are used as indicators for a certain type of contamination of the living environment in a period of 1 year

Month	Bacterial Flora
December 2019	Indicative microorganisms
January	
February	
March	
April	
May	<i>Bacillus subtilis</i>
June	
July	<i>Bacillus subtilis</i>
August	<i>Streptococcus fecalis</i>
September	
October	
November	<i>mold</i>
December	

\*\* Indicators of microorganisms

*Bacillus subtilis* - an indicator of increased presence of the right

*Fecal streptococcus* - an indicator of fecal contamination

*Mold* - an indicator of increased viscosity in the space

TABLE 3 shows qualitatively isolated microorganisms that are internationally accepted as indicators that can be associated with the type of hygienic air defect and other inanimate environment.

#### IV. CONCLUSIONS

1. There is a seasonal variation in total air pollution in relation to legal standards.

- Air pollution in the winter months is higher than the legally allowed standard in both outdoor and indoor air.
- In the summer months, the air pollution values are close to the standard allowed values in both outdoor and indoor air.

2. There is a quantitative difference in air pollution between outdoor and indoor air, especially in the winter months.

In the winter months the outside air is twice as polluted with total PM particles as the indoor air (136  $\mu\text{g}/\text{m}^3$  / 66  $\mu\text{g}/\text{m}^3$ ) (see Graphicon 1 and Graphicon 2)

3. There is a seasonal variation in terms of quantitative air pollution of outside air as well as indoor air.

- The winter period for the outside air is at least 7 times higher air pollution than the summer months (136  $\mu\text{g}/\text{m}^3$  vs. 18.42  $\mu\text{g}/\text{m}^3$ ) which directly proportionally affects the increase of the level of microbiological contamination, for the same at least 7 times (24.75 CFU / L - 142.8 CFU / L) times between

the period of lowest summer PM air pollution ( $> 9.82 \mu\text{g}/\text{m}^3$ ) and the period of highest winter air pollution ( $136 \mu\text{g}/\text{m}^3$ )

b. The winter period indoors is at least 3.5 times more polluted than the summer months ( $66 \mu\text{g}/\text{m}^3$  vs.  $19 \mu\text{g}/\text{m}^3$ ), which directly affects the increase in the degree of microbiological contamination of indoor air ( $139, 07 \text{ CFU} / \text{L}$  vs.  $64.5 \text{ CFU} / \text{L}$ )

4. In the summer months there is no visible difference between the total air pollution outside the indoor air, but in the summer month's microorganisms are more often present indicators of the presence of dust and fecal contamination

5. Air pollution with PM particles proportionally affects the microbiological contamination of indoor air. The reduction of air pollution is proportionally followed by a reduction of microbiological air contamination in both seasons and in both measured air samples.

a. The 7-fold increase in air pollution in the outdoor air in winter conditions directly-proportionally affects the increase in the degree of microbiological contamination, for the same at least 7 times ( $24.75 \text{ CFU} / \text{L}$  -  $142.8 \text{ CFU} / \text{L}$ ) times between the period of the lowest summer air pollution ( $> 9.82 \mu\text{g}/\text{m}^3$ ) and the period of highest winter air pollution ( $136 \mu\text{g}/\text{m}^3$ )

b. Increasing the air pollution of indoor air in winter conditions by 3.5 times in winter conditions, directly affects the increase in the degree of microbiological contamination of indoor air ( $139.07 \text{ CFU} / \text{L}$  versus  $64.5 \text{ CFU} / \text{L}$ )

c. There is no visible association of microbiological contamination with the origin of increased air pollution, i.e. outside / indoor air

6. PM values 2.5 Particles  $> 46 \mu\text{g}/\text{m}^3$  -  $96 \mu\text{g}/\text{m}^3$  affect inversely proportional to the amount of microbiological contamination.

7. Only values of pm 2.5 of  $> 10 \mu\text{g}/\text{m}^3$  but not  $> \mu\text{g}/\text{m}^3$  3 are sent with a greater reduction of microbiological contamination.

#### REFERENCES

- [1] Air quality deteriorating in many of the world's cities. World Health Organisation(WHO). <https://www.who.int/mediacentre/news/releases/2014/air-quality/en/>
- [2] 7 million premature deaths annually linked to air pollution (WHO). <https://www.who.int/mediacentre/news/releases/2014/air-pollution/en/>
- [3] Isabella Gandolfi I, Valentina Bertolini, Roberto Ambrosini, Giuseppina Bestetti, Andrea Franzetti  
Unravelling the bacterial diversity in the atmosphere. Applied microbiology and biotechnology, 97 (11), pp.4727-4736, 2013
- [4] General Information about Staphylococcus aureus, <https://www.cdc.gov/hai/organisms/staph.html>
- [5] Agglomeration Skopje, [http://air.moepp.gov.mk/?page\\_id=193](http://air.moepp.gov.mk/?page_id=193)
- [6] Medical Faculty Skopje - Microbiology and Parasitology, <https://izm.edu.mk/>
- [7] US EPA, stay of federal water quality criteria for metals; water quality standards; establishment of numeric criteria for priority toxic pollutants states' compliance revision of metals criteria; final rules, federal, register 60(86), pp. 22228-22237, US Environmental Protection Agency, Washington D.C., 1955
- [8] M. Garofalakis, D. Hyun, R. Rastogi, K. Shim, "Building decision trees with constraints," Data Mining and Knowledge Discovery, vol. 7. no. 2, 2003, pp. 187-212

# Application of the Blockchain Technology in Medicine and Healthcare

Panche Tashevski, Blagoj Ristevski, Snezana Savoska, Ilija Jolevski, Mimoza Mijoska

Faculty of Information and Communication Technologies – Bitola

University “St. Kliment Ohridski” University – Bitola, North Macedonia

pance\_05@yahoo.com, { blagoj.ristevski, snezana.savoska, ilija.jolevski, mijoska.mimoza }@uklo.edu.mk

**Abstract** – Applications in the field of medicine and healthcare would include electronic health records, health insurance, biomedical research, drug supply and procurement, and medical education. One of the requirements for health and medicine is interoperability, which represents the ability to exchange data or information accurately, efficiently and consistently. There are currently problems with the amount and exchange of data in healthcare. The potential for blockchain in healthcare is to solve the challenges connected with the security, privacy, sharing and storage of data. A proposed model in this paper describes a new procedure of design and implementation on a decentralized platform for managing data with blockchain for providing more secure, transparent and significant medical assistance for patients and providers of healthcare globally.

**Keywords:** blockchain, data decentralization, electronic health data, e-health.

## I. INTRODUCTION

The first blockchain was created by the unknown persons behind the online cash currency bitcoin, under the pseudonym of Satoshi Nakamoto in 2008. He described: “A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution”. The word blockchain is a combination of two words, the first is “block” and the second is “chain”, and this is a combination of more blocks in a chain equal to the blockchain. The first block was developed at the beginning of 2009, and that is the beginning of the new revolution in informational technologies. Decentralization as the main feature on blockchain was introduced in Web 3.0, where the “Dapp” Application was decentralized and they will be everywhere [2].

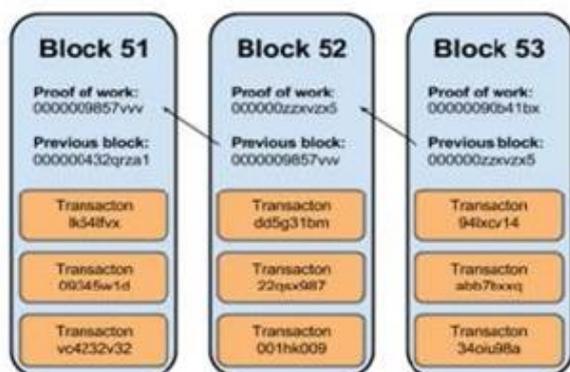


Figure 1. Blockchain blocks [3].

Blockchain is registered for all transactions because every system keeps a record of the transaction and is stored as a database. Important for blockchain systems is how the transaction information is sent and stored. At bitcoin transactions the data have stored in blocks, this is shown in Fig. 1. Cryptography is an integral part of the inner workings of blockchain technology, more precisely the hash function in such a way where is impossible to change content on one block without a content change for all of the blocks who follow him. This is an extremely important feature of blockchain's, as it ensures the immutability of the data entered into the block.

Bitcoin is a decentralized system which means that there are no central servers. All users are connected in a peer-to-peer network and every user presents one node on this network is shown in Fig.2. Because a peer-to-peer is that every user makes a direct connection with several other users, it means that the user is indirectly connected to all other users. Information was travelling through a peer-to-peer network to each of its participants with sent a message only to users who are connected directly. It means each of the users when sent a message to users who are directly connected, the message passes to all of the participants in the network, but only those users who are directly connected to the sender have an overview.



Figure 2. A peer-to-peer network

The Byzantine generals' problem is one of the most well-known and classic problems faced by decentralized networks. Solving this problem was one of the key developments in the creation of Bitcoin and, by extension, all other cryptocurrencies. Some of the nodes in the network have a complete blockchain. Because the nodes constantly communicate with each other, they always check that their copy of the block matched to block on the other nodes. If the block doesn't match with others, then the node updates automatically its version for agreement with the rest of the nodes in the network. This is a kind of database that is connected to thousands of

computers that are constantly updated in real-time. Every bitcoin user has a private key, public key and bitcoin address.

The public key has been created from the private key, but on the way that is impossible to make a reverse process that is to make a private key from the public key. Then, a public key is created, and a bitcoin address that's not in danger of abuse does therefore not need to hide. From the other side, very important is a private key to be kept well, because it is needed upon signing transactions with bitcoin address with key what is connected. For example, if the bitcoin address is signing as banking number of user account and private key has signed pin of the user who is using to confirm payment from that account.

There is one very important difference around this functionality: If the user lost his pin who confirm a transaction from his banking account, then the bank will create a new pin, but if the user loses his private key with it loses his bitcoin address forever. These bitcoins exist, but no one can access them without a private key.

When the user will want to send his bitcoins to another user, he must enter the recipient bitcoin address, the amount he wants to send and sign the transaction with his private key. Then, the information for that transaction is sent to participants with whom a user connects directly with a peer-to-peer network, so that will be forwarded until it reaches the recipient. Bitcoin transactions were forwarded only by a mediator. Every node in a bitcoin network is a mediator, which checks the transaction and forwarded it when doesn't reach all the nodes. Every time a transaction occurs it has to be approved by the nodes, each of them checks its validity. Once every node has checked a transaction there is a sort of electronic vote, as some nodes may think the transaction is valid and others think it is a fraud. The nodes referred to above are computers. Each node has a copy of the digital ledger or blockchain. Each node checks the validity of each transaction. If a majority of nodes say that a transaction is valid then it is written into a block.

The process of mining or creating blocks is:

The users take verified transactions from the pools to create a new block. The number of transactions that will be potentially added in the block, check the way of the payment and other criteria are decided by the user. The users created Merkle Tree from the transaction in the first step, until they take root in Merkle. Merkle tree, also known as a hash tree, is a data structure used for data verification and synchronization. It is a tree data structure where each non-leaf node is a hash of its child nodes. All the leaf nodes are at the same depth and are as far left as possible. It maintains data integrity and uses hash functions for this purpose [12]. The structure of the Merkle tree and Merkle root is shown in Fig. 3 [5].

If a parameter in the transaction changes, his hash will change, that way root hash will not correspond to the

written root in the block. Miners hash the block header with the corresponding hash function.

The rest of the paper is organized as follows. blockchain technology in healthcare and medicine is explained in Section II, while electronic healthcare records (EHR) is explained in the third Section. Internet on medical things (IoMT) is described in the subsequent Section. The process for issuing and filling out medical prescriptions is explained in Section V, while the sharing laboratory test/ data of result is explained in Section VI. Smart contracts based on Ethereum for clinical trials are described in the seventh Section. Data flow for compensation in healthcare is explained in Section VIII. The proposed model of the blockchain system in medicine and healthcare is highlighted in the subsequent Section. The last Section gives concluding remarks.

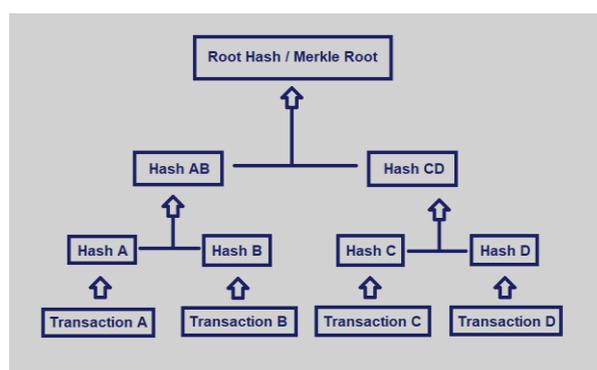


Figure 3. Merkle Tree [5].

## II. BLOCKCHAIN TECHNOLOGY IN HEALTHCARE AND MEDICINE

The potential for blockchain in healthcare is to solve the challenges connected with the security, privacy, sharing and storage of data.

One of the requirements for health and medicine in the industry is interoperability, which represents the ability to exchange data or information accurately, efficiently and consistently. Interoperability in healthcare has a purpose to facilitate the exchange of information connected with healthcare, between providers on healthcare services and patients.

In healthcare, the blockchain has a positive impact on the result of the healthcare on the companies and interesting parties to optimize the business process to improve the patient result, management with data of the patient, lower costs and better use of records connected with healthcare protection. The ability to share information for medical healthcare without risk of privacy and security of records of endangerment on users and patients is one of the primary steps for improvement in the quality of healthcare service for patients and users.

## III. ELECTRONIC HEALTHCARE RECORD (EHR)

Traditional medical files known as PHR (Personal Healthcare Records) are written on paper. Information

technologies allow the introduction of electronic healthcare records known as EHR [6].

EHR system is developed with help of blockchain technology, where data will be able shared securely or kept in the frame of the institutions. EHR content complete data of the patients like the history of examinations, tests, results, diseases, drugs and prevention, etc. Electronic access to health records allows the doctor to get a significant improvement of quality treatment on patients [7].

EHR allows better measures to treat the disease, improvement of healthcare delivery and increased level of preventive care to patients.

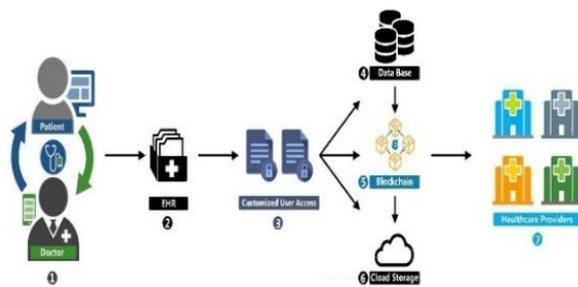


Figure 4. EHR in Blockchain medical system [8].

The primary data was generating of interaction between the patient and his doctor and specialist. This data consists of medical history, current problems and other physiologic information. EHR was created for every patient by using primary data collected in the first step, as shown in Fig.4. Other medical information that is generated by medical care, medical picture, drugs and history are input in the EHR.

Individual patients can access their EHR, while a customized control is given only on the institution for which the patient was referred [12]. Institutions that want to access information must ask for the permission of the patient, and a patient is a person who decided to whom give access control to his data. These three steps were a part of the core of all processes, including database, blockchain and storage data in the cloud. Healthcare providers, such as ad hoc clinics, hospitals are the end-users who gain access to secure data delivery and are authorized by the owner or the patient.

#### IV. INTERNET OF MEDICAL THING (IoMT)

In the IoMT area, the patient is a source of all data. Medical IoT devices are normally connected or in the distance where they follow patients through generating a big data volume. The created data are stored on a block or they stored in the cloud. Artificial Intelligence will help the blockchain to make intelligent virtual agents, which will be able to automatically create new records [10].

In the case of sensitive medical data, where the priority was security, a decentralized system will help with blockchain to reach better security. Lenders for health care are the last users who want access to health and secure delivery. The blockchain application for the Internet of Medical Things is shown in Fig.5.

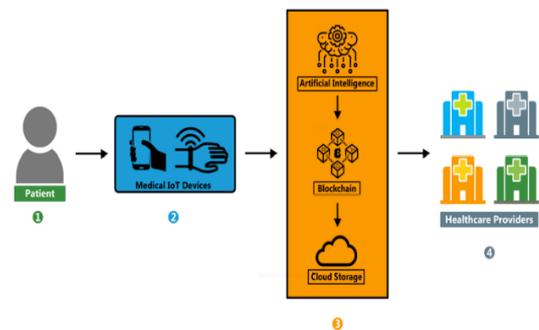


Figure 5. Blockchain for the Internet of Medical Things [8].

#### V. PROCESS FOR ISSUING AND FILLING OUT MEDICAL PRESCRIPTIONS

The process of medical prescription management was going to eliminate a long time process for waiting, remove an element for fraud of system and reduce the mistake made on the side of misinterpretations by doctors.

The doctor prescribes for the patient and he enters a patient's data in health evidence with a smart contract.

The pharmacy through a smart contract on the Ethereum blockchain access the prescription with permission given by the primary doctor and patient.

After access to the prescription, the pharmacy issues medicine and the way of usage sending by a smart contract on Ethereum blockchain [9].

#### VI. SHARING LABORATORY TEST/ DATA OF RESULT

The patient visits a laboratory for blood testing, and after the data processing and examination, the laboratory shares results and put them into the patient health record. Patients through the Ethereum blockchain will get a notification where the processed result of the test is available.

The patient permits to share a piece of information on a blockchain. When a patient will allow to share a medical record of the blockchain medical system, the result will be available in the medical system.

Laboratory through the Ethereum blockchain healthcare system will access patient information for his working place where they commit payment for this service.

#### VII. SMART CONTRACT BASED ON THE ETHEREUM FOR CLINIC TRIALS

To enable users to execute smart contracts connected with clinics trials on the Ethereum network was resulting in more secure drugs.

In this process, more metadata, including registration on protocols, previous set details, diaries for screening and writing through a smart contract are used.

Pharmaceutical companies' metadata will be stored in the Ethereum blockchain for the identification of potential patients for inclusion in the clinical trials.

The organization was sending a message to selected patients, including applications for reading and access to their medical records, including any relevant results of laboratory studies.

If the patient is allowed access to the account, the pharmaceutical company will be processed through a smart contract. The patients, meanwhile, besides compensations for participation in trials, will be accessible to the new treatment.

#### VIII. DATA FLOW FOR COMPENSATION IN HEALTHCARE

The doctors will quickly resume patient care instead to maintain the treatment of their patients while they are waiting for an answer from the payer.

The companies from healthcare insurance are sharing their policies through the smart contracts on Ethereum blockchain technology.

The supplier submits to blockchain application for previously permission for meeting on specialist, treatment or recipe.

Smart contract for a medical decision-maker of the payer determines automatic payment approves with using patient medical information. The patient and all of the laboratories, pharmacy and specialists on which patient delegated access will be checked insurance with having authorization in real-time.

#### IX. PROPOSED MODEL OF BLOCKCHAIN SYSTEM IN MEDICINE AND HEALTHCARE

The users such as doctors, nurses through applications connected with blockchain will access data when they have a patient code. The user to get a code of data of the patient has to check the PHR application and access with

patient code which is required to input to required records.

The medical user has received code and inputs him in the blockchain system where he was performing a check in the medical database in which they are enrolled patient records. The received data with blockchain technology are decrypted with a code entered in medical user who gets them EHR records of the patient, but a user can check patient's insurance. When a medical user will finish the task with a patient, he/she writes a result through a blockchain in the medical database or on the cloud storage.

During the payment, the medical users and hospitals where they are employed through the blockchain access to data in the healthcare fund, and the funds are distributed appropriately. If the patient has different insurance, it will calculate whether the cost will be covered by the healthcare fund, and/or by the company where the patient worked.

Access to medical data may have and Pharmacy when patients are depending on medicines. Also, patients who have IoT devices connected through an application where entering data their code through their devices and will gain access to the data.

The Pharmacy companies through blockchain technologies will receive additional information's for finding new pharmacy solutions, but also to improve the old ones through organized reports.

Governmental and other institutions can access medicine data through blockchain technology through execution on different statistic procedures that are performed during check-in the healthcare system in one same country. The proposed blockchain system in Medicine and Healthcare is shown in Fig.6.

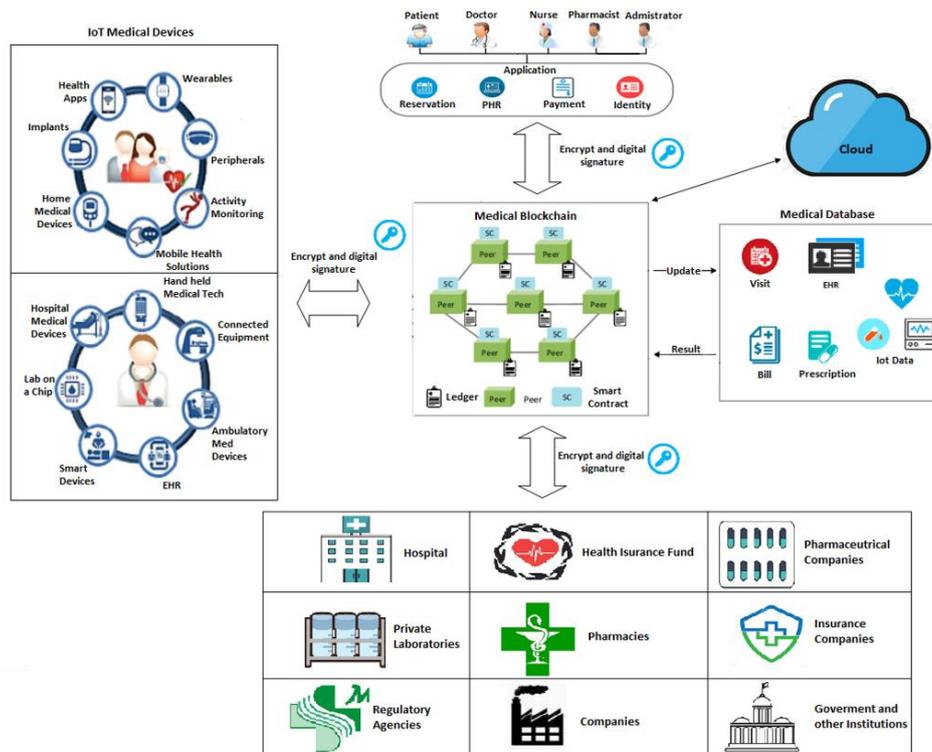


Figure 6. Proposed Blockchain System in Medicine and Healthcare.

## X. CONCLUSION

Blockchain has the potential to solve many problems in the healthcare sector today. Through this technology, new solutions in healthcare and medicine can be provided. This new model uses new dynamics through different stakeholders in medicine, such as patients and providers of services.

A proposed model in this paper describes a new procedure of design and implementation on a decentralized platform for managing data with blockchain for providing more secure, transparent and significant medical assistance for patients and providers of healthcare globally and in the hospital. Using a blockchain in this model improves the process of healthcare for the patient.

Blockchain can help in many ways through reduced transaction costs by using smart contract which is embedded for the general purpose of simplifying procedures, reducing administrative burdens and removing intermediaries.

## REFERENCES

- [1] Satoshi Nakamoto "Bitcoin: A Peer-to-Peer Electronic Cash System".2008

- [2] Faten Adel Alabdulwahhab "Web 3.0: The Decentralized Web Blockchain networks and Protocol Innovation" 2018
- [3] Paul Graham "Bitcoin by analogy" April 2014
- [4] Lutpin " Dos and Don'ts of Peer-to-Peer Trading" 2019
- [5] JAKE FRANKENFIELD "Ralph Merkle Tree" 2020
- [6] Wang, S.J.; Middleton, B.; Prosser, L.A.; Bardon, C.G.; Spurr, C.D.; Carchidi, P.J.; Kittler, A.F.; Goldszer, R.C.; Fairchild, D.G.; Sussman, A.J. A cost-benefit analysis of electronic medical records in primary care. *Am. J. Med.* 2003, 114, 397–403.
- [7] Terry, A.L.; Thorpe, C.F.; Giles, G.; Brown, J.B.; Harris, S.B.; Reid, G.J.; Thind, A.; Stewart, M. Implementing electronic health records: Key factors in primary care. *Can. Fam. Phys.* 2008, 54, 730–736.
- [8] Seyednima Khezr, Md Moniruzzaman, Abdulsalam Yassine and Rachid Benlamr "Blockchain Technology in Healthcare: A Comprehensive Review and Directions for Future Research". 2019
- [9] Zheng, Zibin, Shaoan Xie, Hong-Ning Dai, Xiangping Chen, and Huaimin Wang. "Blockchain challenges and opportunities: A survey." *International Journal of Web and Grid Services* 14, no. 4 (2018): 352-375.
- [10] Yaeger, K.; Martini, M.; Rasouli, J.; Costa, A. Emerging "Blockchain Technology Solutions for Modern Healthcare Infrastructure". *J. Sci. Innov. Med.* 2019.
- [11] Mijoska, M., and Risteovski B. "Blockchain Technology and its Application in the Finance and Economics." (2020): *International Conference on Applied Internet and Information Technologies AIIT 2020, Zrenjanin, Serbia*, pp. 197-202.
- [12] Mijoska, M. and Risteovski, B., 2021. Possibilities for applying blockchain technology—a survey. *Informatica*, 45(3).

# A Survey of Energy-efficient Solutions for 5G Networks

M. Ilić\*, V. Mikić\*, A. Zakić\* and D. Zlatković\*\*

\* Alfa University/Faculty of Information Technology, Belgrade, Serbia

\*\* Alfa University/Faculty of Mathematics and Computer science, Belgrade, Serbia

milos.ilic@alfa.edu.rs, vladimir.mikic@alfa.edu.rs, aleksandar.zakic@alfa.edu.rs, dragan.zlatkovic@alfa.edu.rs

**Abstract - 5G is a 5th generation mobile network which represents a new global wireless standard. 5G wireless technology brings greater bandwidth, increased speed, lower latency, higher performance, more reliability, etc. Because of its numerous business benefits, wireless communication is reaching new heights. With all of this comes higher electricity consumption, which will inevitably result in global environmental issues, and that is why sustainable 5G networks are necessary. Green 5G networks aim to make communication as green as possible. In this paper the review of network generations and comparison of their power consumption is given. Also, the presentation of green 5G network solutions for energy savings is presented. The results of these solutions in terms of reducing energy consumption were compared and discussed.**

## I. INTRODUCTION

When compared to previous communication system generations, the deployment of 5G networks is projected to deliver substantial improvements in capacity, the number of connected devices, energy efficiency, and speeds [1]. These characteristics will be achieved, among other things, by a mixture of larger bandwidths, improved antenna technology, and flexible radio access options. As a result of the increase in data traffic, energy consumption will increase as well, leading to an increase in worldwide CO<sub>2</sub> emissions. Energy consumption will also represent a vital economic cost element for future wireless networks, posing significant difficulties for the development of green networks aimed at reducing carbon footprints [2]. Due to these reasons, network operators and device suppliers are actively promoting an energy-efficient network evolution. Taking all of these factors into account, energy efficiency and sustainability are identified as important objectives for the deployment of 5G. Numerous innovative technologies have been proposed to minimize traditional energy usage in order to achieve sustainable 5G. To minimize the reliance on conventional energy, green energy alternatives are being researched.

In addition to summarising and defining all network generations from 1G to 5G, we will discuss and present an overview of recent academic and industrial research on 5G energy-saving solutions and their energy efficiency.

## II. GENERATIONS OF WIRELESS CELLULAR TECHNOLOGY

Networks for Wireless Communication have undergone a radical transformation in the past several decades. Wireless communication generation (G) refers to systemic changes within the speed, technology and frequency of the networking system. Standards, capabilities, new technologies, and new features distinguish each generation from the previous one. Throughout the history of network development, four main generations emerged, with the addition of the fifth, currently under deployment.

### A. Network Generations from 1G to 4G

The first generation (1G) of wireless mobile communication systems was created in the early 1980s using an analog architecture. Founded on Advanced Mobile Phone System (AMPS) technology, it provided voice services. Considering 1G systems were founded on analog technology, one of the typical issues that emerged was sensitivity to interruption, which affected call quality. Furthermore, there was an inherent vulnerability since analog signals do not permit the use of sophisticated encryption technologies. The considered speed of a 1G network is up to 2.4 kbps. The main limitations of this generation of communication technology include restricted capacity, device size, poor sound quality, battery life, and overall efficiency [3].

The second generation (2G) technology was introduced in the late 1980s, and it transmitted voice data via digital signals at a rate of 64kbps. It also offered services such as SMS, photo messaging, and multimedia message services (MMS). The GSM (Global System for Mobile Communication) represents the most widely used 2G mobile standard, and was the first technology that supported international roaming [4]. This allowed mobile users to utilize their cellular connections in multiple nations across the world with higher quality and stability. However, 2G networks have limitations such as interference difficulties caused by frequency reuse, geographical circumstances that can cause occasional call drops, and other issues. The main major difference between the first and the second generation of wireless networks was that radio signals in 1G were analog, whereas in 2G, the signals were digital.

In the early 2000s, the third generation (3G) of wireless mobile communication technology was introduced. The objective of 3G networks was to provide higher data speeds ranging from 144 kbps to 384 kbps in large coverage regions. 3G offered advanced services in comparison to previous generations. In addition to standard voice communication, it provided internet services like: TV access, e-mail, web browsing, navigation maps and video calls [5]. The main disadvantage of the technology is its inefficiency in terms of energy use, because 3G consumes substantially more power than most 2G models and is less cost-effective to run and maintain in comparison to previous generations of networks [6]. The transition from Universal Mobile Telecommunication System (UMTS) to high-speed packet access (HSPA) gave considerably improved end-to-end network performance, eventually leading to the creation of the fourth generation.

The fourth generation (4G) mobile networking system, which was entirely IP-based, was launched in the late 2000s. The primary objective of 4G technology is to deliver high-speed, secure, and low-priced voice and data services over IP. The purpose of the shift to all IP is to provide a single platform for all the other technologies that have been created thus far [7]. The speeds for this generation vary from 100 Mbps up to 1 Gbps. Terminal mobility is an important aspect of 4G since it allows wireless services to be provided at any time and from any location. From a structural viewpoint, the 4G long-term evolution (LTE) framework is built to provide support for packet-switched communication with low latency and continuous mobility [8].

#### B. 5G network

Generation after generation has seen periods of solely analog frameworks with no data capabilities (1G), digital systems optimized for excellent voice telephony (2G), broadband and multimedia systems (3G), the IP-based network revolution, and the emerging era of unified IP, enormous and effortless end-to-end connectivity, and mobility (5G) [9]. Moreover, a comprehensive examination of the timeline from the first generation to the present one reveals an obvious tendency for the arrival of new wireless technology to appear roughly every decade. Modern, more challenging use cases like high-resolution video content, traffic accident prevention, remote access, and real-time control impose new demands on the network in terms of end-to-end latency, dependability, and resilience. Despite the advances in the architecture and development of 4G cellular networks, these modern communication requirements demand the deployment of a fifth-generation mobile network. 5G aspires to give infinite access to data and the opportunity for anybody, everywhere, at any time, to exchange data for the benefit of the whole planet. It also features all of the sophisticated characteristics that make this generation of mobile technology the most powerful yet and in high demand. The 5G standardization processes have already resulted in the commercial availability of this technology. The increased output rate loss in 5G technology is addressed by using the Open Transport Protocol (OTP). 5G outperforms the current 4G in terms of speed and coverage, as it uses a 5 GHz signal and is designed to

provide speeds of up to 1 Gb/s for a handful of connections or tens of Mb/s for thousands of connections. Radio Access Networks (RANs) using fifth-generation technology would no longer be constrained by base station proximity or a complicated architecture, as 5G paves the path for virtual RAN by introducing new protocols that enable the creation of extra data access points. The implementation of other important technology concepts like massive MIMO (Multiple-Input-Multiple-Output) is a cornerstone for the deployment of 5G. MIMO is based on the fact that a radio signal between a transmitter and receiver is processed by reflections from barriers, resulting in numerous signal pathways [10]. Another 5G trend that offers low latency and represents a new paradigm in wireless networks is Device-to-Device communication (D2D). It represents direct communication between users (devices) without incorporating base stations. Based on the evolution between generations of wireless networks, the fifth-generation represents next-gen technology which is able to satisfy the most demanding networking tasks and provide a stable platform for incorporating new, demanding use cases.

### III. A GREEN 5G NETWORK

The high energy usage of mobile networks has become a serious issue for sustainable development due to the rapid growth of mobile devices [11, 12]. The importance of energy efficiency in future mobile generations is increasing for mobile network operators. The need for green communication has never been more important in the matter of reducing energy consumption. The establishment of green 5G networks will not only result in energy savings, but also in cost savings, carbon dioxide emissions reductions, and other benefits. The overall power consumption of 5G networks is also one of the key factors affecting network performance and user experience, which are of great importance [13]. For researchers, green 5G networks have become an intriguing topic in the past few years, particularly in terms of finding energy-efficient solutions.

Many authors have proposed various solutions to improve energy efficiency in order to enable sustainable 5G networks. Hawasli et al. [14] propose three algorithms to dynamically adapt the operation of small base stations to active/sleep (on/off) for non-uniform user distribution in heterogeneous networks, which is a promising technology for 5G networks, and investigate the general optimal power minimization problem. Al-Rubaye et al. [15] embedded an intelligent system into 5G ultra-dense networks to provision dense and irregular deployments that maintain extended coverage and also improve energy efficiency. Rostami et al. [16] introduced a novel wake-up signaling for the 5G control plane, aiming to reduce the energy consumption of the cellular module in the downlink. Arbi et al. [17] achieved energy savings in the 5G radio access network based on LTE technology, by implementing an adaptive sectorisation technique. Al-Quzweeni et al. [18] proposed an optical network supported architecture and investigated it to provide the wired infrastructure needed for 5G networks, and to support network function virtualization towards an

energy efficient 5G network. Ryoo et al. [19] designed the RRC connection control for 5G to enhance energy efficiency with moderate control signaling overhead. Saxena et al. [20] identified the complexity of optimal traffic awareness in 5G cloud radio access networks and designed a framework for traffic-aware energy optimization. Zhang et al. [21], benefiting from the assistance of macro cells, proposed a novel access scheme for heterogeneous cache-enabled 5G hyper cellular networks, according to both user interest and fairness of service, where the small cells can turn into a semi-sleep mode. Bouras et al. [22] present novel variations of sleep mode for 5G femtocells combined with hybrid access strategies and estimate capacity and energy benefits. Xiao et al. [23] proposed a unified placement model for energy efficiency of baseband functions and mobile edge computing in 5G networks.

#### IV. RESULTS AND DISCUSSION

The comparison of power consumption in different generations of mobile networks can be seen in Table I. It shows the development and progress of the frequency band, power density, mobile station, and base station power level through generations.

TABLE I. POWER TRENDS FROM 1G TO 5G [24]

Technology	Frequency band	Power density (Watt/M <sup>2</sup> )	Mobile station power level	Base station power level
1G	800 MHz	4.0	Low	Low
2G	850/900/1800/1900 MHz	4.5-9.0	GSM 850/900: 33dBm	Macro: 46 dBm
			GSM 1800/1900: 24-39dBm	Micro: 14-32 dBm
3G	800/850/900/1800/1900/2100 MHz	4.5-10	21-33 dBm	24-38 dBm
4G	1.8 GHZ, 2.6 GHZ	10	23 dBm	43-48 dBm
5G	30-300 GHz	10	High	High

From the 1980s, when the 1G network appeared, to the present day, when the 5G network is available, there has been significant progress in expanding the frequency range. Of course, such advancement has resulted in an increase in electricity consumption. Higher electricity consumption leads to higher carbon dioxide emissions but also to higher costs. That is why energy efficiency is of great importance, both for preserving the environment and for reducing costs for network companies.

Figure 1 shows the expected electricity usage of wireless access networks (2G-5G) by 2030. Based on the graph, it can be seen that the 2G network had the highest electricity consumption until 2021, when it was equaled by 4G. Following that, the 4G network is expected to be the largest consumer of electricity from 2022 until 2029, while in 2030, 5G is expected to overtake it.

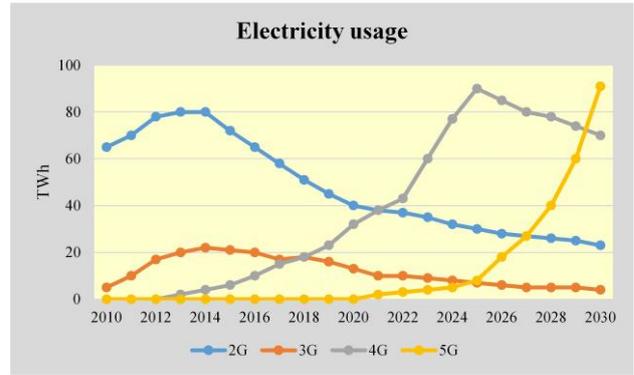


Figure 1. Expected case electricity usage of wireless access networks [25]

Taking into account this trend of increased energy consumption by 5G networks, an urgent need for energy-efficient solutions has emerged. Figure 2 summarizes the energy-saving findings, obtained from ten relevant publications listed in the green 5G network section. The results show a percentage improvement in energy savings after the implementation of the proposed solutions.

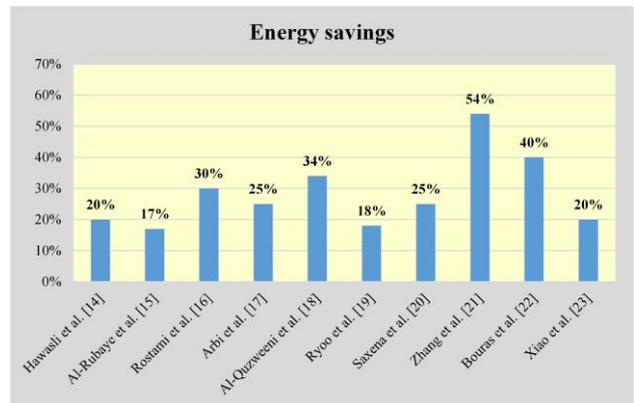


Figure 2. 5G network energy-saving findings comparison

This graph shows that the proposed solution in [21] reduces energy consumption by 54%, which is the largest energy savings among the alternatives. With only 17% energy efficiency improvements in a given sample, [15] had the lowest energy savings. The proposed solutions achieved 28.3% of energy savings on average. It is clear from the standard deviation of 11.67 that the energy savings percentages of the proposed solutions do not deviate significantly from the average energy savings percentage. With a value of 29.1%, the coefficient of variation suggests that the arithmetic mean is a satisfactory measure of central tendency, and that the data is satisfactorily grouped around the arithmetic mean.

#### V. CONCLUSION

The increased number of wireless networks together with the needed service quality for expanding data services has generated the need to innovate green solutions that effectively satisfy essential energy efficiency requirements. In this paper, an overview of energy efficiency initiatives in 5G is presented. The comparison of the results achieved by these solutions clearly indicates that 5G has leveraged a unique

paradigm, which has led to the creation of various concerns for future wireless networks, stressing the importance of better energy consumption optimization and overall sustainability.

#### REFERENCES

- [1] Ge, X., & Zhang, W. (2019). 5G Green Mobile Communication Networks. *Springer*.
- [2] Wu, Q., Li, G. Y., Chen, W., Ng, D. W. K., & Schober, R. (2017). An overview of sustainable green 5G networks. *IEEE Wireless Communications*, 24(4), 72-80.
- [3] Ashiho, L. S. (2003). Mobile technology: Evolution from 1G to 4G. *Electronics for you*, 6, 968-976.
- [4] Churi, J. R., Surendran, T. S., Tigdi, S. A., & Yewale, S. (2012). Evolution of networks (2G-5G). In *International Conference on Advances in Communication and Computing Technologies (ICACACT)* (Vol. 51, No. 4, pp. 8-13).
- [5] Kumaravel, K. (2011). Comparative study of 3G and 4G in mobile technology. *International Journal of Computer Science Issues (IJCSI)*, 8(5), 256-263.
- [6] Holma, H., & Toskala, A. (Eds.). (2005). WCDMA for UMTS: Radio access for third generation mobile communications. *John Wiley & Sons*.
- [7] Hwang, J. S., Consulta, R. R., & Yoon, H. Y. (2007). 4G mobile networks—Technology beyond 2.5 G and 3G. *PTC (Pacific Telecommunications Council) Proceedings, Hawaii*.
- [8] Khan, F. (2009). LTE for 4G mobile broadband: air interface technologies and performance. *Cambridge University Press*.
- [9] Bhalla, M. R., & Bhalla, A. V. (2010). Generations of mobile wireless technology: A survey. *International Journal of Computer Applications*, 5(4), 26-32.
- [10] Vook, F. W., Ghosh, A., & Thomas, T. A. (2014, June). MIMO and beamforming solutions for 5G technology. In *2014 IEEE MTT-S International Microwave Symposium (IMS2014)* (pp. 1-4). IEEE.
- [11] Chen, M., Hao, Y., Hu, L., Huang, K., & Lau, V. K. (2017). Green and mobility-aware caching in 5G networks. *IEEE Transactions on Wireless Communications*, 16(12), 8347-8361.
- [12] Zhang, S., Cai, X., Zhou, W., & Wang, Y. (2019). Green 5G enabling technologies: an overview. *IET Communications*, 13(2), 135-143.
- [13] Zhang, J., Zhang, X., & Wang, W. (2016). Cache-enabled software defined heterogeneous networks for green and flexible 5G networks. *IEEE Access*, 4, 3591-3604.
- [14] Hawasli, M., & Çolak, S. A. (2017). Toward green 5G heterogeneous small-cell networks: power optimization using load balancing technique. *AEU-International Journal of Electronics and Communications*, 82, 474-485.
- [15] Al-Rubaye, S., & Tsourdos, A. (2020). Airport connectivity optimization for 5G ultra-dense networks. *IEEE Transactions on Cognitive Communications and Networking*, 6(3), 980-989.
- [16] Rostami, S., Heiska, K., Puchko, O., Talvitie, J., Leppanen, K., & Valkama, M. (2018, December). Novel wake-up signaling for enhanced energy-efficiency of 5G and beyond mobile devices. In *2018 IEEE Global Communications Conference (GLOBECOM)* (pp. 1-7). IEEE.
- [17] Arbi, A., & O'Farrell, T. (2015, June). Energy efficiency in 5G access networks: Small cell densification and high order sectorisation. In *2015 IEEE International Conference on Communication Workshop (ICCW)* (pp. 2806-2811). IEEE.
- [18] Al-Quzweeni, A. N., Lawey, A. Q., Elgorashi, T. E., & Elmoghani, J. M. (2019). Optimized energy aware 5G network function virtualization. *IEEE Access*, 7, 44939-44958.
- [19] Ryoo, S., Jung, J., & Ahn, R. (2018, April). Energy efficiency enhancement with RRC connection control for 5G new RAT. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 1-6). IEEE.
- [20] Saxena, N., Roy, A., & Kim, H. (2016). Traffic-aware cloud RAN: A key for green 5G networks. *IEEE Journal on Selected Areas in Communications*, 34(4), 1010-1021.
- [21] Zhang, J., Zhang, X., Imran, M. A., Evans, B., & Wang, W. (2016, December). Energy efficiency analysis of heterogeneous cache-enabled 5G hyper cellular networks. In *2016 IEEE Global Communications Conference (GLOBECOM)* (pp. 1-6). IEEE.
- [22] Bouras, C., & Diles, G. (2017, March). Energy efficiency in sleep mode for 5G femtocells. In *2017 Wireless days* (pp. 143-145). IEEE.
- [23] Xiao, Y., Zhang, J., & Ji, Y. (2018, October). Energy efficient placement of baseband functions and mobile edge computing in 5G networks. In *Asia Communications and Photonics Conference* (pp. M3E-4). Optical Society of America.
- [24] Abrol, A., & Jha, R. K. (2016). Power optimization in 5G networks: A step towards GrEEen communication. *IEEE Access*, 4, 1355-1374.
- [25] Alsharif, M. H., Kelechi, A. H., Kim, J., & Kim, J. H. (2019). Energy efficiency and coverage trade-off in 5G for eco-friendly and sustainable cellular networks. *Symmetry*, 11(3), 408.

# Review Analysis of E-Government in the Republic of Serbia

M. Mazalica\*, B. Radulovic\* and A. Stojkov\*

\* University of Novi Sad / Technical Faculty "Mihajlo Pupin" Zrenjanin, Zrenjanin, Serbia  
milica.mazalica@tfzr.rs, biljana.radulovic@tfzr.rs, sahskica.s@gmail.com

**Abstract** – The new technological era brings many benefits and made everyday life much easier. The E-Government portal is one of them. This portal enables people to submit a request to the state administration from home without waiting in line. This portal has received increasing attention as an innovation method for improved governance and many countries are investing their efforts and resources into its development. The E-Government portal of Republic Serbia was launched in June 2010 and since then the government of Republic Serbia is constantly working to improve it. This paper is present a short analysis of this portal in the Republic of Serbia. The focus of this paper is on the development of this portal in a developing country in this case in the Republic of Serbia.

## I. INTRODUCTION

The modern age has brought many benefits that make everyday lives much easier, and one of them is the possibility to submit a request to the state administration from home without waiting in line. E-Government is the portal that enables finishing all procedures without physical arrival to the counter. This portal has received increasing attention as an innovation method for improved governance and many countries are investing their efforts and resources into its development.

E-Government (electronic government) is the use of technological communications devices to provide public services to citizens and other persons in a country or region, and also offers new opportunities for more direct and convenient citizen access to government and government provision of services directly to citizens. It also allows citizens to interact with public administration to achieved objectives at any time and any location 24 hours a day and seven days a week. There are three levels of interaction in the E-government: informing, communication and transaction. On the other hand interaction in E-Government with participants in the process is divided on the: Government to Citizen (G2C), Citizen to Government (C2G), Government to Business (B2G), Government to Employes (G2E), Government to Government ( G2G), Government to Nonprofit (G2N), Nonprofit to Government (N2G). For the establishment of E-Government, it is necessary to carry out certain preconditions such are socio-cultural prediction, organization, legal prediction, and technical prediction, but the basic prediction is global connectivity at the level of all administrative bodies and access to communication infrastructure for all E-Government users.

The first public service was established by Denmark in 1970 through the Central Tax Return System. In other countries, EU electronic public services begin operating in the late 1990s and early 2000s. The first of the new EU member's electronic public was introduced in 2000 by Estonia. In the Republic of Serbia, E-Government was launched in June 2010 and is published at [ww.euprava.gov.rs](http://ww.euprava.gov.rs). In February 2020 this portal was updated by the office for IT and E-Government. This new version was created based on a questionnaire filled by citizens and which contains answers to 300 of their questions based on real-life problems. The team from different state institutions worked together to write the answers which would be understandable to all citizens. The information and services on this portal are organized according to 8 areas of life events. [2]

This paper gives a short review of the E-Government portal in the Republic of Serbia. In this research, the focus is on the implementation of this portal in our country. This paper is structured as follows: in the introduction, a section is described E-Government and its portal, the next section gives a short review of the development of E-government in the state of the EU and Republic of Serbia. Section three gives services and modules of this portal in the Republic of Serbia, section four shows the technical implementation of this portal in the Republic of Serbia. The discussion and conclusion remarks are given at the end of this paper.

## II. OVERVIEW OF E-GOVERNMENT

### A. E-Government in EU

E-Government implies a way of organizing the public sector to increase efficiency, transparency, easier access and response to request citizens, through, intensive and strategic implementation of information and communication technologies in the public sector. [3] The concept of E-Government refers to three main target groups: administration, business sector and citizens. The influence of the public demand for central services depends on the choice of services that the government decides to offers. The EIF is a document published by the EU in which it has given guidance and recommendations to public administrations on how to improve governance of their interoperability activities, and how to establish digital services.[1] The EIF is a common greed approach to the delivery of European public services in an interoperable manner. It defines basic interoperability guidelines in the form of common

principles, models and recommendations.[1] In paper [4] is given research studies about interoperability and particularly interoperability frameworks in state of EU and state which are in progress for membership. The aim of this paper was to describe the effort and producer done in order to conduct the worldwide survey and presents the results achieved whit it.

The most important part of the implementation of E-Government is the contentment of socio-technical factors which affect the success of E-Government. The paper [5] presents the influence of Critical success factors (CSFs) on success implementation E-Government in both term technology and non-technology factors which were mapped in seven dimensions ITPOSMO (Information, technology, Process, Objective, Staffing and skill, Management and other Resource). This study in paper [5] showed that non-technical factors are more dominant than technical factors in the implementation of successful E-Government. In paper [6] is given quality model for measuring e-portal quality in a unified, reliable and easy manner, and identify as the missing best practices that could improve the quality for those e-portals. Security of e-portal is another important part of the implementation of E-government. Paper [7] presents a security framework implementing centralized access control for providing authentication and authorization functions to Web applications and services delivered by Public Administration and hosted in the cloud. The framework shown in [7] implements authentication according to the paradigm of federated SSO using SPID complaint certifies IdPs and authorization which is enforced according to the RBAC model. But the main method for integrating applications and services with the framework is via the OAuth2/OpenID Connect protocol. [7]

The implementation of E-Government in developing countries is a very complex and demanding process. The paper [8] presents a method for assessing e-government implementation in developing countries to help them better identify problems and determine the appropriate responses to achieve their goals.

#### B. E-Government in the Republic of Serbia

The internet portal is the "front door" to infinite virtual space (cyberspace). E-Government portal in Republic Serbia is a central access point to electronic services of the state administration. The users of the portal are citizens, legal entities and employees in state administration bodies and local government. The E-Government portal consists of several subsystems: identification, registrations and log in (using all qualified certificates in Republic Serbia), services generator, submitting a request, request processing, e-Schedule, e-Delivery, CMS (Content manager system), personalization of portal users based on a personal profile, E-Survey, E-Notifications. This portal was made on the EIF strategy for development government for the member states of the EU. Besides the E-Government portal exist M-Government application. In paper [8] is shown one solution of M-

Government application designed for finding answers for citizens questions, asked in the Serbian language. This paper also describes how to overcome the problem caused by legislation and also by the specific features of the Serbian language and as well as connect with existing e-Government applications. E-Government portal is based on big-data and cloud computing. In paper [9] is given an overview of current concepts related to cloud infrastructure design that should provide support for big data and a model of the big data infrastructure in e-government based on cloud computing and SDN concepts. The paper [10] and [11] shows the implementation of local e-government in the Republic of Serbia. In this paper are given an analysis of services that administration at the local level gives to citizens and businesses electronically.

### III. E-GOVERNMENT PORTAL IN THE REPUBLIC OF SERBIA

The E-Government portal of Republic Serbia is organized by a life according to events i.e. items such as family life, kindergarten, primary and secondary education retirement, studying, opportunities and opportunities for young people, looking for a job, purchase of the real estate, starting a business and retirement. The e-government portal has three types of users citizens, economy and country. Every model provides its users certain services. State bodies that provide services are local self-government, state administration, public services, preschool institution, primary education institutions, secondary education institutions, higher education institutions, centers for social work. [12]

The e-government portal offers services to citizens related to the issuance of personal documents ( such as passports and ID cards, drivers licenses, birth registration, weapons documents, qualified electronic certificate), children-education ( such as baby welcome, e-kindergarten, rejection, my high school, e-diary), health (such as health insurance, health care, information about appointment and result of the test on covid-19 all this is shown in Figure 3. ), traffic-tourism (such as vehicle registration, registration of residence of foreigners), certificates (criminal certificate), finance ( such as local tax administration, incentives, salaries, my data for the bank, digital assets), work ( such as personal training, competitions for public places), projects ( such as intellectual property, public calls), foreigner ( such as application for a work visa, residence registration). [12]

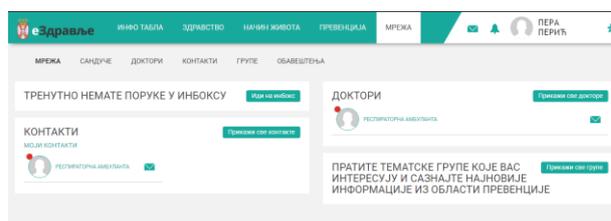


Figure 3. Users interface for E-Health

Services that are available to entrepreneurs on the e-government portal are: business establishment (company registration), trade (tobacco products), agriculture (plant protection), conviction (criminal convictions), health (such as regulation of drugs in human medicine, regulation of drugs in veterinary medicine, regulation of medical devices), property ( e-building permits), inspection services (e-inspector), tourism (tourist registration), traffic ( vehicle registration), weapons (weapons document), projects (project financing), digital assets ( such as white paper, providing services to entrepreneurs, providing services to legal entities).

On this portal are also available services for country administration such as trust services ( as timestamp, e-payment), registers ( as data exchange system, local tax administration, real estate in public ownership, open data, central population register), training (exam for inspector) and seal (consent).

E-government allows three types of user registration: registration with a qualification electronic certificate, registration with two-factor authentication and registration with username and password. In Figure 1. is shown steps for user registration with user name and passwords. After registration with username and password, users can see the interface which is shown in Figure 2.

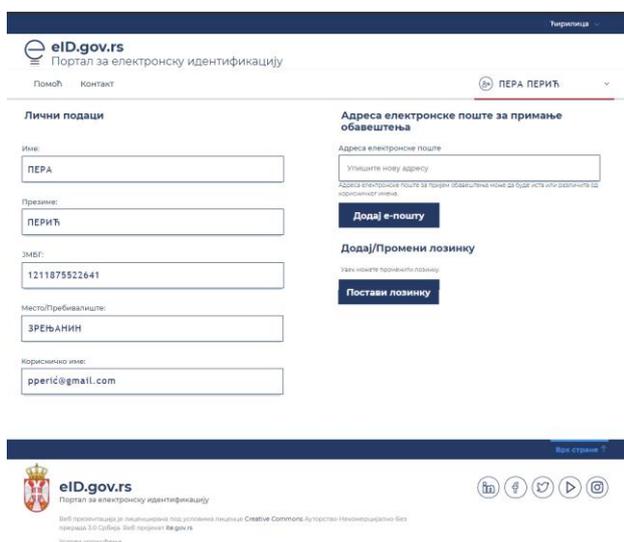


Figure 1. Users steps for registration

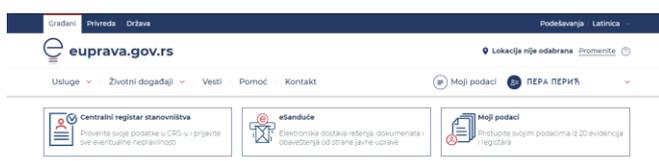


Figure 2. User interface after registration

### A. CMS Modules and External systems

Content Management System is software that helps with creating and managing the content of the website. Also

help in tracking the user sessions, handling the search queries, collecting visitor feedback and comments, hosting forums on the website [13]. The E-Government represents the collection of separate modules which are all developed in CMS technologies and combine in one functional system. These are Module for generating electronic services, Electronic application submission module, eSchedule, Module for processing electronic requests, ePolls. In addition to these four modules, there are also those that are separate units in relation to the e-Government portal. The modules in the structure of the E-Government portal: [14]

- **Module for generating electronic services** are the central part of the system. This module is for state employees and provides simple and successful service creation. Include defining the level of authentication at startup, service description structured by paragraphs, defining the attachments, creating an electronic form and defining payment parameters.
- **The electronic application submission module** provides online application requests.
- **eSchedule** is used to schedule an appointment when the user needs to come to the institution in person.
- **Module for processing electronic requests** takes electronic requests to be processed. Also Used for change status of the case, communication with citizens and delivery of solutions in electronic form.
- **ePolls** allow state authorities to create questionnaires. After users respond, data are presented in the form of statistical reports.

The separate systems attached in the E-Government portal: [14]

- **The system for issuing registration stickers** is a complete solution used to register vehicles. Authorized users carry out the entire procedure without going to the police station. The module exchanges information with systems in the Ministry of Internal Affairs, the Business Registers Agency and the Association of Insurers of Serbia.
- **Newborn registration system in maternity hospitals** used by authorized persons which keeps records of newborns and registers them in the registry book assigns residence and health and social insurance. This system uses the resources of the Ministry of State Administration and Local Self-Government, the Ministry of Interior, Central Register of Compulsory Social Insurance, Pension and Disability Insurance and the Republic Institute for statistics.
- **The candidate training system for driving** contains modules that are in charge of recording all the processes of training candidates for driving. Candidates have their own profile which is used for checking the passed lessons, simulation of tests, final testing, results.... The

system is connected to the information system of the Ministry of Internal Affairs and actively exchanges data.

- **eKindergarten** enables parents to register their children for preschool education and employees in preschool institutions provide insight into the registry books and a central register of compulsory social insurance.
- **eReminder** electronic service intended to remind citizens about the forthcoming expiration date of the documents.
- **eParticipations** enables electronic enforcement of public discussion and has two modules. The first is used by authorized persons from state administration bodies. The second module is used to submit comments.

The government service bus is the central part of exchanging data between applications. This system is extended by a special application that allows parameterization authorization parameters for granting access rights to web services located on the platform. The interoperability platform allows applications to communicate via unique platforms. [14]

#### IV. DEVELOPMENT TECHNOLOGIES

The e-government portal consists of a basic application and many other systems that complement the functions of the portal. The portal and most other systems rest on the same development technology but some of the systems have different developing technologies. The core of the portal is developed on Microsoft .NET platform version of Microsoft .NET Framework 4.0 with C # programming language and for the user interface was used JavaScript, jQuery, AJAX and CSS. For database is used SQL Server 2008 R2 and as application server Internet Information Services 7.5 (IIS 7.5). [14]

Three subsystems are developed in different technologies they are System for registration of newborns in maternity hospitals, a Candidate training system for driving and eParticipations. The difference is in the framework. These three systems instead of .NET Framework 4.0 uses ASP.NET MVC Framework. [14]

The government service bus uses BizTalk Server 2010 as a basis. The application was developed using the Microsoft .NET platform with C # programming language, while Microsoft SQL Server 2008 R2 is used for the database. [14]

#### V. CONCLUSION

E-Government is the portal that enables finishing all procedures without physical arrival to the counter. This portal has received increasing attention as an innovation method for improved governance and many countries are investing their efforts and resources into its development.

E-Government portal in Republic Serbia is a central access point to electronic services of the state administration. The users of the portal are citizens, legal entities and employees in state administration bodies and local government. The E-Government portal of Republic

Serbia is organized by a life according to events i.e. items such as family life, kindergarten, primary and secondary education retirement, studying, opportunities and opportunities for young people, looking for a job, purchase of the real estate, starting a business and retirement The E-Government Republic of Serbia contains separate modules which are all develops in CMS technologies and combine in one functional system. The core of the portal is developed on Microsoft .NET platform version of Microsoft .NET Framework 4.0 with C # programming language and for the user interface was used JavaScript, jQuery, AJAX and CSS. For database is used SQL Server 2008 R2 and as application server Internet Information Services 7.5 (IIS 7.5).

In our country after the appearance of Covid-19 this portal came to life. The pandemic forced people to use this portal, and also our government to develop and improve this portal. Nowadays E-government portal of Republic Serbia enable people to finish they administrative procedure without physical arrive to the counters such as E-Kindergarten, but for finishing some administrative procedure for example personal documents they still need to visit public counters. Our government is hard-working to overcome this problem and improve this portal as soon as.

#### REFERENCES

- [1] New European Interoperability Framework, European Union, 2017, ISBN 978-92-79-63756-8, doi:10.2799/78681
- [2] The office for information technologies and E-Government, Abot as and Inforamtion, <https://www.ite.gov.rs/>, Accessed August 2021.
- [3] J. F. Affisco, K. S. Soliman, "E-government: a strategic operations management framework for service delivery", 2006, Business Process Management Journal, Vol. 12, pp. 13 – 21, doi: <http://dx.doi.org/10.1108/14637150610643724>
- [4] A. Lisboa, D. Soares, "E-Government interoperability frameworks: a worldwide inventory", Procedia Technology, Vol. 16, pp 638 – 648, 2014.
- [5] D. Naptupulu, D. I. Sensure, Y. G. Sucahyo, "Socio-Technical Factors of E-Government Implementation", 2017 International Conference on Electrical Engineering Computer Science and Informatics (EECSI 2017), Yogyakarta, Indonesia, 2017.
- [6] A. Fath-Allah, L. Cheikhi, A. Idri, R. E. Al-Qutaish, "A Best Practice Based E-Government E-portals Quality model: A detailed view", 2018 6th International Conference on Control Engineering & Information Technology (CEIT), Istanbul, Turkey, 2018.
- [7] A. Bettacchi, B. Re, A. Polzonetti, "E-Government and Cloud: Security Implementation for Services", 2017 Fourth International Conference on a eDemocracy & eGovernment (ICEDEG), Quito, Ecuador, 2017.
- [8] Dz. Avdic, A. Avdic, Z. Spalevic, U. Marovac, A. Crisanin, "M-Government Application Intented to Serach Documents Written in Serbian Language", Sinteza, 2014.
- [9] J. Suh, V. Vujin, D. Barac, Z. Bogdanovic, B. Radenkovic, "Desinging Cloud Infrastructure for Big Data in E-Government", Journal of Universal Excellence, Vol. 1, pp. A26–A38, 2015.
- [10] T. Lolic, D. Dakic, I. Mikovic, D. Stefanovic, "Electronic Governemnt at the Local Level in the Republic of Serbia", 10th International Scientific Conference "Science and Higher Education in Function of Sustainable Development", Uzice, Serbia, 2017.

- [11] M. Paroski, Z. Konjovic, D. Surla, V. Popovic, "Development of eGovernment services in the Autonomous Province of Vojvodina", Information Technology for Development, Vol. 21, No. 3, pp 492– 510, 2015.
- [12] Porta E-Goveremnt in the Republic of Serbia, <https://euprava.gov.rs/>, Accessed August 2021.
- [13] Velocity Consultancy, CMS website, <https://www.velocityconsultancy.com/what-is-a-cms-website/>, Accessed August 2021.
- [14] The office for information technoogies and E-Governemnet, Abot as and Inforamtion, Archive, <http://arhiva.ite.gov.rs/doc/tenderi/2018/JNO1/KD%20Odrzavanje%20Portala%20eUprava%2010042018%20konacno.pdf>, Accessed August 2021.

# A Survey of Machine Learning Techniques Used in Recommender Systems

Nora PireciSejdiu, Blagoj Ristevski, Ilija Jolevski

University "St. Kliment Ohridski" – Bitola

Faculty of Information and Communication Technologies – Bitola, North Macedonia

{pireci.nora, blagoj.ristevski, ilija.jolevski}@uklo.edu.mk

**Abstract** - The amount of data on the internet is increasing every day, and this is leading to information overload. Recommendation systems have been very efficient by significantly improving information retrieval in data mining. Machine learning techniques have been used to improve the performance of recommendation systems, which are typically trained to maximize user satisfaction and increase sales in business companies. This paper deals with the research of machine learning techniques, which are used in recommendation systems in different fields improving performance.

**Keywords** – recommender systems, content-based filtering, collaborative filtering, hybrid filtering, machine learning techniques, decision tree, naïve Bayes, k-nearest neighbor, neural networks, support vector machines.

## I. INTRODUCTION

The rapid increase in the amount of information in digital technologies has created a challenge for users in accessing relevant documents. Recommender systems (RS) are information filtering systems that deal with the problem of information overload by filtering vital information fragment out of a large amount of dynamically generated information according to the user's preferences, interest, or observed behavior about an item. Recommender systems are beneficial to both service providers and users [21]. Many Researchers and organizations use data mining to extract useful knowledge regarding their needs. Data mining covers many techniques such as classification, clustering, regression, association rules, summarization, time series analysis etc. Each technique has some algorithms like classification has decision trees, naïve Bayes, neural networks etc., while clustering has k-means etc. [8] [26] [27] [28].

Recommender Systems propose useful and interesting items to users in order to increase both seller's profit and the buyer's satisfaction. They contribute to the commercial success of many online ventures such as Amazon or NetFlix and are a very active research area. Examples of recommended items include movies, web pages, books, news items and more. Often an RS attempts to predict the rating a user will give to items based on her past ratings and the ratings of other (similar) users [12]. Due to the explosion of e-commerce, recommender systems are rapidly becoming a core tool to accelerate cross-selling and strengthen customer loyalty. There are two prevalent approaches for building recommender

systems – content-based recommending and collaborative filtering (CF). The CF algorithm is one of the most common recommender system algorithms. This study focuses on improving the performance of recommender systems by using data mining techniques [19]. To improve the quality of recommendations, machine learning techniques are used in recommendation systems. The most popular techniques are decision tree classifier, naïve Bayes, k-nearest neighbor, neural network and support vector machines.

This paper surveys the application of machine learning techniques in recommendation systems based on existing literature where recommendation systems have been proven to improve the process and quality of decision making depending on the technique used.

The rest of the paper is structured as follows. Section II describes the recommendation systems. The commonly used machine learning techniques are described in Section III. The subsequent section depicts the related work and compares the obtained results when different techniques are used. Concluding remarks are highlighted in the last section.

## II. RECOMMENDATION SYSTEMS

The explosive growth in the amount of available digital technologies and information and the number of visitors to the Internet has created a potential challenge of information overload which hinders timely access to items of interest on the Internet. Recommender systems are beneficial to both service providers and users. They reduce transaction costs of finding and selecting items in an online shopping environment. Recommendation systems have also proved to improve the decision making process and quality [21].

The use of efficient and accurate recommendation techniques is very important for a system that will provide a good and useful recommendation to its individual users. This explains the importance of understanding the features and potentials of different recommendation techniques. Figure 1 shows the anatomy of different recommendation filtering techniques [21].

Methods that use the former are referred to as collaborative filtering methods, whereas methods that use the latter are referred to as content-based recommender methods.

Note that content-based systems also use the rating matrices in most cases, although the model is usually focused on the ratings of a single user rather than those of all users. In knowledge-based recommender systems, the recommendations are based on explicitly specified user requirements. Instead of using historical rating or buying data, external knowledge bases and constraints are used to create the recommendation. Some recommender systems combine these different aspects to create hybrid systems. Hybrid systems can combine the strengths of various types of recommender systems to create techniques that can perform more robustly in a wide variety of settings. In the following, we will discuss these basic models briefly, and also provide pointers to the relevant chapters in the book where they are discussed [2].

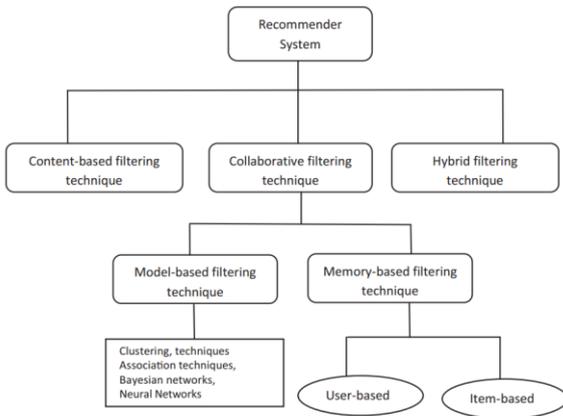


Figure 1. Recommendation techniques.

A. Content-based filtering

In content-based recommender systems, the content plays a primary role in the recommendation process, in which the ratings of users and the attribute descriptions of items are leveraged in order to make predictions. The basic idea is that user interests can be modeled based on the properties (or attributes) of the items they have rated or accessed in the past [2]. Content-based systems examine the properties of the items recommended. As shown in Figure 2 [25] content-based recommenders rely on the fact that a user is interested in items similar to those he liked (purchased, searched, browsed, etc.) in the past.

They entail the description of items that may be recommended, the creation of a profile describing the types of items the user likes, and a strategy that compares

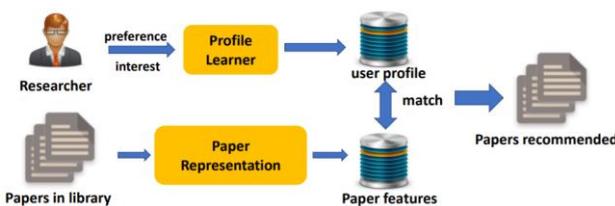


Figure 2. Content-based system for a paper recommendation.

item and user profiles to determine what to recommend [17]. Content-based recommendation systems may be used in a variety of domains ranging from recommending web pages, news articles, restaurants, television programs, and items for sale.

B. Collaborative filtering

Collaborative filtering (CF) is a popular recommendation algorithm that bases its predictions and recommendations on the ratings or behavior of other users in the system [13]. Collaborative filtering is also referred to as social filtering as it filters information by using the recommendations of other people. Collaborative filtering recommender systems recommend items by identifying other users with similar tastes and use their opinions for a recommendation. Collaborative filtering explores techniques for matching people with similar interests and making recommendations on this basis [16]. The term “collaborative filtering” refers to the use of ratings from multiple users in a collaborative way to predict missing ratings. In practice, recommender systems can be more complex and data-rich, with a wide variety of auxiliary data types [2]. Collaborative filtering systems focus on the relationship between users and items as illustrated in Figure 3 [25]. The similarity of items is determined by the similarity of the ratings of those items by the users who have rated both items [1]. There are two types of methods that are commonly used in collaborative filtering, which are referred to as memory-based methods and model-based methods.

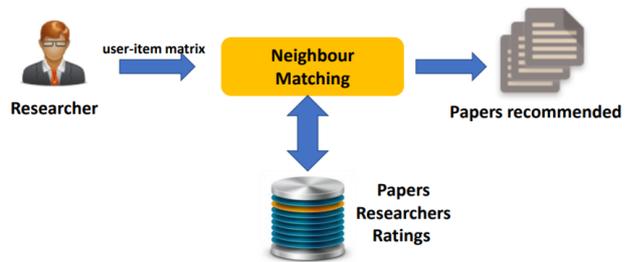


Figure 3. Collaborative filtering system for a paper recommendation.

C. Hybrid filtering

Hybrid filtering, in many cases where a wider variety of inputs is available, one has the flexibility of using different types of recommender systems for the same task. In such cases, many opportunities exist for hybridization, where the various aspects from different types of systems are combined to achieve the best of all worlds. Hybrid recommender systems are close to the field of ensemble analysis, in which the power of multiple types of machine learning algorithms is combined to create a more robust model. Ensemble-based recommender systems can combine not only the power of multiple data sources, but they are also able to improve the effectiveness of a particular class of recommender systems (e.g., collaborative systems) by combining multiple models of the same type as shown in Figure 4 [25]. This scenario is not very different from that of ensemble analysis in the field of data classification.

The hybrid method of recommendation is also employed in many applications, such as the temporal purchase patterns derived from sequential pattern analysis (SPA) [4]. The hybrid method of recommendation is also employed in many applications, such as the temporal purchase patterns derived from sequential pattern analysis (SPA). On one hand, these applications derived implicit ratings that can be used in online transaction data for collaborative filtering. On the other hand, these applications used temporal purchase patterns to eliminate the harmful effect on recommendation services through SPA [17].

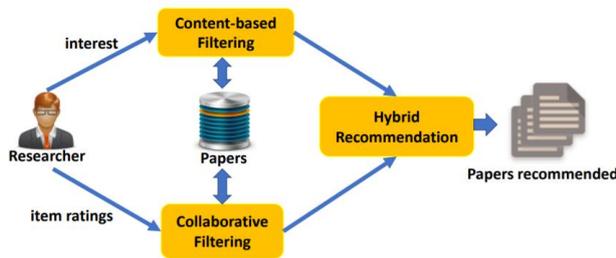


Figure 4. Content-based system for a paper recommendation.

### III. MACHINE LEARNING TECHNIQUES

Many algorithms are today classified as “machine learning”. All algorithms for the analysis of data are designed to produce a useful summary of the data, from which decisions are made. However, algorithms called “machine learning” not only summarize our data; they are perceived as learning a model or classifier from the data and thus discover something about data that will be seen in the future. Machine learning enthusiasts often speak of clustering with the neologism “unsupervised learning”; the term unsupervised refers to the fact that the input data does not tell the clustering algorithm what the clusters should be [27] [28]. In supervised machine learning the available data includes information about the correct way to classify at least some of the data. The data classified already is called the training set [1].

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention [13].

There are many different types of classification tasks in machine learning and specialized approaches to modeling but the most popular Machine Learning techniques are:

- Decision trees have been previously used as a model-based approach for recommender systems. The use of decision trees for building recommendation models offers several benefits, such as efficiency and interpretability and flexibility in handling a variety of input data types (ratings, demographic, contextual, etc.) [12]. The general idea of the algorithm Tree structure that has been widely used is to represent classification

models. Most decision tree induction algorithms are based on a greedy top-down recursive partitioning strategy for tree growth. They use different variants of impurity measures, like; information gain, gain ratio, and distance-based measures to select an input attribute to be associated with an internal node [6]. Decision Trees are classifiers on a target attribute or class in the form of a tree structure. The observations or items to classify are composed of attributes and their target value. The nodes of the tree can be: a) decision nodes, in these nodes a single attribute-value is tested to determine to which branch of the sub-tree applies or b) leaf nodes that indicate the value of the target attribute [22].

- Another very popular linear classification algorithm is naïve Bayes. It has been applied to text categorization with reasonable performance, although the performance is significantly worse than that achieved by regularized linear classifiers such as support vector machines. It is also very suitable for online updating, which could be important in practice [7]. One can treat the items as features and users as instances in order to infer the missing entries with a classification model [2]. Naïve Bayesian classifiers assume that there are no dependencies amongst attributes. This assumption is called class conditional independence. It is made to simplify the computations involved and, hence is called "naive" [3].
- K-nearest neighbor algorithm is one of the simplest non-parametric lazy algorithms called "Closest Point Search" is a mechanism that is used to identify the unknown data point based on the nearest neighbor whose value is already known. It is easy to understand but has incredible work in fields and practice especially in classification. It does not use training data to do generalization, that and in the best case, it makes a decision based on the entire training data set [6]. The k-nearest neighbor algorithm (k-NN) is a method to classify an object based on the majority class amongst its k-nearest neighbors. k-NN algorithm usually uses the Euclidean or the Manhattan distance. However, any other distance such as the Chebyshev norm or the Mahalanobis distance can also be used [3].
- A neural network is a set of connected input and output units in which each connection has a weight associated with it. Neural network learning is also referred to as connectionist learning due to the connections between units. It involves a long training process that requires a number of parameters for the classification of categories. A backpropagation neural network is a multilayer, feed-forward neural network consisting of the input layer, a hidden layer and an output layer. The neurons present in the hidden layer and output layer have biases, which are connected from units whose activation function is always 1. The bias term also acts as weights [9]. Artificial

neural networks (ANNs) are also usually applied to traffic prediction problems because of their advantages, such as their capability to work with multi-dimensional data, implementation flexibility, generalizability, and strong forecasting power [23].

- Support Vector Machines (SVMs) are one of the most theoretically well-motivated and practically most effective classification algorithms in modern machine learning [24]. Support Vector Machines (SVM) improves upon perceptrons by finding a separating hyperplane that not only separates the positive and negative points but does so in a way that maximizes the margin – the distance perpendicular to the hyperplane to the nearest points. The points that lie exactly at this minimum distance are the support vectors. Alternatively, the SVM can be designed to allow points that are too close to the hyperplane, or even on the wrong side of the hyperplane, but minimize the error due to such misplaced points

#### IV. RELATED WORKS

Due to the extensive use of computers, smartphones and high-speed Internet, people are now using the web for social contacts, business correspondence, e-marketing, e-commerce, e-surveys, etc. [2]. In recent years, recommendation systems have changed the way of communication between both websites and users. The recommendation system sorts through massive amounts of data to identify the interest of users and makes the information search easier [20].

The application of machine learning techniques in recommender systems has shown very good results and has significantly improved the performance of recommender systems. This is evidenced in related works that have been done so far. In the following, we can see the application of machine learning techniques in different areas of recommendation systems

In paper [3] the authors used naïve Bayes, decision Tree, and k-Nearest Neighbor algorithms to propose a new method to search alternative designs in an energy simulation. Their experiment shows that the decision tree has the fastest classification time followed by naïve Bayes and k-nearest neighbor. Based on the Precision, Recall, F-measure, Accuracy, and AUC value, the performance of naïve Bayes is the best.

In the paper [10] authors compared machine learning methods depending very much on the characteristics of a particular data set and the requirements of the respective business domain. This case study provides an assessment of the predictive performance of different classification methods for campaign management. The evaluation of data mining methods for marketing campaigns has special requirements. Whereas, typically the overall performance is an important selection criterion, for campaign management it is more important to select the technique which performs best on the first few quantiles. The C4.5 decision tree inducer in WEKA produced good results, in particular for the first few percentiles of the Gain Curve. K-nearest neighbor (k-NN) methods provided good results

only after irrelevant attributes were removed from the data set. There are a number of useful extensions one can perform in addition to the steps described in this paper.

In [12] a new method for decision tree-based recommender systems is proposed. The proposed method includes two new major innovations. First, the Decision Tree produces lists of recommended items at its leaf nodes, instead of single items. The second major contribution of the paper is the splitting method for constructing the decision tree. Splitting is based on a new criterion - the least probable intersection size. The proposed decision-tree-based recommendation system was evaluated on a large sample of the MovieLens dataset and is shown to outperform the quality of recommendations produced by the well-known information gain splitting criterion.

In this research work [13] a movie recommender system is built using the k-means clustering and k-nearest neighbor algorithms using The MovieLens dataset. The proposed work deals with the introduction of various concepts related to machine learning and recommendation system. In this work, various tools and techniques have been used to build recommender systems. It is seen that after implementing the system in the python programming language the root mean square error (RMSE) value of the proposed technique is better than the existing technique. It is also seen that the RMSE value of the proposed system is achieving the same value as the existing technique but with less no of clusters.

In paper [14] a comparison has been performed between different collaborative filtering algorithms to assess their performance. They evaluated k-nearest neighbor (k-NN), Slope One, co-clustering and non-negative matrix factorization (NMF) algorithms. k-NN algorithm is representative of the memory-based collaborative filtering approach (both user-based and item-based). The other three algorithms, on the other hand, are under the model-based collaborative filtering approach. They used the MovieLens dataset based on six evaluation metrics and resulted that the k-NN algorithm for item-based collaborative filtering outperformed all other algorithms examined in this paper.

In [5] authors propose a Bayesian methodology for recommender systems that incorporates user ratings, user features, and item features in a single unified framework. The key advantage of this approach is that it can use all the available information in a unified, coherent model.

In paper [15], a collaborative filtering based recommender system is improved by the ask-to-rate technique to solve the cold start problem. This paper determines the optimal number of neighbors in the item-based collaborative filtering k-NN algorithm after the login of the new user to the recommender system. After implementing the new user signup process framework, the results indicate that an optimal number of neighbors for the new user is 5 to 15 following the used standard dataset. If the number of neighbors is considered greater than 15, more neighbors with negative similarity will be involved in calculating the item rate prediction for the new user, reducing the accuracy of recommendations. If the number of neighbors is considered less than 5, no

neighbors may be found for the user. In these conditions, it is proposed that the  $k$  value is not constant for everyone, and only positive neighbors for each user are considered,  $k$  or number of the new user's neighbors is determined by an experimental evaluation.

Paper [11] evaluates the performance of ten different recurrent neural networks (RNN) structures on the task of generating recommendations using written reviews. The RNN structures they studied include well know implementations such as multi-stacked bi-directional Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) as well as the novel implementation of attention-based RNN structure. The attention-based structures are not only among the best models in terms of prediction accuracy, they also assign an attention weight to each word in the review.

A unique switching hybrid recommendation approach is proposed in [18] by combining a naïve Bayes classification approach with collaborative filtering. Experimental results on two different data sets have shown that the proposed algorithm is scalable and provide better performance—in terms of accuracy and coverage—than other algorithms while at the same time eliminating some recorded problems with the recommender systems.

## V. CONCLUSION

This survey focused on the existing literature and explored the application of machine learning techniques in recommender systems. Based on our research and in previous works as well as recent ones, it is well known that no algorithm can perform and give satisfactory results in all areas of use. Some algorithms may perform very well in a given field but in a specific field, another algorithm may perform better.

The biggest challenge of recommender systems is the so-called Cold Start problem, and it continues to be one of the key areas of current and future research.

## REFERENCES

- [1] Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, "Mining of Massive Datasets", Chapter 9, (2010)
- [2] Charu C. Aggarwa. "Recommender Systems: The Textbook". Springer International Publishing Switzerland, (2016)
- [3] Ahmad Ashari, Iman Paryudi, A Min Tjoa, "Performance Comparison between Naïve Bayes, Decision Tree and k-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool". International Journal of Advanced Computer Science and Applications, Vol. 4, No. 11, (2013)
- [4] Shouxian Wei, Xiaolin Zheng, Deren Chen, Chaochao Chen, "A hybrid approach for movie recommendation via tags and ratings". Electronic Commerce Research and Applications, Volume 18, Pages 83-94, (2016)
- [5] M. Condliff, D. Lewis, D. Madigan and C. Posse, "Bayesian Mixed-Effects Models for Recommender Systems", Proc. ACM SIGIR '99 Workshop Recommender Systems: Algorithms and Evaluation, (1999)
- [6] Delveen Luqman Abd AL-Nabi, Shereen Shukri Ahmed, "Survey on Classification Algorithms for Data Mining: (Comparison and Evaluation)". Computer Engineering and Intelligent Systems www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online) Vol.4, No.8, (2013)
- [7] T. Zhang, S.I. Vijay, "Recommender systems using linear classifiers". Journal of Machine Learning Research 2, pp. 313-334, (2002)
- [8] M Mohanapriya, J Lekha, "Comparative study between decision tree and knn of data mining classification technique". Second National Conference on Computational Intelligence, (2018)
- [9] Anuradha Patra, Divakar Singh, "A Survey Report on Text Classification with Different Term Weighing Methods and Comparison between Classification Algorithms". International Journal of Computer Applications (0975 – 8887) Volume 75– No.7, (2013)
- [10] Bichler, Martin and Kiss, Christine, "A Comparison of Logistic Regression, k-Nearest Neighbor, and Decision Tree Induction for Campaign Management". AMCIS 2004 Proceedings. 230, (2004)
- [11] David Zhan Liu, Gurbir Singh, "A Recurrent Neural Network Based Recommendation System". Technical Report: Stanford University, (2016)
- [12] Amir Gershman, Amnon Meisels, Karl-Heinz Luke, Lior Rokach, Alon Schclar, Arnon Sturm, "A Decision Tree Based Recommender System". 10th International Conference on Innovative Internet Community Systems (I2CS) – Jubilee Edition 2010 –. Bonn: Gesellschaft für Informatik e.V. (S. 170-179), (2010)
- [13] Rishabh Ahuja, Arun Solanki, Anand Nayyar, "Movie Recommender System Using K-Means Clustering AND K-Nearest Neighbor". 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), (2019)
- [14] Maryam Al-Ghamdi, Hanan Elazhary, Aalaa Mojahed, "Evaluation of Collaborative Filtering for Recommender Systems". (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 3, (2021)
- [15] A.Mojdeh Bahadorpour, B.Behzad Soleimani Neysiani, C.Mohammad Nadimi Shahraki, "Determining Optimal Number of Neighbors in Item-based kNN Collaborative Filtering Algorithm for Learning Preferences of New Users". ISSN: 2180 – 1843 e-ISSN: 2289-8131 Vol. 9 No. 3, (2017)
- [16] Sayali D. Jadhav, H. P. Channe, "Efficient Recommendation System Using Decision Tree Classifier and Collaborative Filtering". International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 03 Issue: 08, (2016)
- [17] Iván Cantador, Desmond Elliott, Joemon M. Jose, "A Case Study of Exploiting Decision Trees for an Industrial Recommender System". European Commission under contracts FP6-027122-SALERO, FP6-033715-MIAUCE, and FP6-045032 SEMEDIA, (2009)
- [18] M Ghazanfar, A Prugel-Bennett "An Improved Switching Hybrid Recommender System Using Naive Bayes Classifier and Collaborative Filtering". Proceedings of the International MultiConference of Engineers and Computer Scientists, IMECS, Vol. 1, (2010)
- [19] Min SH., Han I. "Recommender Systems Using Support Vector Machines". In: Lowe D., Gaedke M. (eds) Web Engineering. ICWE 2005. Lecture Notes in Computer Science, vol 3579. Springer, Berlin, Heidelberg, (2005)
- [20] Paritosh Nagarnaik, A.Thomas, "Survey on Recommendation System Methods". IEEE sponsored 2nd International Conference on Eletronics and Communication System, (2015)
- [21] F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh, "Recommendation systems: Principles, methods and evaluation". Egyptian Informatics Journal, Volume 16, Pages 261-273, (2015)
- [22] Maizan Mat Amin, Jannifer Yep Ai Lan, Mokhairi Makhtar1, Abd Rasid Mamat, "A Decision Tree Based Recommender System for Backpackers Accommodations". International Journal of Engineering & Technology, 7 (2.15) 45-48, (2018)
- [23] Xiaolei Ma, Zhuang Dai, Zhengbing He, Jihui Ma, Yong Wang and Yunpeng Wang, "Learning Traffic as Images: A deep convolutional neural network for large-scale transportation network speed prediction". Special Issue Sensors for Transportation, (2017)
- [24] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar, "Foundations of Machine Learning". Massachusetts Institute of Technology, (2012)
- [25] Xiaomei Bai, Ivan Lee, Xiangjie Kong, Feng Xia, "Scientific Paper Recommendation: A Survey" IEEE Access PP(99):1-1, (2019).

- [26] Beyene, Solomon Shiferaw, Tianyi Ling, Blagoj Risteovski, and Ming Chen. "A novel riboswitch classification based on imbalanced sequences achieved by machine learning." *PLoS computational biology* 16, no. 7 (2020): e1007760.
- [27] Risteovski, Blagoj, Suzana Loshkovska, Sasho Dzeroski, and Ivica Slavkov. "A comparison of validation indices for evaluation of clustering results of DNA microarray data." In *2008 2nd International Conference on Bioinformatics and Biomedical Engineering*, pp. 587-591. IEEE, 2008.
- [28] Risteovski, Blagoj, and Suzana Loskovska. "Bayesian networks application for representation and structure learning of gene regulatory networks." In *2009 12th International Conference on Computers and Information Technology*, pp. 609-613. IEEE, 2009.

# Storage, selection and visualization raster models in spatial databases

A. Vasiljević\*, M. Čeliković\*, S. Popov\*

\*University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia  
alexva02@uns.ac.rs, milancel@uns.ac.rs, srdjanpopov@uns.ac.rs

**Abstract** – The smart city project requires constant collecting data of various formats and informativeness, whose processing should result in software solutions intended for smart management. Project specificity of such proportions implies reliance on narrow specialized database management system. PostgreSQL is a relational database that could solve the problem of storing enormous amount of data, but it cannot be adequately used for storage purposes data such as raster images without additional extensions. Realization of systems based on spatial data implies the application of specially optimized systems for data storage, of which it is currently the most popular PostGIS. Geoinformatics is the area in which is dominant working with geospatial data – the way they are modeling, storage, retrieval, selection and visualization occupies a central place in the work.

**Keywords** – spatial data, spatial databases, raster data, PostgreSQL, PostGIS, QGIS

## I. INTRODUCTION

The modern age has brought numerous innovations in science and industry: new scientific disciplines have developed, they have emerged new potential problems to be solved, the idea of applying computer intelligence has crystallized, the need to monitor various types of geographical terrain has developed, the presumption of smart city existence is formed, etc.

Positive consequence of the mentioned shifts with the largest potential spectrum of appliance is the possibility of working with spatial data. Thanks to constant observing areas of interest, it is now possible to execute prediction and visualization of (a)typical terrain. In order to define the preconditions for implementing a vision of smart city, it is necessary to research and study the existing database types, their extensions and optimizations for narrow specific scope. Nowadays, the database that proved to be the most optimal in working with geospatial data is PostgreSQL. [1]

Main feature of relational databases is their interconnection with relations that are manifested in the form of tabular data. As a representative specimen the PostgreSQL database was taken, due to open source and extensibility. The smart city project is unique because the data which it manipulates with must be geospatial character, which means that any data individually carries a greater amount of information. The following chapters will explain the concepts of spatial data and spatial databases, the process of modeling spatial data and the

query creation, with detail explanation of possible ways for indexing data – this will be illustrated with a few examples. This is followed by a detailed review of PostgreSQL and its specially optimized extension called PostGIS. Additionally, a solution for efficiently of working with geospatial data called QGIS will be expound.

## II. SPATIAL DATA AND SPATIAL DATABASES

The term spatial data is defined as a special type of data which is characterized by the definition of attributes, but also geometry of real models of real world occurrences. Main reason for applying spatial data in geographical area of interest is their close correlation to space. Important property of spatial data is indirect and/or immediate description of individual positions. This group of computer data can be presented graphically or non-graphically. Information carried by members of this group are also a form, descriptive attributes and relation to other spatial data.

Geoinformation systems have created a good basis for appliance of spatial data - support for managing data itself and for managing their properties is now enabled. Similar systems organize and store real world information in the form of a collection of thematic layers interconnected by geography. The need for data standardization, how it could have an expanded range of appliance, occurs as a consequence of the development of information and communication technology, spatial databases, capabilities publication and further studies, requests for visualization and distribution over the network, etc. [2]

What makes spatial data different from other data groups are the following properties: coordinates in coordinate system, latitude and longitude, name of observed geographic area, zip code (if any), distances and orientation in relation to other entities and similar essential descriptive attributes. In addition to what spatial data have different, next to characteristics, there are specific format and standards and possibility of storing all necessary and important features into a database.

Spatial databases are specially designed that way to fulfill all the conditions necessary for placing spatial objects that are stored in the form of a table whose columns represent spatial data. Databases of this type are most often included in the group of relational databases, in order to provide support for setting spatial queries and spatial indexing data based on a declarative language such

as SQL. There is support for working with spatial data types (basic concepts are: point, line and polygon), for geometric calculations such as distance or connections between objects. Examples of spatial databases are Oracle Spatial, SQL Server and PostGIS. [3]

The advantages of spatial databases are: spatial data are viewed like any other data in the database, the database management system performs the organization and indexing which results in the user not having to repeat the implementation of operators and functions, determination of spatial connections and appliance of spatial operations are possible thanks to the support for the use of simple SQL queries. Main shortcomings of such organized databases are: expensive implementation, reduced flexibility, incompatibility with some of the existing GIS software, lower speed compared to local, existence of specific structures, as well as the lack of user experience in working with spatial databases. By a combination of good and bad properties of spatial databases, it is possible to come up with a conceptual solution to the problem of working with spatial data – future upgrades and optimizations would affect to the fact that spatial data have an increasing scope in domains that are not closely related to geoinformation systems. An example of where all this could be used is the process of developing the smart city project.

### III. RASTER MODELING AND QUERY CREATION OVER SPATIAL DATA

As spatial data are determined by metric, topological and thematic component, it is possible to describe them using two basic types of models, depending on whether their properties are primary or their position is primary: raster (surface, areal) and vector model (line, linear). Figure 1 shows a comparative view of the raster and vector spatial data model representation.

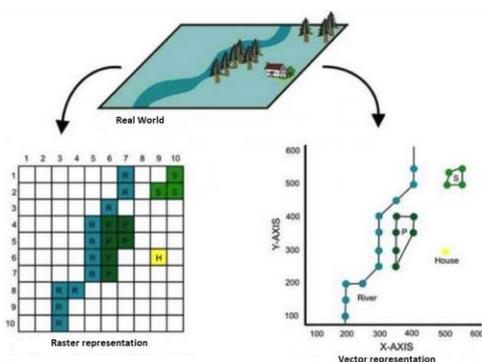


Figure 1. Comparative view of the raster and vector spatial data model representation

The base of raster data are surfaces, while the pixel is basic geometric element. Raster data are commonly displayed in the form of a matrix, where the position of each pixel is determined by row and column. Raster image format represents a representation of point color amplitude values. The basic feature of raster systems is representation of space by a number of layers, whereby the space shown on each of the layers is divided on rectangles of the same shape and size. Cells of different layers overlap, which is the main one prerequisite for the

appliance of various analyzes by combining values using mathematical and other operations. Raster data highlights features (positional accuracy is not essential), and therefore exceptional well describe objects or occurrences with vaguely defined borders.

The pixel-based data model is distinctive demanding memory usage but simplified structure of data itself, a layout of defined cells that does not follow strictly natural boundaries but generating from satellite image is simplified, less precision than vector model but data analysis is not a complicated process. [4]

Raster model of spatial data, as already stated, is described as the surface covered most often by a network of square cells of equal dimensions. Thanks to the fact that endpoint coordinates for the network are known, the dataset expressed as a 2D matrix is uniquely determined. In spatial databases the attributes are explained as the values assigned to each individual cells, which together form a single layer. There can be multiple such layers. Raster data extraction follows after applying special procedures such as remote detection, scanning of plans and maps and creation of digital model of the observed terrain. Continuous spatial entities are most closely modeled using a data model based on pixels, such as height, temperature, type of land, way of using the land, etc.

Despite the fact that raster graphics has serious problem manifested in the form of pixel mixing, it is commonly used in domains where the surface is an object of interest. In order to archive raster files a number of compression techniques are applied. Computer supported file formats for storing this type of model are: jpeg, tiff, bitmap and bil. A format that allows not only storing but also processing raster data is a grid. [4]

Querying spatial database data is a complex process since it needs to solve the problem of connecting spatial algebra operations with query language used by the database management system. Additionally, spatial data required inputs modeled by raster graphics (inputs can be entered via graphic input devices or can be calculated in the query), as well as graphic representation of the obtained results. From the standpoint of algebra, operations that must be conducted in order to prepare spatial data are spatial selection, spatial merging, appliance of spatial functions and operations that represent the relationship between the objects of spatial algebra and database management system facilities. The problem of the need for graphical presentation of values of spatial data types can be solved by using graphical user interface that enables input data transformation.

The integration of geometry into the query language implies next steps: marking values of spatial data type as constants in the query and graphical input such constants, expressing the previously mentioned classes of algebraic operations and describing the presentation of results.

Spatial indexing is supported by the spatial database management system and its goal is supporting spatial selection. This indexing method should organize space and objects so that they can be considered as parts of space and subsets of objects taken into account when

answering queries. In addition to that, spatial indices enable spatial selection, they also allow the performance of operations such as spatial merging and searching for objects close to the values from the query. Thanks to the appliance of approximative methods, spatial data can be indexed and searched. [5]

From the aspect of object-relational databases, software implementation of spatial data models consists of a set of data types and operations enforceable on them. Executing queries over spatial databases became possible because of developing an abstract data type. Approach like that was introduced in order to overcome limitations in modeling entities of interest.

Loading a raster image into a spatial database can be executed using the *raster2pgsql* command. Since it works with a relational database, it is possible to use a .sql file or graphical user environment pgAdmin. Example of using the specified software command is: *raster2pgsql raster\_options\_go\_here raster\_file someschema.sometable > out.sql*. Question language optimized for work with spatial databases allows specifying exclusive operations, defining constraints for correct binding of occurrences in raster catalogs, specifying optional parameters intended for handling input raster data set and enumeration of optional parameters for manipulating objects stored in the database.

An example of how to use a loader to create input file and its upload (with resizing to a size of 100x100 pixels) is shown in Figure 2.

```
raster2pgsql -s 4326 -I -C -M *.tif -F -t 100x100 public.demelevation > elev.sql  
psql -d gisdb -f elev.sql
```

Figure 2. Creating an input file and uploading it

The -s parameter is a reference identifier of the spatial data. If the value is not explicitly specified or is equal to zero, metadata of the raster data model will assign appropriate identifier. The view to a raster image is created using the -I parameter, which is why the generated elev.sql file will also contain the main table and view. In order to be able to apply restrictions above the spatial data, it is necessary to specify the parameter -C. Frequently used restrictions are applied to assignment of unique identifier and pixel size in order to properly register raster in the columns of the table. Vacuum raster table analysis will be performed due to explicit specifying the -M parameter. In concrete case, the input can have an arbitrary name, but must be specified in .tiff format. Adding a column with the file name can be considered as done by specifying the optional parameter -F, while -t 100x100 performs cutting the grid into square cells that will then be represented in separate rows of the table. The cell is initially set to the value „auto“ in order to the loader could calculate the appropriate cell size using the first raster and then apply that size to everything others. Important feature of PostgreSQL is that the schema name does not have to be stated. If this is the case, the table with given name will be created in the default public database schema. To avoid the repetitive part of the job, redirection can be applied, for which the operator „>“ is

used. Thereby achieving the effect that all parameters that appear in the command which is executed over the database are listed in file with the given name (in the described example it is a file under the name elev.sql). The second command in Figure 2 is for deleting the existing table and for creating a new one that will be filled with raster data (parameter -d). Additionally, a column for storing information about raster data will be created – its name is given above previously saved .sql file. [6]

Query over the data stored in the spatial database is realized by applying the appropriate commands. Thus, it is possible to select a part of a raster image by specifying the operation that seeks the intersection between the original image and its region of interest which should be segregated. During the realization of the query, non-unique results for some operations are possible, so it must be taken into account. This means that it is necessary to perform conversion from text to geometric representation of spatial data. An example of a more complex SQL query

```
SELECT rast  
FROM my_raster  
WHERE ST_Intersects(rast, 'SRID=4326;POINT(-10 10)::geometry');
```

is given in Figure 3. [7]

Figure 3. Query for selecting a part of a raster image

#### IV. POSTGRESQL

PostgreSQL is open-source object-relational database management system, reliable architecture, integrity and accuracy. It relies on the SQL query language (not only uses it, but also expands it) in combination with many functions that securely store and scale working data of a high degree of complexity. It allows the use of foreign keys, merging tables, creating views and working with stored procedures written in various programming languages. PostgreSQL has support to work with most SQL:2008 data types such as integer, numeric, boolean, char, varchar, date, interval and timestamp. It also allows storing binary large objects (so-called blobs) which include photographs, sound and video records.

The database management system has the task to provide the ability to manage data and their processing, and consists of three units: logical, conceptual and physical. The specificity of spatial databases is reflected in the fact that it adds spatial data types that represent geometric properties, and include spatial structures such as boundaries and spatial data dimension. Additionally, this type of database defines a set of function for analyzing of geometric components, for determining spatial relationships, as well as for handling geometry. Such databases can store and manipulate with spatial objects in the same way as they do with every other facility.

Three basic aspects that represent the link between spatial data and databases are: type of spatial data (point, line, polygon), multidimensional spatial indexing, efficient processing and spatial operations, as well as

spatial functions in SQL for queries about spatial properties and relationships. [1]

The evolution of PostgreSQL has affected to the fact that today it is considered a tool of solid reputation thanks to the architecture it relies on, thanks to high reliability, data integrity, huge set of features, scalability and the open-source community's dedication which stands behind this tool and allows consistent delivery of efficient and innovative solutions. It allows you to work on significant operative systems and is compliant with ACID principles. The development of this software has resulted in the construction of powerful plugins very common nowadays such as PostGIS geospatial database extension. [8]

## V. POSTGIS I QGIS

PostGIS is defined as a spatial object-relational PostgreSQL database extension. It enables work with geospatial data and provides support for geographic features since it leaves space to execute geospatial queries in SQL. Innovations that introduce PostGIS into PostgreSQL are manifested in the new data types such as geography, geometry, raster and others. It also adds functions, operators and indexes extensions corresponding to these specific data types. These extensions make PostgreSQL fast and a powerful spatial database management system.

Before the advent of PostGIS, PostgreSQL provided support for work with geometry data types, but its basic problem was that such types were too restricted for the data encountered by the geoinformation systems. Working on the construction of geometric object gets more intense over time, which resulted in the creation of a basis for the introduction of spatial index, and possibility to define the basic ones spatial data storage components. They include spatial objects and standardized text (well known text), elementary functions such as distance (Length()) and space (Area()), an extension for import/export to in Java and spatial indexes for fast data access. [5]

Today, PostGIS is considered widespread database. Number of programs that store and download data based on this PostgreSQL extension is on the rise. Numerous reasons for the increasing appliance are contained in the fact that PostGIS is supported by both, open source and closed source software, as well as server and desktop systems. Easy installation and use of tools also contribute to the increase in the number of users. An example of how you can create a general query by location for an arbitrary geographic object can be seen in Figure 4.

```
SELECT superhero.name
FROM city, superhero
WHERE ST_Contains(city.geom, superhero.geom)
AND city.name = 'Gotham';
```

Figure 4. SQL query by location for an arbitrary geographic object

The specified query should select the column called name from superhero table by considering only that torque from superhero table located in the same location as torque from table called city, with additional request to select torque which belongs to a town called Gotham.

Since version 2.0 PostGIS fully supports working with raster data sets. The command by which the loading of raster model is performed is earlier mentioned: *raster2pgsql*. This command is able to generate sql query that will load spatial data for any raster supported format in the same way on which the *shp2pgsql* command works for files which contain data about the shape of the object. If it is necessary to get additional information about the raster object, it can be invoked by calling command *gdalinfo* that will list attribute values such as driver information, file size, coordinates, pixels, color, type and

```
$ gdalinfo worldclim/tmax09.bil
Driver: EHdr/ESRI .hdr Labelled
Files: worldclim/tmax9.bil
      worldclim/tmax9.hdr
Size is 2160, 900
Coordinate System is:
GEOGCS["WGS 84",DATUM["WGS_1984",SPHEROID["WGS 84",6378137,298.257223663,
AUTHORITY["EPSG","7030"]],TOWGS84[0,0,0,0,0,0],AUTHORITY["EPSG","6326"]],
PRIMEM["Greenwich",0,AUTHORITY["EPSG","8901"]],UNIT["degree",0.0174532925199433,
AUTHORITY["EPSG","9108"]],AUTHORITY["EPSG","4326"]]
Origin = (-180.000000000000057,90.000000000000000)
Pixel Size = (0.166666666666667,-0.166666666666667)
Corner Coordinates:
Upper Left (-180.0000000, 90.0000000) (180d 0' 0.00"W, 90d0' 0.00"N)
Lower Left (-180.0000000, -60.0000000) (180d 0' 0.00"W, 60d0' 0.00"S)
Upper Right ( 180.0000000, 90.0000000) (180d 0' 0.00"E, 90d0' 0.00"N)
Lower Right ( 180.0000000, -60.0000000) (180d 0' 0.00"E, 60d0' 0.00"S)
Center ( 0.0000000, 15.0000000) ( 0d 0' 0.00"E, 15d0' 0.00"N)
Band 1 Block=2160x1 Type=Int16, ColorInterp=Undefined
Min=-153.000 Max=441.000
NoData Value=-9999
```

similarly (Figure 5). [10]

Figure 5. Additional information about loaded raster data

After loading raster image, access is checked whether records of it are really present among the relevant ones metadata. If it turns out to have nonzero values in the raster *raster\_columns* column, this is the signal that the operation is executed successfully. The described effect can be achieved by applying the simple SQL select command shown on Figure 6.

```
postgis_example = SELECT * FROM raster_columns;
```

Figure 6. An example of simple SQL select command

Preparations for the future visualization of loaded raster data involves converting the range of all rows in the table (each row of the table represents one segment raster block grid of square cells) into a file that contains information about shape of spatial data, for which the function *ogr2ogr* can be used for. Program code for the described transformation is given in Figure 7.

```
ogr2ogr temp_example.shp PG: "host=localhost port=5432
dbname='postgis_example' user='me' password='mypassword'" -sql
"SELECT rid, filename, ST_Envelope(rast) as the_geom FROM chp01.tmax01"
```

Figure 7. Preparation of a raster model for visualization

QGIS is a free open-source application that provides support for collaboration with a large number of different platforms optimized to work with geographic information systems. It is intended for desktop computers and allows you to view, edit and analyze geospatial data. QGIS is the most commonly used with PostGIS. This plugin allows displaying raster images presented in the form of a table together with generated network of cells representing image pixels. Figure 8 shows the result of the QGIS tool operation.

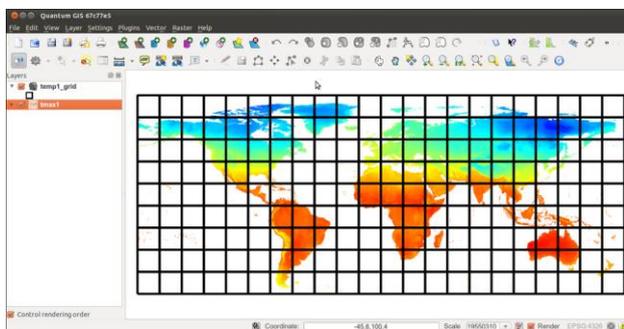


Figure 8. Visualization of a raster model of spatial data using QGIS tool

The advantage of working with QGIS is reflected in the fact that the previously described process of preparation for visualization is not complex, and it is enough just to load the forms obtained as a result of applying the transformation over the input raster image (*temp\_example.shp*). Thereafter the application will know how to interpret and then how to visualize the obtained data. [9]

## VI. CONCLUSION

The paper presents the issues the geoinformation system is faced with. This system can be described as a system specifically designed to support spatial data processing that unites the field of informatics and geodesy. In order to explain the concept of working with spatial databases, it was necessary to process the procedure of creating spatial data models as well as to explain concept itself and base concepts of object-relational databases in the general case.

Numerous studies have influenced the creation of a strong bond between information technology and geographic and geodetic aspects. The obtained result is reflected in developing spatial databases designed to enable storing, searching and visual displaying of modeled data. Vector and raster model of spatial data are mentioned, but the accent was on raster images since they dominantly figure in the domain of the smart city concept.

A system that combines hardware and software components, spatial data, human factor, scientific disciplines of informatics and geography, as well as different approaches of processing is the basis of manipulation with spatial data in the usual way, as it does with other standard data types. Because spatial databases exhibit characteristic properties, this paper briefly

explains and illustrates way of creating queries and spatial indexing procedure.

As a starting point for placing the collected raster data into database, PostgreSQL expanded with PostGIS enhancement was chosen. This geospatial extension affected into its appliance: it gradually begins to take precedence as it offers various possibilities of working with data and it contains numerous functions that use spatial data. PostGIS distinguishes between geometric and geographical type of data which becomes obvious during presentation of obtained query results. A tool that allows you to visualize the inputs given in the raster image form is QGIS which also becomes more widespread due to the simplicity of use.

Work on software development of smart city project solution will be based on the principles of spatial databases because this completely new concept successfully overcomes many limitations and shortcoming characteristic for classical databases. Problems of working with spatial data expressed by a raster model will be less and less present and thus will gradually move towards satisfying a large number of necessary prerequisites for software support for the idea of a smart city. As well, future upgrades will affect the spatial database become the conceptual basis for designing solutions which would connect information technology and many other scientific disciplines.

## REFERENCES

- [1] M. Pranjić, „Prostorni podaci i geografske baze podataka”, 2017.
- [2] P. Crickard, E. van Rees, S. Toms, „Mastering Geospatial Analysis with Python“, 2018.
- [3] Baze prostornih podataka, [Geovizualizacija i Web kartografija \(bg.ac.rs\)](http://Geovizualizacija.i.Web.kartografija.bg.ac.rs)
- [4] A. Hamzić, „Geoprostorne baze podataka“, 2018.
- [5] D. Mikiwicz, M. Mackiewicz, T. Nycz, „Mastering PostGIS“, 2017.
- [6] Raster Data Management, Queries, and Applications, [Chapter 5. Raster Data Management, Queries, and Applications \(postgis.net\)](http://Chapter.5.Raster.Data.Management.Queries.and.Applications.postgis.net)
- [7] PostGIS Raster Frequently Asked Questions, [Chapter 8. PostGIS Raster Frequently Asked Questions](http://Chapter.8.PostGIS.Raster.Frequently.Asked.Questions)
- [8] What is PostgreSQL, [PostgreSQL: About](http://PostgreSQL.About)
- [9] Importing raster data with the raster2pgsql PostGIS command, [Importing raster data with the raster2pgsql PostGIS command | PostGIS Cookbook \(packtpub.com\)](http://Importing.raster.data.with.the.raster2pgsql.PostGIS.command.PostGIS.Cookbook(packtpub.com))
- [10] L. S. Hsu, R. O. Obe, „PostGIS in Action“, 2021.

# Secure Communication in Early Fire Detection Systems

R. Radišić, D. Dinu and S. Popov

University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia  
radisic.rade@uns.ac.rs, dinud@uns.ac.rs, srdjanpopov@uns.ac.rs

**Abstract** - Current hardware quality to price ratio has enabled mass production of small, cheap electrical devices. It provided the use of computers in monitoring and automating buildings and cities, creating the Internet of Things concept. In such systems, various software and network protocols are being used because of an incoherency of used software, opening the space for malicious attacks. Out of all available types of systems, the ones with critical mission are the most affected with this problem. This paper explores the possibility of a system with a single, secure communication layer over all elements in a network on an example of an early fire detection system.

## I. INTRODUCTION

Hardware technology has advanced enough to mass produce various kinds of high processing power devices. Combined with global networking, it created a perfect environment for developing an Internet of Things (IoT) systems, which should help in everyday tasks. Apart from that, it gives another possibility to monitor the health of the system environment, such as residential buildings for example, for a set of characteristics. One set of those characteristics may be related for an early detection of a fire hazard in order to be early detected, or even better, automatically put down, thus minimizing the chances of human casualties.

Early fire hazard detection/prevention systems are relatively complex systems composed of multiple hardware and software components. Regarding hardware, there are multiple types of sensors involved, which are sending measured data through communication devices to the final destination, which could be a local server, or a cloud provider, where the results are aggregated and the calculations are done in real time. After the raw data is processed, it can be displayed to the end user through various channels, for example, sending the data using the Internet with a dedicated client mobile or desktop application, or using the SMS protocol etc.

One recognized problem with IoT systems is that acquired data most often goes through multiple different communication protocols, because different devices involved often are not supporting the same protocol, or there are some physical limitations involved in device usage, so a specific wired or wireless protocol has to be used. Security in IoT systems is a very important topic and there cannot be outside involvement (for example, tampering sensor measurements) of any kind, because it can cause the system with a critical mission to misbehave and cause damage or even human casualties. Since there are multiple hops in between acquiring a device (sensor), aggregating device (data server) and processing device

(application server), there are multiple entry points for the attacker that are exploitable. This paper will try to explain how to prevent this issue with available software technology using an example of a fire hazard detection system.

## II. OVERVIEW OF AN INTERNET OF THINGS SYSTEM

### A. Hardware

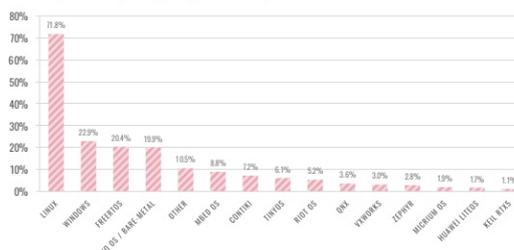
There are multiple device groups involved in an IoT system. Starting from the area of interest there are sensors, which are used to measure specific data (for example air humidity) and actuators, which are used to physically manipulate certain objects (for example, dim the lights). Usually, there is a microcontroller device that is capable of both communicating with sensors/actuators on one side using specialized, often simpler protocols, and the rest of the system on the other using TCP/IP protocol. Rest of the hardware includes network devices such as routers and switches and server computers that have much more computing power than microcontrollers and are capable of faster aggregation and deriving conclusions from aggregating data. Alternative to local server usage is to use cloud services for doing the calculations, but that can potentially introduce additional latency in the system which sometimes is not acceptable.

### B. Software

Since the IoT system can be composed of many different devices, the number of used software may also be big. There are device and communication drivers that can be a part of an operating system, which usually is the case. Custom IoT software solutions use these drivers and protocols to send/receive data to the servers, where decisions are made by doing the calculations in specialized programs. In order for systems to be able to react in real time, the Real Time Operating Systems (RTOS) are being used. Most popular RTOSes are listed below:

## IoT OPERATING SYSTEMS

Which operating system(s) do you use for your IoT devices?



Copyright (c) 2018, Eclipse Foundation, Inc. / Made available under a [Creative Commons Attribution 4.0 International License](#) (CC BY 4.0)

Figure 1. Operating systems used in IoT solutions in 2018 [4]

Figure 1 shows that Linux and Windows are the most popular choices. These are, of course, stripped down versions relative to PC operating systems so it could run on devices with less resources. The third place belongs to FreeRTOS, which is a specialized operating system for real time devices. Interestingly, in the fourth place, there is a non-OS variant and after it there are some less used RTOSes such as TinyOS etc.

### III. SECURING COMMUNICATION IN IOT SYSTEMS

#### C. Ockam Framework

Ockam is a set of libraries grouped into a framework that offers safe communication between devices in an IoT system. Messages are encrypted and to enable end-to-end encrypted secure channels that don't expose application data on network intermediaries, Ockam includes a lightweight, binary, application layer routing protocol [1]. Using this protocol makes it possible to make a single channel that is on multiple underlying network protocols, thus being independent of the security level of a sole network protocol. Ockam is available as a library dependency in two programming languages: Rust and Elixir. In the case of this paper, Rust language will be used, but the examples will be simple, so anyone with a basic understanding of programming should be able to further follow.

Ockam has a few concepts that need to be defined: a Node, a Worker, Routing and Transport between nodes using a Secure Channel. An Ockam Node is an asynchronous execution environment that can run very lightweight, concurrent, stateful actors called Ockam Workers [1].

```
#[ockam::node]
async fn main(mut ctx: Context) -> Result<()> {
    ctx.stop().await
}
```

Figure 2. Simple Ockam Node example

Figure 2. shows the simplest example of an Ockam Node definition. Declaring dependencies are removed for the clarity of an example. It is an asynchronous main function that accepts a Context as a parameter. The whole function is marked as Ockam Node using Rust's macros. This will generate all the necessary boilerplate code in order to create a new node.

As it was already mentioned, Ockam Node may contain multiple Workers, which have the following properties [1]:

- Run in an Ockam Node.
- Have an application-defined address (like a postal mail or email address).
- Can maintain an internal state.
- Can start other new workers.
- Can handle messages from other workers running on the same or a different node.
- Can send messages to other workers running on the same or a different node.

The following code is used for creating a simple Echoer Worker:

```
pub struct Echoer;

#[ockam::worker]
impl Worker for Echoer {
    type Context = Context;
    type Message = String;

    async fn handle_message(
        &mut self, ctx: &mut Context,
        msg: Routed<String>) -> Result<()> {

        println!("Address: {}, Received: {}",
            ctx.address(), msg);

        ctx.send(msg.return_route(),
            msg.body()).await
    }
}
```

Figure 3. Simple Ockam Worker example

A struct is defined for a new worker. If it contains fields, then it is a stateful Worker. New struct has to implement a Worker trait, then define an asynchronous method called `handle_message`. Inside that method is what the worker will do when it receives a message. In this case, it will print out an address from where the message is being received and the message content. After that, it will send the message body back to the sender. In order for a worker to be included in the node, the node itself has to be defined like this:

```
#[ockam::node]
async fn main(mut ctx: Context) -> Result<()> {
    ctx.start_worker("echoer", Echoer).await?;
    ctx.send("echoer",
        "Hello Ockam!".to_string()).await?;
    let reply = ctx.receive:::<String>().await?;
    println!("App Received: {}", reply);
    ctx.stop().await
}
```

Figure 4. Ockam Node with registered Worker

The node will start the Echoer worker with an address "echoer" and then send the "Hello Ockam!" message to it. The Echoer worker will then execute, printing the message and sending it back, so the node will print the same message once again. After that, the context will stop all the active workers, the node itself and do the required cleanup. It is worth mentioning that every node has an implicit worker called "app" which is responsible for executing the main function.

The communication between workers in the example below was direct, or one hop away. Ockam gives the possibility to send messages through multiple hops of the same or multiple nodes. This is obtained through Routing.

```
pub struct Hop;

#[ockam::worker]
impl Worker for Hop {
    type Context = Context;
    type Message = Any;

    async fn handle_message(
        &mut self, ctx: &mut Context,
        msg: Routed<Any>) -> Result<()> {

        println!("Address: {}, Received: {}",
            ctx.address(), msg);

        let mut message = msg
    }
}
```

```

        .into_local_message();
    let transport_message = message
        .transport_mut();

    transport_message.onward_route.step()?;
    transport_message.return_route
        .modify()
        .prepend(ctx.address());

    ctx.forward(message).await
}
}

```

Figure 5. Hop Worker example

Figure 5. is an example of a Hop worker, whose handling message will remove its address from the onward route of the received message and put it in first place of a returning route of the message. After that, the message will be forwarded to the next worker. Node definition will contain information where the message should go.

```

#[ockam::node]
async fn main(mut ctx: Context) -> Result<()> {
    ctx.start_worker("echoer", Echoer).await?;
    ctx.start_worker("h1", Hop).await?;

    ctx.send(route!["h1", "echoer"],
        "Hello Ockam!".to_string()).await?;

    let reply = ctx.receive:::<String>().await?;

    println!("App Received: {}", reply);

    ctx.stop().await
}

```

Figure 6. Node using Hop Worker example

There are now two workers registered in the node, with addresses “echoer” for an Echoer worker, and “h1” for a Hop worker. The message is sent with “route!” macro containing the array of workers. The message is sent to workers in the order that is specified in the “route!” array, so “h1” worker will receive the message first, and, as previously said, remove itself from onward\_route and prepend to return\_route. That way, once the worker with the address “echoer” receives the message and sends the answer back, it will go back through “h1” worker first. There can be an arbitrary number of in-between hops. In that case, the Hop worker can be declared multiple times with a unique address for each worker (for example, “h1”, “h2”, “h3”, etc.) and they all can be added to “route!” array in a particular order.

The provided examples were exchanging messages routed inside one node. Usually, there is a need to send messages to other nodes using different protocols. This is where Ockam Transport is used. It is a plugin for Ockam Routes. It moves Ockam Routing messages using a specific transport protocol like TCP, UDP, WebSockets, Bluetooth etc. [1].

```

#[ockam::node]
async fn main(ctx: Context) -> Result<()> {
    let tcp = TcpTransport::create(&ctx)
        .await?;

    tcp.listen("127.0.0.1:4000")
        .await?;

    ctx.start_worker("echoer", Echoer).await?;
    Ok(())
}

```

Figure 7. TCP responder node example

Figure 7. shows a responder node, which uses TCP protocol and listens on port 4000 for incoming messages. In order for “app” worker to never stop working (unless the program is terminated), instead of stopping the context as it was the case in previous examples, there is only an Ok(()) which is a value of generic type Result.

```

#[ockam::node]
async fn main(mut ctx: Context) -> Result<()> {
    let _tcp = TcpTransport::create(&ctx)
        .await?;

    let r = route![(TCP, "localhost:4000"),
        "echoer"];
    ctx.send(r, "Hello Ockam!".to_string())
        .await?;

    let reply = ctx.receive:::<String>().await?;

    println!("App Received: {}", reply);
    ctx.stop().await
}

```

Figure 8. TCP initiator node example

Figure 8. shows an initiator node that will trigger the responder node from Figure 7. The context is set to use the TCP protocol, then a message is sent to a worker with the address “echoer” through the TCP protocol route, at address “localhost:4000”, so the responder node worker “echoer” is triggered through the TCP protocol. Similarly with multiple hop workers, there can also be multiple TCP node hops. This is done by running new nodes to listen at different ports and using these ports in a route while sending a new message.

All the previous examples showed how communication is done in Ockam, but the messages still aren’t encrypted. Secure channels are used for sending the messages that are encrypted. Ockam currently supports two different key agreement protocols - one based on the Noise Protocol Framework and another based on Signal’s X3DH design [1].

Secure channels are obtained in a way that nodes get their identity and using that identity, the handshake is done between nodes. Once the handshakes are done, all other messages can be exchanged with encryption through a secure channel. The example configuration will be similar to the examples in Figures 7 and 8, with a middle hop between them.

```

#[ockam::node]
async fn main(ctx: Context) -> Result<()> {
    ctx.start_worker("echoer", Echoer).await?;
    let tcp = TcpTransport::create(&ctx).await?;
    tcp.listen("127.0.0.1:4000").await?;

    let vault = Vault::create(&ctx)?;
    let mut bob = Entity::create(&ctx, &vault)?;

    bob.create_secure_channel_listener(
        "bob_listener",
        TrustEveryonePolicy
    )?;

    Ok(())
}

```

Figure 9. TCP responder through secure channel node example

Responder node still has an “echoer” worker and listens on the port 4000, but beside that, it has an additional Entity object that has its secret contained in the

Vault object. A secure channel listener will be created for an Entity and every new connection to a node must go through the secure channel first to perform a handshake.

```
#[ockam::node]
async fn main(ctx: Context) -> Result<()> {
  let tcp = TcpTransport::create(&ctx).await?;
  tcp.listen("127.0.0.1:3000").await?;

  Ok(())
}
```

Figure 10. TCP middle node example

Middle node listens to port 3000 for incoming messages. However, since it hasn't been handshaked with any other node, it can only serve as a mediator for encrypted messages that are going to be passed through it.

```
#[ockam::node]
async fn main(mut ctx: Context) -> Result<()> {
  TcpTransport::create(&ctx).await?;
  let vault = Vault::create(&ctx)?;
  let mut alice = Entity::create(&ctx,
  &vault)?;

  let r = route![(TCP, "localhost:3000"),
                (TCP, "localhost:4000"),
                "bob_listener"
  ];

  let channel = alice.create_secure_channel(
    r,
    TrustEveryonePolicy
  )?;
  ctx.send(route![channel, "echoer"],
    "Hello Ockam!".to_string()).await?;
  let reply = ctx.receive::<String>().await?;

  println!("App Received: {}", reply);

  ctx.stop().await
}
```

Figure 11. TCP initiator using secure channel node example

The initiator node will have its own Entity object, which is going to be used to create a secure channel. This will trigger responder node's secure channel listener and the handshake will be performed between nodes. After the handshake is done, messages between nodes will be encrypted end-to-end, so any hop nodes in between wouldn't be able to see message content. This concludes the core infrastructure overview that Ockam offers for secure message exchange between nodes.

#### IV. EARLY FIRE DETECTION SYSTEM EXAMPLE

##### D. System Architecture

The example configuration is similar to the system in [2]. It contains a flammable gas detector, a smoke detector, GPS module and all that is connected to an Arduino programmable microcontroller. Beside that, it has a relay attached to the microcontroller in case there is an electrical installation that can cause fire, since fires caused due to electrical and heating equipment are the top-reasons for any residential fire [3]. The faulty electrical installation can then be quickly turned off in case of an emergency.

Residential buildings contain multiple apartments, so one of these early detectors could be installed per apartment, or even per apartment room. This mesh of microcontrollers are all connected to a computer that's inside the building, that can be a Raspberry Pi. That computer serves as a hop between the controller and the rest of the system. It can eventually decide whether to temporarily turn off power in a part of the building. The message is further sent on a cloud system that should contain the statuses of all residential buildings with the possibility for the user to see it with some kind of dashboard view. Also, alarm triggering notification with the exact information where the hazard is happening would be of use to, for example, firefighters, so they could react much more quickly.

##### E. Communication overview

The system will use Ockam as a tool for secure communication between system elements. Microcontrollers will have secure channels established with both residential computer and the system on a cloud. The connection with the residential computer will serve as a quick responding channel (for example, temporarily cutting off electricity in the room/apartment), while in the communication with the system on a cloud, the residential computer will serve as a Gateway Node, just passing the encrypted messages, without ability to see their content.

So far, three types of nodes are identified in the system, the microcontroller, the residential computer and the cloud node. Each of them will have its own workers, which will do certain tasks. For example, a microcontroller node has a worker for periodically checking every sensor state. Sensor worker returns sensor state in an answer to the default "app" worker that's previously mentioned, and if there's a need to trigger other devices through secure channels, "app" worker will be responsible for doing that. Similarly, once the execution command comes in from a higher instance, the "app" worker will delegate the message to the corresponding worker. Residential computer node has a worker that checks if there are any potential fire hazards, and if there is one, acts temporarily by cutting off the electricity in the area of hazard. Cloud system node worker notifies the closest firefighters crew to the hazard area with the collected hazard information and waits for the confirmation. If there is no confirmation in a certain time or the firefighters crew is currently unavailable, the next best possible fire fighters crew is notified.

#### V. CONCLUSION

IoT systems are being actively developed and are more and more present in everyday life. Smart cities are the aspirations whom the researchers and industry are trying to get to more than ever. Early detection/prevention of hazardous events has a very important place in IoT research. Existing systems are upgraded to be more reliable, robust and secure. Security is a very important aspect of every IoT system, since an easily breachable IoT system can do more damage than not using an IoT system at all.

The problem that rises with implementation of IoT systems lies in relying on multiple protocols that are being used in communication. If those are poorly configured, or not implemented properly, it can be an entry point for malicious attacks that could disrupt the system behavior. Ockam is a set of open source libraries whose mission is to provide communication through secure channels over multiple protocols and networks that are being used. Its main philosophy is to put zero trust in network communication and to encrypt the messages that are being exchanged through secure channels. With more sensors and devices watching us and reporting data to the Internet, the privacy of third parties who cross our sensors' paths (either by accident or design) is an important consideration [5].

In this paper, a variant of the early fire detection system has been described. This is a form of a system with critical mission whose misbehaving can lead to fatal consequences including huge material damage and, more importantly, the cost of human lives. It is imperative that these kinds of systems securely communicate, so it is a good use case for Ockam. In Ockam, communicating computers are represented as nodes, which contain workers that can be triggered by sending a message, thus immediately reacting to the events in the system. Using this kind of communication, there is a broad spectrum of

responses, including automated turning off the cause of a fire, or notification to the firefighters that will help benefit first responders reaction time to be drastically smaller. Next step for this system is further and broader design with concrete hardware components and an early prototype that will confirm or deny the starting assumptions of a stable, robust and secure system.

#### REFERENCES

- [1] Official Ockam website, [Online], <https://www.ockam.io/>, last access 18th September 2021
- [2] L. P. Maguluri, T. Srinivasarao, M. Syamala, R. Ragupathy, Nalini N.J "Efficient smart emergency response system for fire hazards using IoT", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 1, 2018
- [3] R. Sinha, Analytics Insights official website article "How IoT can be used for preventing fires", [Online], <https://www.analyticsinsight.net/how-iot-can-be-used-for-preventing-fires/>, last access 19th September 2021
- [4] Devopedia article "IoT Operating Systems" , [Online], <https://devopedia.org/iot-operating-systems>, last access 18th September 2021
- [5] A. McEven, H. Cassimally, "Designing the internet of things", Wiley, 2014

# Fog Computing architecture for IoT Smart traffic applications

M. Mazalica\*, D. Ivetic\*\*

\* University of Novi Sad / Technical Faculty "Mihajlo Pupin" Zrenjanin, Zrenjanin, Serbia

\*\* University of Novi Sad/ Faculty of Technical Science Novi Sad, Novi Sad, Serbia

milica.mazalica@tfzr.rs, ivetic@uns.ac.rs

**Abstract** – Nowadays Internet of Things (IoT) represents a technological revolution able to change the world by connecting physical and virtual objects to the Internet. This technology can offer many advantages, but it still has some limitations such as overused local networks, uncontrolled data loss, etc. All these limitations can be overcome by using emerging IoT architecture which are consisting of sensor nodes, edge layer (devices), fog layer (gateways) and cloud layer. This paper has proposed a model for smart traffic applications which is based on fog architecture. This fog-based model architecture helps reduce the network load by processing data before transmission and helps in reducing the energy consumption of sensor nodes by reproof computation from sensor nodes to fog layer (gateways). The most important advantage of this model is that only processing data are transmitted to the cloud, so data loss is avoiding.

## I. INTRODUCTION

Internet of Things (IoT) is a global infrastructure that allows people and things (e.g. sensors, actuators and smart devices) to be connected anytime and anywhere. [1] The virtual connection of IoT devices is based on interoperable information and communication technologies. The term "Internet of Things" was suggested by Kevin Ashton in 1999. but this term first becomes popular through the MIT-related Auto-ID Center at MIT. However, today the term "Internet of Things" is used to determine advanced connections of devices, systems and services, that reach out from machine to machine-to-machine and overs various protocols, domains and applications. Nowadays The Internet of Things developed very quickly and these IoT technologies enable the connection of a large number of users, devices, services and applications on the Internet.

The IoT platform is a software and hardware environment for the efficient development of IoT systems. This platform is complex for management because it includes several processes such as resource discovery, resource tracking and platform reconfiguration. During the development of IoT platform is necessary to consider the necessary functionalities elements such as devices, communication components, services, management of IoT platform functionalities, security and application. For development, IoT is used next technologies: network technologies and protocols, sensory networks, mobile technologies, cloud computing and big data.

IoT application is the layer of IoT architecture where users can use processed data from the data warehouse on the cloud for one purpose. This is a spot where all the value of IoT materializes and gets the most practical possible application. IoT applications can be web and mobile, and they provide the user interface for interacting and using data from IoT devices, but the same IoT devices can have their user interfaces. For developing IoT applications popular languages such as Java, Swift and Node Js are used.

IoT is applied in various areas such as agriculture, industry, medical and healthcare, transportation, building and home automation, industrial, energy management, environmental monitoring, living lab, military and others. The IoT application can be divided into personal and business purposes, urban environment, education, health, logistics, industry, environment, energy systems, retail, agriculture, robotics. One such category of application IoT is the Intelligent transport systems (ITS). ITS is an advanced application that aims is to prove innovative services relating to different modes of transport and traffic management. In these systems, the vehicle has sensing power, storage capabilities, onboard computing facilities and communication systems. The idea behind ITS was to solve the issues raised by the legacy systems efficiently.

This paper gives a short review of IoT architecture based on Fog computing architecture for IoT smart traffic applications. In this research, the focus is on analyses of applicability of this architecture in IoT smart traffic. This paper is structured as follows: in the introduction, a section is given a short description of IoT and IoT applications. The next section gives different architectural approaches such as Cloud, Fog and Edge computing. In section 3 is given the most important technologies for enabling IoT applications. Section 4 gives examples of the application of Fog IoT in smart traffic. The discussion and conclusion remarks are given at the end of this paper.

## II. OVERVIEW OF IoT ARCHITECTURE

The presence of IoT demands even more careful consideration of orchestrations due to the latency sensitivity of many IoT applications (or services) and limitations specific to IoT. [4] IoT applications require Edge and Fog computing for reducing overall delay. In consequence, fog and edge layers not only minimize latency data, but also bandwidth consumption of IoT applications. [4]

In paper [4] is given a generic model for the cloud to things landscape, which is based on research community efforts and industrial advances. According to [4] fully-implemented landscape for IoT applications would consist of four layers: IoT Mist computing layer, Edge computing layer, Fog computing layer and Cloud computing layer. Mist computing layer or extreme edge is formed by IoT devices that are computation-enabled and can be coopered with each other similar to a mesh network. This increases the capabilities of IoT devices to constitute a mesh network that can be managed centrally by a nominated controller or distributed. This layer can only be realized without any communication with fog and edge or cloud computing layer. The edge computing layer enables gateways such as dedicated routers, switches or based stations located close to IoT devices to act as the gateway to the fog or cloud layer. This layer aims to minimize and manage data traffic. Edge can also be equipped with micro data centers (or so-called cloudlets) which can collect and filter the sensed data of IoT and send only the reduced analyzed data to the fog or cloud computing layer. In this way, minimization is achieved. The fog computing layer is an intermediary layer to alleviated edge and cloud limitation. This layer is benefiting from computations closer to the IoT layer than the cloud layer and is more powerful than the edge layer.

The paper [2] more describes concepts, architecture, standards. Tools and application of this layer, as well as new research aspects about IoT-Fog-Cloud ecosystems. The Cloud computing layer is acting as the backbone and providing persistence data storage and powerful and unlimited computation resources which are not available in others computing layers. This layer can be categorized as private, public, hybrid, single-provider, multi-provider and federated cloud.

In papers [5] and [6] is given a more detailed description of Fog computing and its related computing paradigms. The paper [3] investigates the role and applicability of both IoT and AI technologies in Edge computing based on Arduino. The result shows that Arduino and its clone boards can be used as an edge mode with AI codes for the Arduino Edge node.

### III. ENABLING TECHNOLOGIES FOR FOG IOT APPLICATION

For efficient and cost-effective deployment of Fog-based IoT systems with distributed architecture the standards play an important role. It is necessary to consider protocols and technologies to support devices that have limited bandwidth and energy. Technologies and protocols which can be used for efficient communication of IoT devices in fog-based IoT architecture include Radio Frequency Identification (RFID), Wireless Identification and Sensing Platform (WISP), WSN (Wireless Sensor Networks), Bluetooth Low Energy (BLE), Near Field Communication (NFC), IEE 802.15.4, IEE 802.11 ah, Z-Wave, Long Term Evolution-Advanced (LTE-A), LoRaWAN, IPv6, IPv6 over Low Power Wireless Personal Area Networks (6LoWPAN), NarrowBand IoT (NB-IoT) and SigFox.

The RFID systems operate on the frequency band of 125 kHz and require 12V power support. These systems are made up of RFID tags and readers. The RFID tags use technology to reflect radio waves and pass on the data to the readers. On the other hand, RFID readers can read and extract the stored information inside the RFID tags. This system has the ability to pick up tag IDs automatically from a distance without considering the Line of Sight (LoS) operations. [1]

WISP (WSPI) is a battery-free and wireless platform used for sensing and computation. These devices are powered by ultra-high frequency RFID readers. WSIP-based systems use the same communication technology as RFID-based systems, but they are unique with a fully programmable microcontroller.[1]

WSNs systems include small nodes with sensing capabilities, which can be easily deployed into IoT infrastructure with no (or little) modifications since IoT supports interoperability of various networks including WSNs. These systems communicate in a Peer-to-Peer manner. The WSNs sink nodes can be utilized to collect sensed data from other nodes in the network.

BLE is a wireless technology for short-range communication that operates on the 2.4GHz frequency band. This technology can be utilized in various IoT scenarios such as medical monitoring, public transportation systems and monitoring industrial environments.

NFC has very short-range communication and operates on a frequency band of 13.56 MHz, and this standard enables devices to communicate to each other only in close vicinity. This standard can be used in social networks, museums and mobile ticketing systems.

Z-Wave is a low-power MAC protocol that operates on the frequency of 908 MHz and is utilized by small data packets with the range of 30m at low speeds up to 100kbps.

LTA-A is an enhanced version of LTE which provides higher throughput and lower latencies as well as improved coverage. [1] This technology supports higher bandwidth up to 100 MHz with aim of obtaining a higher level of system performance. LTE-A is suitable for fog-based IoT infrastructures because fog-based devices may be used to offer relay services to end-devices or other fog nodes in the network.

LoRaWAN is a Low Power Wide Area (LPWA) technology that supports low power and low data rate. In fog-based IoT, this technology can be used by the end-devices for communication with gateways using a single hop.

NB-IoT is a low-power cellular technology specifically designed for IoT to improve coverage with respect to LTE. This technology allows connection between different objects with a small amount of data over a long period. The integration of NB-IoT and fog computing can save network bandwidth, ensure the quality of data analysis, improve the response time, and enhance the efficiency of data storage compared to traditional cloud computing models. [1]

SigFox is a network protocol that provides an M2M WAN communication solution that operates on the 868MHz frequency band. In IoT applications, this protocol is used to enhance the network capacity, increase the life cycle of the devices, reduce the cost of devices and improve communication range and minimize energy consumption. Because of the capability of fog-based devices to perform tasks closer to the network edges, SigFox will have better potential in fog-based IoT systems.

IPv6 is an Internet protocol that supports Internet Protocol Security (IPSec) and has the ability to provide a unique address to a large number of IoT devices. It also offers support for neighbor discovery which enables neighboring nodes to communicate and determine the presence of each other. [1] All these properties make this protocol suitable for fog-based IoT systems.

For improving interoperability technologies, topologies and applications in fog-based IoT systems, IEEE Std 1905.1a and IEE 1451 are two standards that can be used. IEEE Std 1905.1a is a standard that supports a common interface, by defining an abstraction layer, in order to deploy multiple networking technologies at smart homes. An important characteristic of the IEEE 1451 standard is that communicates on the Journal Pre-proof Internet in the same way for all sensors and actuators regardless of the type of the network which can be either wired or wireless. [1]

#### IV. EXAMPLES OF FOG IOT IN SMART TRAFFIC

A smart light traffic system is a network of connected traffic lights which is used to minimize traffic congestion, prevent an accident, reduce noise and fuel consumption. One example of using this system is given in paper [1] and there is described their application in health monitoring systems. Street cameras that sense the flashing light of an ambulance can change the street-light for the ambulance to pass through the traffic. It means that street lights communicate with sensors and detect the presence of vehicles and pedestrians and regulate the light accordingly. According to [1] fog devices can coordinate to provide green traffic waves and send a warning signal to vehicles approaching the traffic. In ITS fog-based devices which are placed at a certain intersection can be used to analyzed local data and inform people of the updated information about the routes.

In paper [7] is presented research of the Fog computing approach for improving the performance of the Cloud computing technologies. This approach is implemented to a level of a prototype of a smart traffic monitoring systems (STMS). The purpose of this system is for congestion monitoring traffic and traffic lights management. In this research, a computer-on-module (Intel Edison) serves as a Fog node to collect real-time data from geographically distributed sensors and to transfer it to the Cloud platform (ThingSpeak) for storage and processing. This integrated platform in [7] is interfaced with Tweeter for sending alters in case of high traffic congestion.

The paper [8] proposes a novel traffic control architecture that is based on the Fog computing paradigm

and reinforcement learning technologies. This study has the aim to generate traffic control flow and communication flow for each intersection to avoid a traffic jam. The Fog computing paradigm and distributed reinforcement learning algorithm are designed in [8] to overcome communication bandwidth limitation and optimal traffic control flow with the purpose to make the whole city's traffic highly efficient.

In paper [9] is presented a decentralized application of Fog nodes towards smart-cities traffic monitoring and forecasting. The architecture shown in paper [9] combines a data distribution layer, connecting the Fog nodes with a centralized Cloud focusing on resilience and near real-time communication and an online Machine Learning modeling technique. This application is designed to be used in a Fog-based infrastructure wherein real-time computing capabilities are proved by network antennas in combination with Fog nodes. This approach was tested in several experiments and the result shows that the data distribution process running in the Fog nodes is flexible to back-haul connectivity and is able to deliver data to the Cloud even in the presence of severe connectivity problems.

Edge computing is used in smart traffic systems for time optimization smart navigation of vehicles optimal traffic load balancing in real-time. In paper [10] Edge-Cloud-centric IoT based smart traffic management systems is developed for traffic inflow prediction and time-optimized smart navigation of the vehicles. This approach allows optimal traffic load and improved road safety.

#### V. MODEL FOR TRAFFIC MONITORING IN SMART CITY

In this section is the proposed model for traffic monitoring in the city of Zrenjanin. The model is proposed based on review and the analysis of the available literature and existing solution and modeled on paper [7] and [8]. This solution represents the application of IoT architecture in combination with Fog computing in smart traffic monitoring. This proposed model monitors the traffic load by the number of vehicles on the main road in Zrenjanin. In this study, the main roads are divided into three sections: Veljko Vlahovic street, Cara Dusana Street and Zitni trg. The first two sections contain two and the third section contains three crossroads. Each crossroads has three magnetic sensors in every traffic line for monitoring the number, speed and dimensions of vehicles.

##### A. Model architecture

The architecture of this model is given in Figure 1. This model has three layers: Edge, Fog and Cloud. The Edge layer contains the magnetic sensors which are placed near the busiest streets for collecting the data. The collected data are sent to the Fog layer via a Wireless network (Wi-Fi). These data are processed and stored on the Local Server and Local Database. The display (positioned in the street for informing the drivers and pedestrians) in this layer shows the number of vehicles, and statistical data presenting covering analytical reports on average speed and size of vehicles in defined longer

intervals (e.g. every 5 minutes). The analyses data from the Fog layer are sent via LPWAN or another long-range technology is defined longer intervals (e.g. each hour) to the main server, centralized storage and application server which are on the Cloud layer.

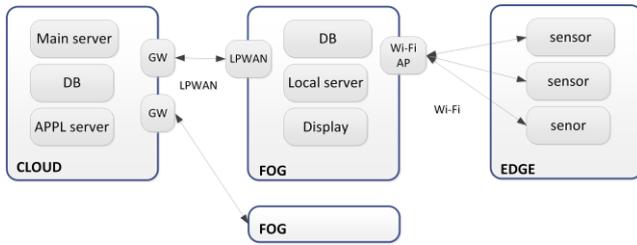


Figure 1. Model architecture

**B. Model scenario**

The model scenario is given in Figure 2. The three sections: Veljko Vlahovic street, Cara Dusana street and Zitni trg, are shown in the Zrenjanin city plan. As we said before, the first two sections contain two and the third section contains three crossroads. Each crossroads has three magnetic sensors in every traffic line for monitoring the number, speed and dimensions of vehicles. This proposed model monitors the traffic load by the number of vehicles on the main road in Zrenjanin.

The proposed scenario with the system element deployment for this model is given in Figure 2. The numbers 1, 2 and 3 in Figure 2. marks three sections: Veljko Vlahovic street, Cara Dusana street and Zitni trg. The blue dots represent locations of crossroads and where the sensors are deployed. The numbers 4 and 5 on this figure present the Fog and Cloud layer with their server, database and application server. All these three layers (Edge, Fog and Cloud) are connected via an LPWAN network or some other long-range technology.

The position of the magnetic sensors on the crossroad is given in Figure 3. Each traffic line in the incoming direction to the crossroad (right side) has three sensors.

Locations marked with the number 3 and 4 in Figure 3. present traffic lines which each has three sensors for traffic monitoring. While locations marked with number 1 and 2 has six sensors because there have two traffic lanes in one direction.

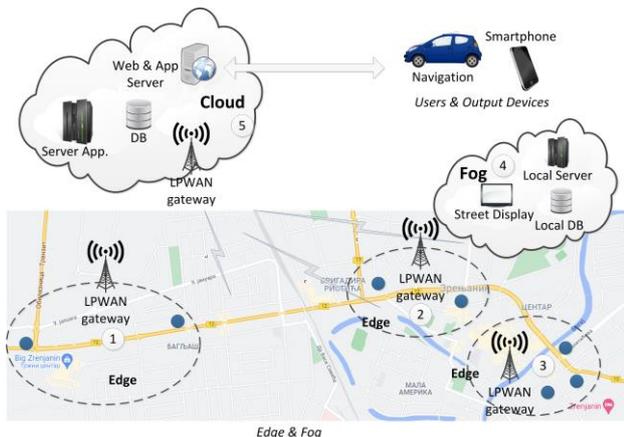


Figure 2. The position of the sensors on the crossroads

The edge sensor devices are connected to the fog layer with Wi-Fi or some other short-range technology such as ZigBee, BLE, etc. The Wi-Fi access point collects all data and stores it in a local database. The collected data are analyzed and displayed in short intervals for a quick review on-street displays located near crossroads. Their purpose is to inform citizens, both drivers and pedestrians about the traffic intensity level.

The LPWAN end-node sends the collected and summarized data in longer intervals (e.g. one hour) to the core of the system – the Cloud layer. In the Cloud layer collected data sent from all three sections are concentrated. The data are stored in the central database, analyzed, and provided to the end-user via the web, PC, tablet and smartphone applications.

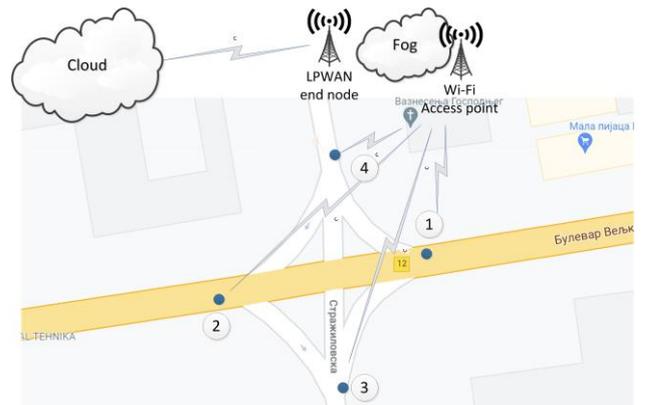


Figure 3. Model scenario

**VI. CONCLUSION**

In this paper is proposed the model of the system architecture with integrated LoRaWAN technology and Fog computing as a solution for reducing network load and distributing computational capabilities. The proposed model presents a model for traffic monitoring in the city of Zrenjanin.

In this model sensor collected data and sent them via Wi-Fi network to the Fog layer where the processing data took place. These data are sent to the Cloud layer with LPWAN technology to the centralized storage. Sending only processed important data to the Cloud has many benefits such as energy consumption of sensors, reduction in the network load, and the most important avoiding data loss.

The future works will include a traffic density estimation algorithm, also the incorporation of more types of sensors in this model (e.g. temperature, humidity, pollution, loudness, etc.).

**REFERENCES**

[1] H.Zahmatkesh, F. Al-Turjman, "Fog Computing for Sustainable Smart Cities in the IoT Era: Caching Techniques and Enabling Technologies - An Overview", Sustainable Cities and Society (2020), Vol 59, 2020, <https://doi.org/10.1016/j.scs.2020.102139>.  
 [2] A. A. Alli, M. M. Alam, "The fog cloud of things: A survey on concepts, architecture, standards, tools, and applications", Internet of Things (2020), [doi:https://doi.org/10.1016/j.iot.2020.100177](https://doi.org/10.1016/j.iot.2020.100177)

- [3] Smartcities, Open-Source Hardware Performance in Edge Computing, <https://smartcities.ieee.org/newsletter/july-2021/open-source-hardware-performance-in-edge-computing>, Accessed August 2021
- [4] M.S. Aslanpour, S. S. Gill, A. N. Toosi , "Performance Evaluation Metrics for Cloud, Fog and Edge Computing: A Review, Taxonomy, Benchmarks and Standards for Future Research", *Internet of Things* (2020), 2020, doi:<https://doi.org/10.1016/j.iot.2020.100273>
- [5] H. A. Kumar, R. J. R. Shetty, S. Roy, D. Sitaram, "Comparison Of IoT Architectures Using A Smart City Benchmark", Third International Conference on Computing and Network Communications (CoCoNet'19), 2019, Trivandrum, India.
- [6] A. Yousefpour, C. Fung, T. Nguyen, K. Kadiyala, F. Jalali, A. Niakanlahiji, J. Kong, J. P. Jue, "All One Needs to Know about Fog Computing and Related Edge Computing Paradigms", *Journal of Systems Architecture* (2019), 2019, doi:<https://doi.org/10.1016/j.sysarc.2019.02.009>
- [7] Tuan Nguyen Gia, Jorge Peña Queralta, Tomi Westerlund, 16 - Exploiting LoRa, edge, and fog computing for traffic monitoring in smart cities, Editor(s): Bharat S. Chaudhari, Marco Zennaro, LPWAN Technologies for IoT and M2M Applications, Academic Press, 2020, pp 347-371, <https://doi.org/10.1016/B978-0-12-818880-4.00017-X>.
- [8] Dobrilović, D., Brtka, V., Jotanović, G. et al. The urban traffic noise monitoring system based on LoRaWAN technology. *Wireless Netw* (2021). <https://doi.org/10.1007/s11276-021-02586-2>
- [9] J. L/ Perez, A. Gutierrez-Torre, J. L. Berral, D. Carrera, "A resilient and distributed near real-time traffic forecasting application for Fog computing environments", *Future Generation Computer Systems* (2018), doi: <https://doi.org/10.1016/j.future.2018.05.013>
- [10] Q. Wu, J. Shen, B. Yong et al., Smart fog based workflow for traffic control networks, *Future Generation Computer Systems* (2019), <https://doi.org/10.1016/j.future.2019.02.058>.
- [11] P. Maiti, H. K. Apat, B. Sahoo, A. K. Turuk, An Effective Approach of Latency-aware Fog Smart Gateways Deployment for IoT Services, *Internet of Things* (2019), doi: <https://doi.org/10.1016/j.iot.2019.100091>.
- [12] S. K. Sood, Sahil, Smart Vehicular Traffic Management: An Edge Cloud centric IoT based Framework, *Internet of Things* (2019), doi: <https://doi.org/10.1016/j.iot.2019.100140>

# Choosing the best Python web framework for beginner according to experienced users

Dejan Viduka\*, Boris Ličina\* and Luka Ilić\*

\* Faculty for Applied Management, Economics and Finance, University Business Academy, Belgrade, Serbia.  
dejan@viduka.info, boris.licina@mef.edu.rs and luka.ilic@mef.edu.rs

**Abstract** - The world of programming and the range of programming languages used is very large, and the dynamics of developing new solutions further complicate the selection of the right solution for their needs and knowledge development. The aim of this paper is to help beginners to choose the appropriate framework for learning and further development with the help of experienced users. With this choice, the users experience is very important - the framework that the user chooses can be used later in further development. This selected framework helps the user to dedicate himself exclusively to the solution and to get the most out of it and himself, rather than wandering between different solutions returning each time to the beginning of mastering a new tool. The paper presents all eight frameworks that have been presented and, based on research, offers a recommendation of experienced users that the solution chooses for beginners, while allowing an easier start and advanced web development.

## I. INTRODUCTION

Nowadays, there are many tools used to develop web content. One of the tools that have found its application in the world of programming is the Python programming language [1]. Its application is large and is used from artificial intelligence to web development [2]. Many different libraries are used for the web, allowing users to develop from simple to very complex web solutions [3]. Such libraries are called frameworks that facilitate work for the specific purpose for which they were created [4]. Software frameworks are a set of standardized libraries [5] and programming tools that help create software applications.

The Python programming language itself has many such frameworks that allow users to work [6]. This way of developing various frameworks in a programming language is frequent, so it can be seen in the programming language PHP. In addition to so many possibilities, the user is faced with a big question, and that is the choice of the right framework for the needs of development, but also for the level of knowledge of the user who should use during development [7]. In this paper, we describe several frames that have been selected according to the user's choice. The aim of this paper is to show the attitudes of users and thus make it easier for users in practice to more easily and quickly choose one of the frameworks to use in their development.

Finding the frame quickly and easily facilitates the user and learning, and thus greatly affects the quality and speed of the final product. This is very important when it

is known that a large number of frames are in circulation and that new ones appear every day that offer with them novelties that need to be mastered. Constant learning is something that is expected from a programmer, but it should also be reduced to an optimal measure. In order for a programmer to be good at his job, it is necessary to dedicate himself to a solution and master all the finesse that makes him a good programmer. If a programmer is constantly learning new technologies, then he has less time to dedicate himself more seriously to one of them and to get the most out of it. New programming languages and frameworks always bring some novelties that were once, but not always, better, so there are a large number of old programming languages that are still used today because they have quality solutions.

## II. PYTHON WEB FRAMEWORKS

Python programming language is very modern in this respect and it develops a large number of frameworks that can be used for various purposes [8]. In this paper, we deal with web frameworks that compensate for the "flaws" of Python itself when it comes to web programming. In essence, we divide them into frames with which we get everything we need (full-stack) and those with which we can develop many necessary functionalities (micro-frame) [9] and thus do what we needed that frame for. This trend is very common in programming languages, with JavaScript and PHP leading the way, but Python is also present [10].

Interestingly, these two mentioned programming languages are mainly based on web development, while Python is generally usable in other areas and the framework allows it to be applied on the web [11]. The following are web frames created from the Python programming language.

### A. Django

The Django framework was created in 2003 and is based on the Python programming language. Its application is in web development where it is successfully used by a large number of users connected in a community that provides customer support. Django supports both the older 2.x and newer 3.x versions of the Python programming language. To use this framework, you need a server environment as well as some of the supported databases, which are: PostgreSQL, MySQL, Oracle and SQLite. The user who decides to use Django in

his work has all the necessary tools for the development of a modern web project, regardless of its complexity in development. Its packages also include an administration panel that makes it easier for the user to work with the content. One of the great advantages of this framework is its support by the large community and well-done documentation. This documentation and support in the work enables an easier start as well as advanced development of modern projects. Of course, in addition to many advantages, Django also has its drawbacks, and one of them is that it is monolithic, which makes it poorly adaptable and flexible for newer user requirements. Despite this drawback, this framework is very suitable for the development of standard applications.

#### B. Pyramid

Pyramid is a web application framework developed as part of the Pylons project. It is an open source framework published under a BSD license. The development of this framework was inspired by libraries such as: Zope or Django. One of the advantages, and on the other hand for some it can be a disadvantage, is that this framework comes without predefined libraries. This means that this framework has more flexibility when it comes to adding new features by users. This framework, like Django, can work on both the old 2.x and the new 3.x version of the Python programming language. In addition to the many similarities, Django and Pyramid differ greatly in the possibility that users themselves develop new functionalities that the authors of the framework themselves did not anticipate. In addition to this, there is a big difference that there is no integrated support for databases. This just means that the user has to integrate it, but of course it is possible. This framework is very open which often increases development costs due to the need to independently develop the necessary tools. This makes this framework suitable for very experienced developers, but on the other hand it makes it very difficult for beginners to work with.

#### C. Bottle

Bottle is a fast, simple and lightweight micro web framework created using the Python programming language. It is distributed as a single file and has no dependencies other than the Python standard library, but not many features like other Python frames. One of its main advantages is the process of distributing a single file, which facilitates the sharing and distribution of applications. To start programming, all you need to do is download bottle.py and place it in the project folder. Its design is flexible, easy to use and facilitates the development of a multitude of simple web applications. However, due to its single-file distribution pattern, the framework is best suited for smaller applications, or prototypes, rather than larger projects.

#### D. CherryPy

This framework appeared in mid-2002 making it one of the oldest frameworks for the Python programming language. Despite this fact, it is less known among users than, say, Django and Flask. This frame is also a micro frame that is flexible and expandable by the user. It has

built-in tools such as sessions, authorization, caching, routing, and database support. This framework allows developers to design and implement a project similar to or similar to other object-oriented Python projects. This fact allows for smaller source code, which results in rapid development but also optimal performance. One of its dominant advantages is the ease of configuration. This framework also supports the old 2.x and new 3.x versions of the Python programming language.

#### E. Tornado

Tornado is one of the Python web frames that focuses more on the network (asynchronous network library) and speed, including all the features of the web frame. Not only does it come with performance-enhancing features, but it also has features like stenciling, routing to build a simple to scalable web application. By using non-blocking network I/O, the Tornado can scale to thousands of open connections, making it ideal for long-term testing and applications that require a long-term connection with each user. The tornado runs on all Unix platforms as well as Linux and BSD.

#### F. web2pi

This web framework was created in 2007, and its real purpose was to serve as a teaching tool in working with students. This means that it was developed from the beginning to be easy to use. Its creators were inspired by the Django frame but also by Ruby on Rails. This web framework contains a web server that enables SSL and streaming, work with relational databases as well as an integrated development environment based on the Internet as well as a batch management interface. Support in working with this framework is reflected in the extensive documentation that the user is gradually introducing into the issue. In addition to all its advantages, the community of this web framework is significantly smaller than Django or Pyramid-e, which in practice means that less support can be expected from experienced users.

#### G. Zope

This Python framework originated as an open source social project that is object oriented. Its full name is "Z Object Publishing Environment" or Zope for short. This framework is known for making Python known around the world. Zope has a developed community that in the last few years has developed additional frameworks designed for different uses and using different development principles, while retaining the old philosophy. Some of these frameworks that have been developed are: Plone, BlueBream and Grok. Zope is a leading Open Source Application Server and Content Management Framework, specializing in content management solutions, portal content management and custom applications. Zope is managed by the global Zope community, with thousands of developers and companies around the world. Zope enables teams to collaborate in creating and managing dynamic Internet-based business applications, such as intranets and portals. The Zope web application server and its Content Management Framework form the basis for one of the most popular and powerful content management systems in the world - Plone.

### H. Flask

This framework is a micro web application created in the Python programming language. It is published under the BSD license. Some of the well-known applications created using this framework are LinkedIn and Pinterest. This micro web application comes with much basic functionalities that can be easily extended. The motto of this framework is "one drop at a time", and this is achieved through extensive documentation that gradually introduces the user to the issue. This is one of the virtues that beginners opt for this framework at the very beginning, but they also often remain faithful to it because of its capabilities in working with very complex applications. In addition to its advantages, this framework also has disadvantages that are reflected in a large number of add-ons that are often not updated. In such cases, users have to spend a lot of time looking for replacements that have similar functions and are still active and supported. It is almost always difficult to find documentation or instructions for use for replacement files, so in such cases the help of a large community is necessary.

### III. CHOOSING THE BEST PYTHON WEB FRAMEWORK

The choice of the appropriate framework depends on the domain for which it is planned to be used. Practically, this means that there is no perfect solution, but it is a matter of user choice that depends on the purpose for which it will be used. The choice of users can depend on several different factors such as, ease of use, support and even ease of setting up the user environment.

Some of these frames are also divided into two main categories, full-stack frames and micro-frames (sometimes called non-full-stack). By choosing a full-stack framework, you will have at hand a variety of tools and packages for developing any type of application. These are MVC architecture, ORM, routers, security and more.

In this work, we learned from users who use Python what they prefer in various web development frameworks. We included in the research eight frameworks that our respondents used in their work. We asked users three piranhas regarding ease of use, quality of support, and ease of setting up the user environment from the frameworks they used or prefer to use. For these three questions, users answered us using ratings that ranged from 1 to 5, which we tried to make it easier for them to rate. After these questions, we calculated the average grade and presented it on a scale from 1 to 5.

The obtained results are shown in Table 1, in which the individual ratings can be seen as well as the final ocean for each frame according to the user's opinion. The table also shows the number of users who rated each box. Each user could rate only one framework he used and rate it. 222 developers participated in this research, which used some of the given frameworks in their work.

From the presented results, the table shows the percentage of users who voted for one of the frames. It is clear that the largest number of users used Django, followed by Flask. It is also interesting that Flask is the first in the received grades, followed by Django and the others.

TABLE I. PRESENTATION OF RESEARCH RESULTS AMONG EXPERIENCED DEVELOPERS

id	Frameworks	Usability	Support	Adjustment	Participants	Mark
1.	Django	4.2	4.1	4.0	121	4.1
2.	Tornado	3.9	3.7	3.6	20	3.7
3.	Bottle	4.1	3.8	4.0	12	4.0
4.	web2py	4.2	4.1	3.4	12	3.9
5.	CherriPy	3.9	3.2	3.5	12	3.5
6.	Pyramid	3.6	3.7	3.4	13	3.6
7.	Zope	3.5	3.4	3.9	4	3.6
8.	Flask	4.6	4.2	4.3	28	4.4

### IV. DISCUSSION

There are a large number of frames on the market that serve a similar purpose, but each is different. These differences are especially visible in beginners who are usually very confused and find it difficult to decide on a solution or experiment with everyone. Experienced users have already made a choice of frames that they use more often, which can be seen from our research. All frameworks listed here are used, but each has its own purpose, as well as the user to whom it suits, depending on the knowledge, experience and project for which they use it. Our research confirms what is known among experienced users and can be read on expert forums dedicated to web development and Python. The table clearly shows the number of participants who in large numbers opted for the two leading frameworks, namely Django and Flask. Other frames from our table also have good grades, but they are still used by a smaller number of users, so they recommend them to others, especially beginners.

According to the ratings, two frames stood out, namely: Flaski, Django and Bottle. These ratings were obtained based on the three parameters we used, namely: support, usability, and tuning. These parameters are very important for every user, from professionals to beginners who first encounter these frameworks for the realization of their projects.

What is characteristic of all the frameworks from our research is that they were created using the Python programming language and therefore require users to know that programming language. Another similarity is that they are all intended for web development and generally differ in their purpose and the size of the community that supports them. Community size is very important in open source projects [12] as well as documentation that helps to easily master each individual programming language or in this case the framework. Three separate frames from our research cover almost all segments that we have listed, so it is no wonder that they are used so much and of course recommended to users from the very beginning to more advanced development. This problem helps users to decide from the start on a

framework that can later serve them for advanced development, depending on their knowledge and needs.

## V. CONCLUSION

There are many frameworks on the market that meet most of the needs of our users, and we usually leave the choice to the users themselves. One way of choosing is experiential, so in this research we asked professionals to recommend a framework according to their opinion that beginners should consider for their work. The research did not bring a ready-made solution which framework is better than someone else's, but showed what other users recommend for work. The paper singled out two frameworks according to the grades they received, but also according to the number of users who recommended them. Django is definitely the first choice of the user, which is quite logical considering its application, documentation and the support that the community provides to it. The second frame is Flask and it is the first in terms of ratings, but a smaller number of users recommend it compared to Django. Flask is also very useful, which can be seen from the listed projects for which it is used in the world, but users still recommend Django first and only then Flask.

The choice is up to the user, and this work and research in it has managed to reduce the eight offered frameworks to two from which users can choose, again depending on the need, knowledge, support or some subjective feeling. Each of these frameworks is a good solution and as such should be viewed, but the user's experience has given its recommendation and as such should always be taken into account when choosing and deciding on the framework that the user chooses for learning and further development.

## REFERENCES

- [1] Ruba Abu Khurma, Ibrahim Aljarah and Ahmad Sharieh () EvoloPy-FS: An Open-Source Nature-Inspired Optimization Framework in Python for Feature Selection, School for Information Technology, The University of Jordan, Amman, Jordan.
- [2] Hina Afzal, Arisha Kamran, Asifa Noreen (2020) Survival Analysis of Python and R within the Job Market Trend, Journal of Information Technology and Computing, 2020 Volume 1, Issue 1 : 31 – 40, DOI : 10.48185/jitc.v1i1.94
- [3] Rašić Nino (2018) Web aplikacije za prikupljanje i analizu meteoroloških podataka, University of Pula, Croatia, Master's thesis.
- [4] Vangala Rama Vyshnavi, Amit Malik (2019) Efficient Way of Web Development Using Python and Flask, Vangala Rama Vyshnavi et al. International Journal of Recent Research Aspects ISSN: 2349-7688, Vol. 6, Issue 2, pp. 16-19.
- [5] Devndra Ghimire (2020) Comparative study on Python web frameworks: Flask and Django, Metropolia University of Applied Sciences, Bachelor's Thesis.
- [6] Vogel, P., Klooster, T., Andrikopoulos, V., & Lungu, M-F. (2017). A Low-Effort Analytics Platform for Visualizing Evolving Flask-Based Python Web Services. 109-113. Paper presented at 2017 IEEE Working Conference on Software Visualization, Shanghai, China. <https://doi.org/10.1109/VISSOFT.2017.13>
- [7] Barbares, A., (2020). htmldate: A Python package to extract publication dates from web pages. Journal of Open Source Software, 5(51), 2439. <https://doi.org/10.21105/joss.02439>
- [8] K. R. Srinath (2017) Python – The Fastest Growing Programming Language, International Research Journal of Engineering and Technology (IRJET), Volume: 04 Issue: 12.
- [9] Kameryn Larkins (2020) A Comparison of Python Web Development Microframeworks, Southern Connecticut State University, New Haven, Connecticut, Master of Science.
- [10] Marija Ševković (2020) Mreža mikroservisa sa distribuiranom bazom podataka u okruženju Flask, Matematički fakultet, Univerzitet u Beogradu, Master's thesis.
- [11] Kamalim, O., & Ayub, M. (2017). The Use of PythonTutor on Programming Laboratory Session: Student Perspectives. Kinetik : Game Technology, Information System, Computer Network, Computing, Electronics, and Control, 2(4). doi:<http://dx.doi.org/10.22219/kinetik.v2i4.442>
- [12] Hongyu Zhai, Casey Casalnuovo and Prem Devanbu (2019) Test Coverage in Python Programs, MSR '19: Proceedings of the 16th International Conference on Mining Software Repositories May 2019 Pages 116–120 <https://doi.org/10.1109/MSR.2019.00027>

# Analysis Of Students' Learning And Achievement Based On Data From The University Information Systems

Ilker Ali<sup>\*</sup>, Natasha Blazeska-Tabakovska<sup>\*\*</sup>, Igor Nedelkovski<sup>\*\*</sup>, Blagoj Ristevski<sup>\*\*</sup>

<sup>\*</sup> International Vision University, Gostivar, Republic of North Macedonia

<sup>\*\*</sup> University "St. Kliment Ohridski" - Bitola, Faculty of Information and Communication Technologies  
Republic of North Macedonia

ilker@vizyon.edu.mk, natasa.tabakovska@uklo.edu.mk, igor.nedelkovski@uklo.edu.mk, blagoj.ristevski@uklo.edu.mk

**Abstract** – The coronavirus pandemic upended almost every aspect of life. Education has undergone significant changes. The need for online teaching has emerged. Several questions have arisen as to how effective online teaching is; what are the advantages and disadvantages? This paper analyzes the impact of student attendance and activity on the learning outcomes of students as well as a success during online and regular teaching. The paper analyzes the downloaded log files from the Microsoft Office 365 portal for the used programs Microsoft Teams and Microsoft Forms and the log files from everyday activities from University Information System – UIS in North Macedonia. Using the Python application the data obtained from the University Management System – UMS are extracted about the achieved success of the students. Additionally, log files for student activities are downloaded from the Microsoft portal.

**Keywords:** E-learning, Online Learning, Learning analytics, Microsoft Teams, Microsoft Forms, University Information System.

## INTRODUCTION

The transfer of knowledge or education is almost equal to human history. The history of e-learning is diverse, filled with success stories from the early 1900s and success stories from the 2000s [1]. Until the 80s of the previous century, the classical education era of lecturers lasted -1983 [2]. Between 1994 and 1999, the first wave of e-learning began to be used. With the development of Internet speed, LAN, wireless Internet, and the advent of Web 2.0 in 2000, the second wave of e-learning began. The emergence of mobile phones and the Mobile Web application in 2008 and HTML 5 in 2010 should also be emphasized, which have made a great contribution to the development and use of e-learning [3].

LMS is a multi-user software application usually accessed through a web browser [5] This application manages distance learning, self-study courses, and mixed learning programs. Provides automation that replaces rigorous and expensive manual work, saves time, and allows content, data, and students to be organized in cyberspace. This application also monitors and reports on the activity and training results of a particular course.

LMS grew out of a series of multimedia and internet developments in the 1990s. In the last four years, the systems have matured and been adopted by many

universities worldwide. LMS are scalable systems that can be used to support university-wide learning and learning programs [6]. With the development of LMS, the need for learning data analysis has emerged.

In recent years, there has been a growing interest in automated data analysis to enhance the learning experience, a research area called learning analytics [4].

Learning Analytics (LA) has attracted a great deal of attention in recent years as educational institutions and researchers increasingly see LA's potential to support the learning process. LA approaches share the movement from data to analysis to action to learning [7].

Olga Viberga's analysis shows that the field of LA is still a developmental area of practice and research in which descriptive studies and methods of interpretive data collection predominate, according to Papamiciu and Economid (2014) [8].

The use of student data by AU institutions calls into question students' privacy. Kyle M., with a developed model, argues that students should be more informed about how their institution uses recognizable data and information and what it ends up with and gain purposeful controls over the flow of information. This proposed awareness and consent model ultimately supports student privacy and autonomy [9].

Mihaela Cocea and Stephan Weibelzahl analyzed the educational institution log files and assessed the level of student motivation. According to them, a prediction module should be set up in the basic systems [10].

Information from the International Vision University in the Republic of North Macedonia is used in this paper. The University has a University Management System, which students, professors, and student affairs use. In the following articles, the University Management System will be called UMS. UMS generates data or log files from everyday activities. Microsoft Portal 365 is used during the online teaching, from which the applications Teams, One Drive, One Note, Forms, etc., are used primarily.

From the data obtained from the University Management System - UMS, using the Python application, information is extracted about the achieved success of the students (obtained grades). Additionally,

log files for student activities are downloaded from the Microsoft portal from the admin panel. The paper analyzes the impact of student activity on success during online teaching in the summer semester 2019/2010 academic year and the winter semester 2020/2021. The paper also analyzes the difference between success in online learning and regular learning during the 2020 and pre-2020 pandemics.

The rest of the paper is structured as follows. The second section describes the used system in online education. In section III the used research methodology is explained. The results of data analyses are present in section IV. The final section gives concluding remarks and points out some directions for future works.

## II. SYSTEM USED IN ONLINE EDUCATION

### University Management System

This research used information from the UMS from the International Vision University in Republic of North Macedonia. UMS brings together the common data of university students and faculty members for their intended purpose. Combined data is processed without compromising relational integrity, resulting in significant results. The results from the data processing are prepared as a report. With user role management, data security is ensured, and the data accessed by each role is limited. Microsoft technologies are widely used in the UMS infrastructure. UMS using C #, Asp.Net MVC, MSSQL, HTML, CSS, Bootstrap, JQuery, and Javascript technologies is hosted on Internet Information Services (IIS).

University Management System (UMS) is a web-based application developed with the C # programming language, .Net Framework MVC, which supports users with disabilities. UMS is composed of several modules: Student Management System, Academic Information System, Student System. These modules are interconnected, and the goal is to assist the student service of the academic staff - professors, students, and the administration for the successful implementation of the educational process at the University. Another example of support for users with disabilities where the interface is made with WCAG 2.0 standards, a screen without many objects and elements so as not to confuse users with disabilities are given in the paper [11].

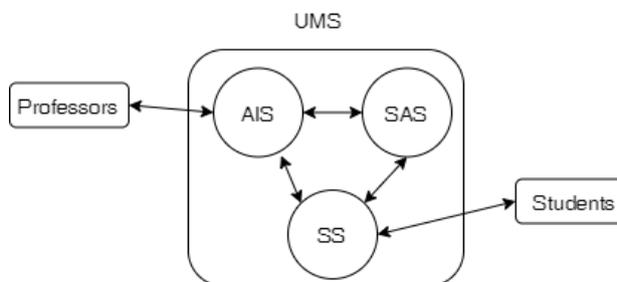


Figure 1- University Management System

### Microsoft Portal 365 as an LMS

During the 2020 pandemic, teaching took place online. Online education took place with the Microsoft Portal, mainly with the Teams application. In addition to

this application, One Drive and One Note applications are used, from where students download the materials set by the professors for their subjects. During the exams was used Microsoft Forms application was used with Microsoft Teams (Figure 1).

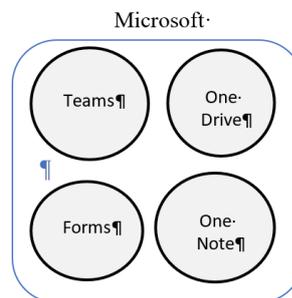


Figure 2- Microsoft LMS.

## III. METHODOLOGY OF SCIENTIFIC RESEARCH

The LMS used at the University consists of two parts. The first is the University Management System - USU explained in more detail above. It can use AIS, students can download or follow from the SS portal or upload to One Drive. The second part is Microsoft applications in the Portal 365 package such as Teams, One Drive, Forms, and One Note. As can be seen from the picture, professors and students can log in to both systems with the same username and password. The professor can decide where to place the materials for the lesson they teach.

Microsoft applications are interconnected, ie, the material placed in One Drive can be easily shared on One Note or Teams. Every action of the professor in AIS, such as posting materials, comments, or messages to students, writing grades are visible in the Student System (SS), i.e., the student can see/access them. Also, the new data are ready to be used in the reports by the Student Affairs Service in SAS (Figure 3).

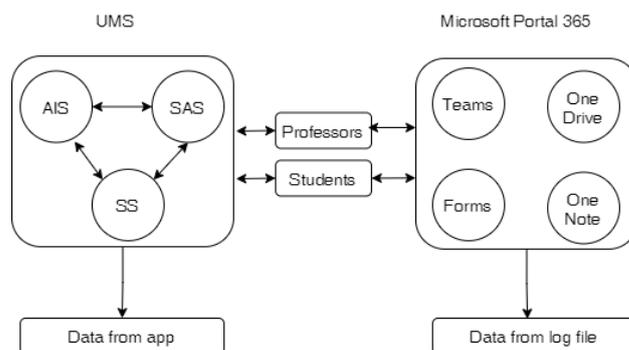


Figure 3 - Used LMS at the University

AIS - Academic Information System

SMS - Student Management System

SS - Student system

### Evaluation

Our goal in this research is to analyze student achievement-related activity. The following figure (Figure 4) shows the database from which the data and log files are taken. The data are taken from a university in North Macedonia.

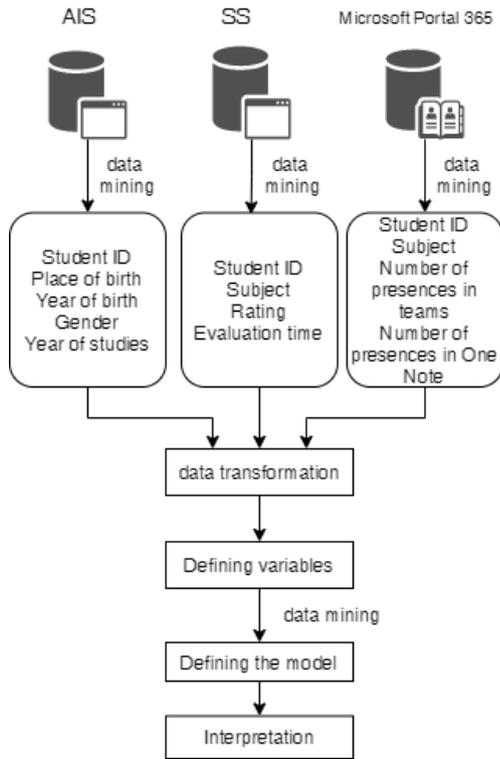


Figure 4 - Data model

The data Student ID, year of birth, gender, and year of studies are taken from SUS. From AIS, we extract the Student ID data and, with this ID, related grades received by the students, subject, and time of assessment (Figure 5).

	A	B	C	D	E	F	G	H	I
1	Numara	Ad	Soyad	Eposta	Cinsiyet	NotYil	NotDonem	DersPlanYil	DersPlan
48848	5304/19	Premtim	Mulaj	premtim.mulaj@vizyon.edu.mk	1	2019	2	2019	2
48849	5304/19	Premtim	Mulaj	premtim.mulaj@vizyon.edu.mk	1	2019	2	2019	2
48850	5307/19	Arife	Türkmen	arife.turkmen@vizyon.edu.mk	2	2019	2	2019	2
48851	5323/19	Nurcan	Sucubasi	nurcan.sucubasi@vizyon.edu.mk	2	2019	2	2019	2
48852	5323/19	Nurcan	Sucubasi	nurcan.sucubasi@vizyon.edu.mk	2	2019	2	2019	2

Figure 5 - Grades from AIS

During the online teaching, the regularity and activities of the students are registered, and that data can be downloaded from the Microsoft portal (Figure 6).

Figure 6 - Regularity of students

Using applications during the pandemic (Figure 7)

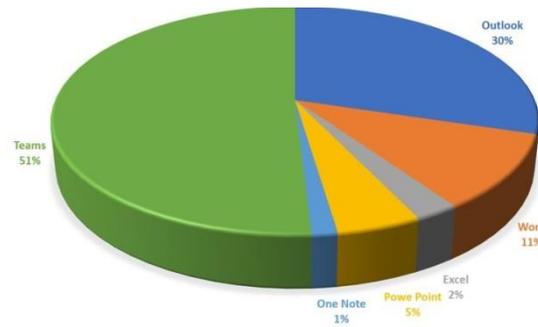


Figure 7- Percentage of applications used

The data is analyzed in the Python programming language. Python is a high-level scripting language, interpreted, interactive, and object-oriented. Python is designed to be very readable. It often uses keywords in English, where other languages use punctuation and have less syntactic structure than others [12].

Python aims to be the most powerful and flexible open-source data analysis/processing tool available in any language. Data mining can also be done with Python, and any kind of analysis of this data can be performed. But to use Python, you need to know a certain level of programming.

For Python visualization, must install special packages. One of the most important packages is the Panda package [13].

The Panda package provides fast, flexible, and expressive data structures designed to make it easier and more intuitive to work with "relational" or "tagged" data. It aims to be the primary blockchain for conducting practical, realistic data analysis in Python.

#### IV. DATA ANALYSIS

The data analysis is explained below.

1. The downloaded data from Microsoft Portal and UMS are merged first. From SMS, we take the data Student ID, gender, and year of studies. From AIS, we extract the data Student ID, exam grades, subject, and assessment time. We merge the two tables into a new file called IlkerQuery.csv shown in Figure 18 - AIS Ratings. Data on student attendance and activities (Figure 6 - Student regularity) can be downloaded from the Microsoft Admin panel and saved under Teams4.csv.

This achieving in the following way. The following codes are writing in the Jupyter editor:

```
import pandas as pd
from matplotlib import pyplot as plt
```

We are importing pandas and matplotlib packages.

```
df1=pd.read_csv('Teams4.csv')
df2=pd.read_csv('IlkerQuery.csv')
df1a=df1.set_index('user')
df2a=df2.set_index('user')
```

This procedure combines the data needed for analysis. CSV files downloaded from UMS, from Microsoft Portal are imported in df1 and df2 modifications. The index is defined in both files.

```
df3=pd.merge(df1, df2, how='right')
```

We merge the data with a right join.

```
[4]: type(df3)
Out[4]: pandas.core.frame.DataFrame
```

D	E	F	G	H	I	J
EPosta	Cinsiyet	NotYil	Notu	DersPlanYil	DersPlanDonem	DersAciklama
mesut.koca@vizyon.edu.mk	2	2019	2	2019	2	Araştırma
nilgun.aydemir	2	2019	2	2019	2	Araştırma
nurcan.sucubasi@vizyon.edu.mk	2	2019	2	2019	2	Araştırma
premtim.mulaj@vizyon.edu.mk	1	2019	2	2019	2	Araştırma
suat.eyuboglu@vizyon.edu.mk	1	2019	2	2019	2	Araştırma
yasemin.sabani@vizyon.edu.mk	2	2019	2	2019	2	Araştırma
turgay.sarac@vizyon.edu.mk	1	2020	1	2020	1	Sosyoloji

Figure 8 - Merged data

2. After merging the data, various processing and analysis can be done on them. The impact of the activity and regularity on the achieved success of the student was analyzed.

```
plt.plot((df3.MeetingCount+df3.CallCount), df3.YuzlukNot, '.')
plt.xlabel('aktivnost')
plt.ylabel('uspeh')
plt.show()
```

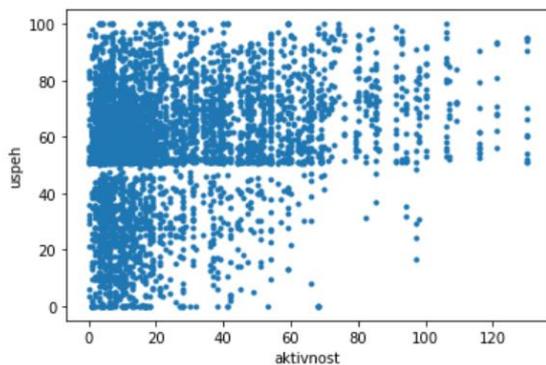


Figure 9 - Graphic representation of the results

The following code is used to extract success at a given time. In the example, the grades in 2019 are filtered for the second half.

```
df4 = df3[(df3.DersPlanYil==2019.0) & (df3.DersPlanDonem==2.0)]
```

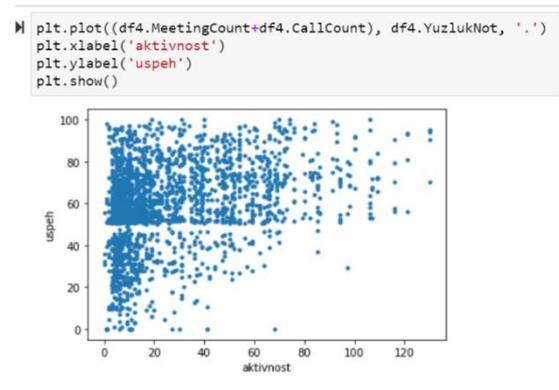


Figure 10 - Success at a given time

To calculate the average, we write the following codes:

```
import array
plt.xlabel('aktivnost')
plt.ylabel('uspeh')
EnYukseDeger=int((max(df4.MeetingCount)))
print(EnYukseDeger)
for i in range (0,EnYukseDeger):
    df6=df4.query("MeetingCount == @i")
    if df6.YuzlukNot is None:
        df6.YuzlukNot=50
    avg=mean(df6.YuzlukNot)
    plt.scatter(i,avg,c='blue')
plt.show()
```

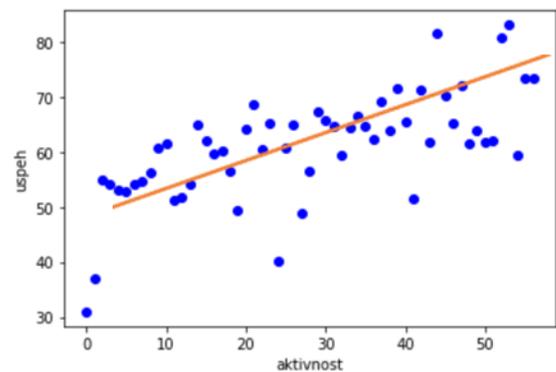


Figure 11 - Mean

The result of the examination, whether there is a connection between the activity and the success of the students, is shown in the graph where the regularity/activity is represented on the x-axis and the achieved success of the students on the y-axis (Figure 11). From the graph, we can see that there is a positive relationship between activity and success. That is, by increasing the value of regularity/activity, the value of success increases. Regularity/activity has a positive impact on student success.

In order to investigate the impact of our platform - UMS and the weight of the conducted exams. In addition, the distribution of student success during teaching online was analyzed. The results of the online exams were downloaded from UMS and Microsoft platforms in csv format. The downloaded results were processed in the Python programming language with the following codes.

```
In [2]: import seaborn as sns
import pandas as pd
import numpy as np

In [3]: df1=pd.read_csv('Teams4.csv')
df2=pd.read_csv('IlkerQuery.csv')
df1a=df1.set_index('user')
df2a=df2.set_index('user')

In [4]: df3=pd.merge(df1, df2, how='right')

In [10]: sns.distplot(df3.YuzlukNot)
```

When processing the exam results, the following graph is obtained (Figure 12).

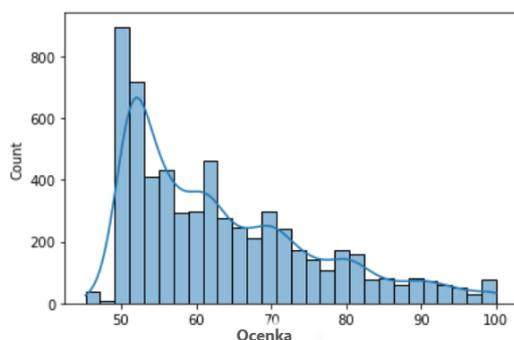


Figure 12 - Distribution of grades

In the graph, we see that the distribution is distorted on the left: Arithmetic mean  $\leq$  median  $\leq$  fashion. According to the results, teaching is appropriate; in other words, students achieve the goals planned in the curriculum. Exams are of the proper weight. This distribution results following the mastery learning model are observed (i.e., all students' material is scientific). The learning level of the students is good.

The second research in this paper is shown in Table 1. Given the success achieved before and after the pandemic with Covid-19, as we can see from the table of our University, there is no significant decline in the average success of students which we can conclude that the UMS system provides us with an opportunity in the successful implementation of teaching at MUV (International University Vision)

**Table 1.** The success of exams at Vision International University.

	Пред COVID-19	Поср COVID-19
Valid	2806	2592
Missing	908	1122
Mean	61.78	59.11
Std. Deviation	11.214	19.659

## V. CONCLUSION

In the most intensive period of digitalization with the development of the USU system, we also aim to respond to the University's needs and give a new meaning to the data being processed. The USU system provides many data from parameters such as course records, group courses, student records, and staff records. Our system is a web-based application with its practical application, and students are provided access from anywhere in the world where they have a connection to the Internet. Also, thanks to the web-based approach, this system provides easier access to students attending courses through distance education and online learning.

In applying the UMS system, strengths and weaknesses have been identified and some issues that are requiring further research. Such a software solution helped us a lot in the realization of the curriculum during the COVID-19 pandemic. Our software application had a positive effect on the realization of the teaching and educational process at our University.

Based on the data we received from the university system and the research we have done with these data, we can say the following: There is not a big difference in students' success between face-to-face education before the pandemic and online education during the pandemic period. Students' success is better in face-to-face education, but there is not a big difference, which shows that it makes sense to apply online education during the pandemic period.

During the pandemic, we calculated the continuity of students' participation in online classes in different periods. We calculated the duration of each student's involvement in the course and the grade he received from that course. We noticed this: As the time to attend class increases, poor grades decrease. This means that participating in online classes increases success.

The problems that arose during the application of the UMS system were related to the inadequacy and weakness in the skills for digital competence of the academic staff and students, and the adaptation process lasted three weeks.

One of the essential findings we have identified is that teachers need to be given more practice to prepare exams in the Microsoft Forms forms and publish the exams on time. There was no problem with the connection during the realization of the exams.

*The recommendations of this research are as follows:* the students should attend online classes for better success; the teachers should pay more attention to online exams, the questions should be timed, a pool of exam questions should be created, and questions will be randomly selected from the question bank; the management and educators should use the results of the data analyses.

*This research also raises some questions that require further study such as:* What type of courses in online education gives better exam results? What types of exam questions provide better results? How does multimedia affect students' abilities and instructors' ability to teach?

## REFERENCES

- [1] S. Tamm, "The History of E-Learning," e-Student, 21 December 2019. [Online]. Available: <https://e-student.org/history-of-e-learning/>.
- [2] D. Janković, P. Rajković and D. Vučković, "MULTIMEDIJALNI SISTEMI KAO DEO SISTEMA ZA ELEKTRONSKO UČENJE," in *XXIII Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju*, Beograd, december, 2005.
- [3] K. Gutierrez, "10 Great Moments in eLearning History," Shift e-Learning, 2014. [Online]. Available: <https://www.shiftelearning.com/blog/bid/343658/10-Great-Moments-in-eLearning-History>.
- [4] L. Johnson, "Horizon Report. Austin," New Media Consortium, Texas, 2011.
- [5] S. D. Foreman, *The LMS Guidebook: Learning Management Systems Demystified*, American Society for Training & Development, 2018.
- [6] H. Coates, R. James and G. Baldwin, "A Critical Examination Of The Effects Of Learning Management Systems On University Teaching And Learning," *Tertiary Education and Management*, vol. 11, p. 19–36, 2005.
- [7] A. C. Mohamed, A. L. Dyckhoff, U. Schroeder and H. This, "A Reference Model for Learning Analytics," *International Journal of Technology Enhanced Learning (IJTEL) – Special Issue on "State-of-the-Art in TEL"*, vol. 4, no. 5/6, 2012.
- [8] O. Viberga, M. Hatakkab, O. Bältera and A. Mavroudia, "The current landscape of learning analytics in higher education," *Computers in Human Behavior*, vol. 89, pp. 98-110, December 2018.
- [9] J. Kyle, "Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy," *International Journal of Educational Technology in Higher Education*, p. 16:24, 2019.
- [10] M. Cocea and S. Weibelzahl, "Can Log Files Analysis Estimate Learners' Level of Motivation?".
- [11] K. Veljanovska, N. Blazeska-Tabakovska, B. Ristevski and S. Savoska, "User Interface for e-learning Platform for Users with Disability," in *The 13-th conference on Information Systems and Grid Technologies ISGT2020*, Sofia, Bulgaria, 2020.
- [12] "Python (programming language)," 2021. [Online]. Available: [https://en.wikipedia.org/wiki/Python\\_\(programming\\_language\)](https://en.wikipedia.org/wiki/Python_(programming_language)).
- [13] "About pandas," 2021. [Online]. Available: <https://pandas.pydata.org/about/>.

# Improving Test Execution Phase through Diversity of Approaches: A Systematic Literature Review

Sara Gračić\* and Vuk Vuković\*

\* University of Novi Sad, Faculty of Economics, Subotica, Serbia  
saritta4u@gmail.com, vuk.vukovic@ef.uns.ac.rs

**Abstract** – Comprehension of what testing used to be has changed over the years. From being considered equivalent to debugging, testing became a separate process with four main phases: planning, design, execution and evaluation. Testing is a resource-consuming activity. Approximately 50% of total projects costs are related to testing, especially with the increase in software size and complexity. By constantly improving testing phases, it is possible to optimize necessary resources. The main focus of this paper is to investigate the latest achievements in the execution phase of software testing. Therefore, a systematic literature review was conducted, four scientific repositories were chosen and after applying inclusion and exclusion criteria, the search resulted in 17 papers. This paper provides deeper insight into the latest trends regarding test execution improvement for both industry and academia. It reports how parallelism, redundancy and similarity of test steps were exploited to improve test execution phase. Further development of test types and acceleration of test execution using GPU were also proposed. Regardless of the approach, test execution time reduction had been reported, which could potentially decrease test execution costs.

## I. INTRODUCTION

At an early stage of development, software testing was considered the same as debugging. However, that understanding of testing has changed. From one of the development phases, it became a standalone process (Software Testing Process). Various software testing models (e.g. ISO/IEC 29119, TMMi, TMap Next) were created to introduce systematic testing in software organizations [1], [2], [3]. As with any other process, software testing can be better. Therefore, its improvement has been in focus of practice and academia. Afzal and Torkar identified four key phases in software testing process: test planning, test design, test execution and test evaluation [4]. This paper is focused on investigating the latest achievements in the phase of test execution.

After the Introduction, the rest of the paper is structured in the following order: Section II presents research methodology (research questions, chosen scientific databases, search strings, inclusion and exclusion criteria, and how the review was conducted); Section III extracts the review data, while Section IV discusses findings. Conclusion is given in Section V.

## II. RESEARCH METHODOLOGY

Systematic literature reviews (SLRs) are focused on presenting detailed information about a specific phenomenon [5] and are vital in scientific research [6]. Fairness and thoroughness are exceptionally important [5]. Otherwise, a review's scientific value would be at risk. SLRs are a form of identification, evaluation, and interpretation of available research [5] regarding a certain topic. A SLR has three main phases: planning a review, conducting a review, and reporting review findings [5].

### A. Planning the Review

Each phase of software testing is of the utmost importance, and both industry and academia are participating in its improvement. This review targets the execution phase of software testing to identify and summarize the available knowledge and current trends related to advances made in this particular phase.

#### 1) Defining Research Questions

Based on the defined objective, the following research questions were constructed. **RQ1:** *How many studies focused on software test execution were published from 2010 till August 23<sup>rd</sup>, 2021?* **RQ2:** *What were the main topics discussed in these papers?* **RQ3:** *What are current trends in the area of software test execution?*

#### 2) Defining Search String for Scientific Databases

To answer research questions in focus of this paper, a SLR was conducted to identify topics and trends discussed in academia. Four scientific repositories were chosen for conducting the SLR: IEEE Xplore, SpringerLink, Science Direct and ACM Digital Library. The research string was set to “test execution” and applied in the above-stated repositories.

#### 3) Defining Inclusion and Exclusion Criteria

Publishing period was set to last 11 years - from 2010 till August 23<sup>rd</sup>, 2021, to collect the sufficient and relatively recent research findings. Only journal and conference papers were included in the initial population. Grey literature (e.g. blog posts, white papers) was not examined for this SLR. Only primary studies in the domain of software test execution were reviewed. Language, although initially not an obstacle for entering

the initial population, may become one in the next stages (if title or abstract are not in English).

### B. Conducting the Review

Advanced search was chosen in all repositories. After applying search string and criteria “document title” in IEEE Xplore, “where the title contains” in SpringerLink, “title” in Science Direct and “title” in ACM Digital Library, the search resulted in 40 papers, structured as follows: IEEE Xplore (18), SpringerLink (6), Science Direct (1) and ACM Digital Library (15).

Titles, keywords and abstracts were examined to confirm that a study is in function of answering research questions. All papers that were not in the field of interest were eliminated.

The only paper from Science Direct was irrelevant to the SLR, so the repository was eliminated. All 6 papers from SpringerLink were eliminated due to format (chapters) and relevance. After eliminating 5 duplicates present in ACM Digital Library and IEEE Xplore, abstract and domain-specific papers were eliminated from ACM Digital Library, while 9 papers from IEEE Xplore were eliminated based on relevancy and format (magazine article).

The final population has 17 papers, 12 from ACM Digital Library and 5 from IEEE Xplore.

## III. EXTRACTING REVIEW RESULTS

Data collected from 17 papers are summarized in the following paragraphs.

It was observed in [7] that functional tests are mostly done manually based on hand-written notes, despite tools that support automation of functional tests. However, these do require programming skills and are hard to use in practice. So, the authors semi-automatically converted hand-written notes into automated tests. In the first stage, machine learning and natural language processing were used to compute “an intermediate representation from test steps”, and in the second “interactively disambiguating that representation to create a fully automated test”. The approach was evaluated and it was shown that 70% of manual test steps could automatically be converted into automated test steps without human interference.

Authors in [8] noted that executions of A/B tests do have theoretical methodologies. However, without case studies and/or guidelines to support them, they have a high level of abstraction. As such, implementing them in industry brings challenges for practitioners. So, IPEAD (Ideation, Prioritization, Execution, Analysis and Documentation) framework for execution of A/B tests was proposed. It is based on the High Testing Tempo framework and Lean Startup book. The authors argued that High Testing Tempo is not detailed in the execution of A/B tests (e.g. when should a test be stopped), as well as how collected results are being analyzed. Their framework has five steps (ideation, prioritization, execution, analysis and documentation), instead of four steps from the High Testing Tempo framework. In execution, their framework provides a set of rules for stopping tests, as well as more detailed guidance

regarding other four steps. IPEAD was also used in a case study, where users were stimulated, via a web page, to share content on Facebook. The case study was presented in detail and results were promising.

E2E (End-to-End) tests are an exceptionally important part of testing process. Improvement of execution of these tests could be done through more efficient usage of necessary testing resources. Based on the previously published approach – RETORCH (Resource-Aware End-to-End Test Orchestration), Augusto investigated how to optimize resources for execution of E2E tests in [9]. Three processes: resource identification, grouping and scheduling were in focus of optimization. His approach was implemented in practice. It succeeded in identifying and orchestrating smart resources, which resulted in deploying only needed functionality (not all). Approximately, orchestrated test suit required 61% less time than non-orchestrated test suite and usage of physical memory was also reduced. This has led to total cost reduction of test suite.

In [10], authors were focused on improving correctness of test execution by implementing checkpointing for test isolation in domain of web applications. Their method enables web applications to support checkpointing and then use that support to isolate and optimize tests. Five PHP applications were used for evaluation of the developed method and it was shown that this method offers test isolation “essentially for free” and it also decreases testing time by 44%.

Authors in [11] pointed out that when testing client-side web applications, different browsers and platforms have to be taken into consideration. Also, evolving web applications require regression testing across all environments, which generates higher testing costs and delayed failure feedback. Therefore, six techniques for more rapid failure feedback in cases of regression testing across various environments were proposed. They applied the same principles of test case prioritization, only they prioritized environments based on recent and frequent failures. The approach was evaluated on four open-source web applications. It outperformed no ordering and random ordering methods, when cost-effectiveness was evaluated. Improvement “rates ranged from -12.24% to 39.05% for no ordering and from -0.04% to 45.85% for random ordering”.

Instruction locality problem in testing was explored in [12]. It was noted that other papers were focused on improving locality over single program runs, but one program could be executed multiple times. Therefore, instruction locality across the execution of test cases was proposed and test cases were also permuted to enable subsequent runs to share as many instructions as possible.

Since quick localization of a fault is important, authors in [13] proposed SPEQTRA, a heuristic that mines executions traces of tests (failed and passed) to find faulty class. It also ranks classes based on their probability of being faulty. SPEQTRA was then compared with AMPLE on a small, representative case. It was shown that SPEQTRA surpassed AMPLE in class ranking (SPEQTRA and AMPLE are both instances of

the class of spectrum based fault localization heuristics) [13].

Bach, Andrzejak and Pannemans [14] researched coverage-based testing approaches in very large industrial software project SAP HANA to reduce time and costs of test execution. Execution time reduction is of vital significance, especially in large software projects, since the cumulative execution time of all tests (around 130,000) is about 250 hours or more than 12 days. Their research has shown that in the observed large software project, an overlap-aware greedy approach for test case selection and prioritization achieved better results than standard greedy (by factor of 1,5). Another important conclusion was that up to 97% of all line coverage lines do not have unique behavior and are redundant, which carries potential for data size reduction.

Authors in [15] developed a prototype TAO (Test Analyzer and Optimizer), for optimization of execution of test units and reduction of wait time from minutes to seconds in large software programs. TAO “implements simple and fast function-level dependency tracking that identifies tests to run on a code change; any false negatives missed by this dependency tracking are caught by running the entire test suite on a test server once the code change is committed”. It also provides developers with dependency information for easier adoption. The prototype showed that it can decrease execution time by over 96% on two large Python projects.

Derivation of test plan was central topic in [16]. The authors noted that parallelism in test suits is not exploited, because main concerns are costs and risks. Test execution plans are not capable of handling dynamic events and test execution order is of at-most-importance. So, the authors presented an approach for creating execution plan that will “facilitate parallel execution, given resource availability and test case dependencies”. The results have shown reduction of test execution time. Although this approach is aligned with automated test execution, in manual execution other things have to be taken into account e.g. a tester’s domain knowledge.

Authors in [17] stated that evolution of test suits leads to accumulation of redundant tests, which has been in focus of researchers. They also noted that more subtle problem, which was not explored, is accumulation of similar tests. Duplicated actions are executed repeatedly, which negatively affects execution time. So the focus of their paper was on reduction of number of duplicated steps. The technique identifies tests that can be merged and generates merged test, which covers all application states, but duplicated steps are only executed once. It also does not affect the detection of faults. The technique was also applied in practice on four open-source web applications and one enterprise web application. Over 3300 test cases and 19600 test steps were analyzed. The test merging technique for GUI decreased the number of test steps by 29%, while execution time was decreased by 39%.

Having in mind reusability and applicability, Rahman, Chen and Gao [18] designed the parallel test execution service framework, with minimal application specific customization needs. For most applications, developers

and testers have to configure test jobs using the standard configuration file. In addition, some environment variables (e.g. application port, database port) need to be generated and injected properly in the worker template for the application under test. The average execution time was reduced by 66.67% using four worker sandboxes. It was concluded that the execution time could further be reduced by increasing the number of workers with higher resource allocation.

Authors in [19] pointed out that research is oriented towards predicting software defects from development process and final product. However, that is not possible when a company is conducting tests for a contractor, since the only source for defect prediction is its own quality of testing process. So, they extracted in-process testing metrics “from acceptance test execution logs of a large scale software application developed at Ericsson Turkey”. Measuring was conducted during 15 weeks and defects were predicted on a weekly basis. It was emphasized, however, that the number of training instances was not enough for solid conclusions regarding the model’s performance and that data accuracy was also problematic. The model was kept as a successful prototype.

A set of test cases can have a significant number of test steps occurring in more than one test case. During execution, these duplicated steps increase test load while providing none or limited additional test results. To solve this, Walter, Schilling and Piechotta [20] proposed test case synthesis. After removal of redundant steps, all (non-redundant) test steps were also rearranged into a new set of test cases. This was achieved by clustering of similar test steps into new test cases. A Path finding algorithm was used to find an optimized test step execution order for each test case. By applying this method in a case study at Mercedes-Benz Passenger Car Development, the authors observed a test load reduction by 15% due to the removal of redundant test steps and an additional reduction of at least 3% due to the rearrangement of test steps. This led to at least 18% overall test load reduction and the authors argued that this was a strong indicator of usefulness of their approach.

Yaneva, Kapoor, Rajan and Dubach presented an approach which was focused on accelerating the execution of functional tests of finite state machines [21]. The acceleration was achieved by running tests parallelly on GPU threads. The authors considered various design options for the encoding of the FSM (Finite State Machines) and its test inputs in order to maximize acceleration achieved by using GPU. They also evaluated the approach by using 13 subject FSMs from the network intrusion detection and signal processing domains. For each of the 13 subject FSMs they generated full test suites based on the all-transition pair coverage criteria. Their approach led to several conclusions. Maximum acceleration of 12× was achieved when compared to a 16-core CPU. GPUs generally achieved higher acceleration for FSMs with long test sequences. Using a dense FSM layout resulted in higher GPU acceleration, particularly for FSMs represented as dense matrices. When test layout was considered, with-offsets layout tended to perform better for FSMs with shorter test sequences, while

padded-transposed layout was more suitable for FSMs with long sequences. Prioritizing test sequences before execution on the GPU improved acceleration.

Authors in [22] stated that testing is execution of one program with multiple data sets and that it, therefore, manifests data parallelism which could be harvested by using GPUs. Their approach “simultaneously executes the program with one test case per GPU thread”. Test executions were accelerated by a factor of 27 when compared to single-core executions on CPU.

In [23], authors stated that test execution techniques were chosen without any specific rules. Experience, rule of thumb and best practice were used by experts, but that may not be cost-effective. So, they proposed a cost model for estimation of economic impact of test execution models. It incorporates expert’s estimation and provides them with additional input in “balancing pros and cons of execution modes at hand”. The authors also applied their model in practice.

#### IV. DISCUSSION

The idea regarding the scope of software testing has evolved over the years, due to the proliferation of software in every part of our lives - from mobile phones to safety-critical systems. This change has also had a tremendous impact on growth of software’s size and complexity [24]. Also, testing software of that robustness consumes a lot of time and effort [24] and up to 50% of total project costs are related to testing [25], [26].

Improving testing process has the potential to reduce resource consumption and ensure the highest possible quality of the final product (software). Therefore, testing is today perceived as a stand-alone process and academia and industry are actively seeking ways to make it better.

According to Afzal and Torkar, testing process has four key phases: test planning, test design, test execution and test evaluation [4]. Improving each phase is important in reducing resource consumption as well as developing a high quality final product.

The authors of this paper decided to focus on test execution phase and a SLR was conducted. Four scientific repositories were chosen - IEEE Xplore, SpringerLink, Science Direct and ACM Digital Library. “Test execution” string was applied in the above-chosen repositories. After evaluation, 17 papers entered the final population. This answers the first research question (**RQ1: How many studies focused on software test execution were published from 2010 till August 23<sup>rd</sup>, 2021?**)

Findings that were extracted from 17 papers provide answers to second and third research question (**RQ2: What were the main topics discussed in these papers?** and **RQ3: What are current trends in the area of software test execution?**). The papers are grouped based on the research topic and discussed in the following paragraphs.

##### A. Test types in execution improvement

Despite automated testing, *functional tests* are conducted manually using hand written notes, so those

notes were semi-automatically converted into automated tests and evaluation showed that 70% of manual test steps could automatically be converted into automated test steps without human interference. It was also noted that automated functional tests require programming skills and are hard to use in practice [7].

Although *A/B tests* are backed up theoretically, they do not provide proper support for implementation. To overcome that, IPEAD framework for execution of A/B tests was proposed. It has 5 steps and provides users with detailed guidance regarding all 5 steps. The framework was applied in practice and its results were promising [8].

Optimization of resources for *E2E tests* was in focus of [9]. Three processes: resource identification, grouping and scheduling were investigated and it was shown that orchestrated test suite required 61% less time than non-orchestrated test suite, physical memory usage was reduced. These led to total cost reduction of test suite [9].

##### B. Test execution in web applications

Checkpointing was used to isolate tests designed for web applications, with the purpose of improving correctness of test execution. After evaluation on 5 PHP applications, it offered test isolation “essentially for free” and it also decreased testing time by 44% [10].

During testing of client-side web applications different browsers and platforms have to be taken into consideration. Also, evolving web applications require regression testing across all environments, which generates higher testing costs and delayed failure feedback. So, six techniques were proposed to tackle with these issues. Prioritization was applied on the environments and the approach outperformed no ordering and random ordering methods, when cost-effectiveness was evaluated [11].

##### C. Exploiting parallelism in test execution

An approach for derivation of execution plan that will take into account parallel execution, while having in mind “resource availability and test case dependencies” has resulted in reduction of test execution time. However, it is more suited for automated test execution. For manual execution other things have to be taken into account e.g. a tester’s domain knowledge [16].

Parallel test execution service framework was designed in [18] with minimal application specific customization needs. The average execution time was reduced by 66.67% using four worker sandboxes and was concluded that it could be further reduced by increasing the number of workers with higher resource allocation [18].

##### D. Redundancy and similarity of steps in test execution

Although evolution of test suits leads to accumulation of redundant tests, no one has explored accumulation of similar tests, since these too increase execution time. So, approach in [17] identifies tests that can be merged, thus creating a test that covers all application states, while eliminating duplicated steps and does not affect detection of faults. The approach was implemented in practice and

decreased the number of test steps by 29%, while execution time was decreased by 39% [17].

Test cases can have a significant number of repeating test steps, which increase load during execution, so test case synthesis was proposed in [20]. After elimination of redundant steps, the remaining steps are rearranged into a new set of test cases. Optimized step execution order was determined by an algorithm. Test load was reduced by 15% and reordering of steps decreased the load by additional 3% [20].

#### E. GPU in acceleration of test execution

Acceleration of the execution of functional tests of finite state machines was achieved by running tests parallelly on GPU threads [21]. 13 subject FSMs from the network intrusion detection and signal processing domains were used for approach evaluation. Maximum acceleration of 12× was achieved when compared to a 16-core CPU. GPUs generally achieved higher acceleration for FSMs with long test sequences. Using a dense FSM layout resulted in higher GPU acceleration, particularly for FSMs represented as dense matrices. When test layout was considered, with-offsets layout tended to perform better for FSMs with shorter test sequences, while padded-transposed layout was more suitable for FSMs with long sequences. Prioritizing test sequences before execution on the GPU improved acceleration [21].

When executing program with multiple data sets, parallelism can be harvested using GPU, so in [22] approach was developed which “simultaneously executes the program with one test case per GPU thread”. Test executions were accelerated by a factor of 27 when compared to single-core executions on CPU.

#### F. Other topics in test execution

As opposed to improving instruction locality over single program, [12] was focused on instruction locality across the execution of test cases and test cases were also permuted to enable subsequent runs to share as many instructions as possible.

SPEQTRA was developed for mining executions traces of tests (failed and passed) to find faulty class and rank classes based on their probability of being faulty. SPEQTRA surpassed AMPLE in class ranking [13].

Reducing time and cost of execution is important, so in [14], overlap-aware greedy approach for test case selection and prioritization was applied and it achieved better results than standard greedy (by factor of 1.5). Also, up to 97% of all line coverage lines were unique redundant, which carries potential for data size reduction [14].

TAO was developed for optimization of test execution in large projects. This prototype decreased execution time by over 96% on two large Python projects [15].

Predicting defects from development process and final product, although extensively researched, is not possible when tests are conducted for a contractor. So, the only source is internal testing process, which has not been investigated. Therefore, in [19] testing metrics “from

acceptance test execution logs of a large scale software application” were extracted and measuring was conducted during 15 weeks and defects were predicted on a weekly basis. However, the results cannot be used to draw strong conclusions due to several issues (number of instances, data accuracy) [19].

Choosing test execution techniques is done without any formal procedure, which could generate higher costs. Therefore, a cost model for estimation of economic impact of test execution models was proposed and applied in practice in [23].

## V. CONCLUSION

Software testing has gained importance over time, especially with the proliferation of software in every sphere of our lives. With the increasing size and complexity of software, testing has become more demanding regarding necessary resources. Testing is considered a separate process with four main phases: planning, design, execution, and evaluation.

The main objective of this paper was to investigate the latest achievements in the execution phase of software testing. The authors selected four scientific repositories for conducting the review. After applying inclusion and exclusion criteria, the search resulted in 17 papers.

Parallelism, redundancy and similarity of test steps were exploited to speed up execution. Test types were further developed, while acceleration of test execution using GPU was performed in two papers.

Despite presenting state-of-the-art improvements in test execution phase, there are several threats to validity. Firstly, as recommended by Kitchenham (2004), the authors selected only four repositories for conducting this review. If the repository selection had been different or if the number of selected repositories had been higher, this SLR could have potentially offered more diversity regarding test execution improvement recommendations and provide a deeper insight into how academia is investing their resources in improving execution phase of testing.

Secondly, the authors decided to exclude gray literature, which could have potentially provided useful recommendations related to test execution improvement.

Finally, only one string was applied in all chosen repositories. If variations such as “software test execution” or “test execution improvement” were applied, it is possible that the SLR would have offered additional valuable information regarding test execution improvement.

To improve software testing process, it is not enough to just focus on the execution phase. Other testing phases have to be taken into consideration as well. In the future, the authors will explore how the rest of the testing phases could be improved.

## REFERENCES

- [1] ISO/IEC/IEEE, “ISO/IEC/IEEE 29119 Software Testing,” 2013. Retrieved from <http://www.softwaretestingstandard.org>

- [2] TMMi Foundation, "Test Maturity Model integration (TMMi) v3.1.," Dublin: TMMi Foundation, 2010
- [3] T. Koomen, L. van der Aalst, B. Broekman, and M. Vroon, "TMap Next," Hertogenbosch: UTN Publishers, 2006
- [4] W. Afzal, and R. Torkar, "Incorporating Metrics in an Organizational Test Strategy," Software Testing Verification and Validation Workshop, ICSTW '08, Washington: IEEE Computer Society, 2008, pp. 304-315.
- [5] B. Kitchenham, "Procedures for Performing Systematic Reviews", Joint Technical Report, Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd. (0400011T.1), 2004.
- [6] A. A. Khan, J. Keung, M. Niazi, S. Hussain, and H. Zhang, "Systematic Literature Reviews of Software Process Improvement: A Tertiary Study", EuroSPI 2017, CCIS 748, 2017, pp. 177-190.
- [7] P. Pedemonte, J. Mahmud, and T. Lau, "Towards automatic functional test execution," Proceedings of the 2012 ACM international conference on Intelligent User Interfaces (IUI '12). Association for Computing Machinery, New York, NY, USA, 2012, pp. 227-236.
- [8] J. G. Siqueira and M. M. V. de Paula, "IPEAD A/B Test Execution Framework," Proceedings of the XIV Brazilian Symposium on Information Systems (SBSI'18). Association for Computing Machinery, New York, NY, USA, Article 14, 2018, pp.1-8.
- [9] C. Augusto, "Efficient test execution in End to End testing: Resource optimization in End to End testing through a smart resource characterization and orchestration," 2020 IEEE/ACM 42nd International Conference on Software Engineering: Companion Proceedings (ICSE-Companion), 2020, pp. 152-154.
- [10] M. Guarnieri, P. Tsankov, T. Buchs, M. T. Dashti, and D. Basin, "Test execution checkpointing for web applications," Proceedings of the 26th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2017). Association for Computing Machinery, New York, NY, USA, 2017, pp. 203-214.
- [11] J.-H. Kwon, I.-Y. Ko, and G. Rothermel, "Prioritizing Browser Environments for Web Application Test Execution," Proceedings of the 40th International Conference on Software Engineering (ICSE '18). Association for Computing Machinery, New York, NY, USA, 2018, pp. 468-479.
- [12] P. Stratis, "Improving Test Execution Time with Improved Cache Locality," Proceedings of the 39th International Conference on Software Engineering Companion (ICSE-C '17). IEEE Press, 2017, pp. 82-84.
- [13] G. Laghari, A. Murgia, and S. Demeyer, "Localising Faults in Test Execution Traces," Proceedings of the 14th International Workshop on Principles of Software Evolution (IWPSE 2015). Association for Computing Machinery, New York, NY, USA, 2015, pp. 1-8.
- [14] T. Bach, A. Andrzejak, and R. Pannemans, "Coverage-Based Reduction of Test Execution Time: Lessons from a Very Large Industrial Project," 2017 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW), 2017, pp. 3-12.
- [15] T. Kim, R. Chandra, and N. Zeldovich, "Optimizing unit test execution in large software programs using dependency analysis," Proceedings of the 4th Asia-Pacific Workshop on Systems (APSys '13). Association for Computing Machinery, New York, NY, USA, Article 19, 2013, pp. 1-6.
- [16] S. S. Chakraborty and V. Shah, "Towards an approach and framework for test-execution plan derivation," 2011 26th IEEE/ACM International Conference on Automated Software Engineering (ASE 2011), 2011, pp. 488-491.
- [17] P. Devaki, S. Thummalapenta, N. Singhanian, and S. Sinha, "Efficient and Flexible GUI Test Execution via Test Merging," Proceedings of the 2013 International Symposium on Software Testing and Analysis (ISSTA 2013). Association for Computing Machinery, New York, NY, USA, 2013, pp. 34-44.
- [18] M. Rahman, Z. Chen, and J. Gao, "A Service Framework for Parallel Test Execution on a Developer's Local Development Workstation," 2015 IEEE Symposium on Service-Oriented System Engineering, 2015, pp. 153-160.
- [19] A. Tosun, O. Turkgulu, D. Razon, H. Y. Aydemir, and A. Gureller, "Predicting defects using test execution logs in an industrial setting," 2017 IEEE/ACM 39th International Conference on Software Engineering Companion (ICSE-C), 2017, pp. 294-296.
- [20] B. Walter, M. Schilling, M. Piechotta, and S. Rudolph, "Improving Test Execution Efficiency Through Clustering and Reordering of Independent Test Steps," 2018 IEEE 11th International Conference on Software Testing, Verification and Validation (ICST), 2018, pp. 363-373.
- [21] V. Yaneva, A. Kapoor, A. Rajan and C. Dubach, "Accelerated Finite State Machine Test Execution Using GPUs," 2018 25th Asia-Pacific Software Engineering Conference (APSEC), 2018, pp. 109-118.
- [22] A. Rajan, S. Sharma, P. Schrammel, and D. Kroening, "Accelerated Test Execution using GPUs," Proceedings of the 29th ACM/IEEE international conference on Automated software engineering (ASE '14). Association for Computing Machinery, New York, NY, USA, 2014, pp. 97-102.
- [23] B. Hauptmann, M. Junker, S. Eder, C. Amann, and R. Vaas, "An Expert-Based Cost Estimation Model for System Test Execution," Proceedings of the 2014 International Conference on Software and System Process (ICSSP 2014). Association for Computing Machinery, New York, NY, USA, 2014, pp. 159-163.
- [24] D. Ghosh and J. Singh, "Spectrum-based multi-fault localization using Chaotic Genetic Algorithm", Information and Software Technology, vol.133, 2021.
- [25] M. M. D. Shahabi, S. P. Badii, S. E. Beheshtian, R. Akbari and S. M. R. Moosavi, "On the performance of EvoPSO: A PSO based algorithm for test data generation in EvoSuite," 2017 2nd Conference on Swarm Intelligence and Evolutionary Computation (CSIEC), Kerman, 2017, pp. 129-134.
- [26] H. L. P. Raj and K. Chandrasekaran, "NEAT Algorithm for Testsuite generation in Automated Software Testing," 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 2018, pp. 2361-2368, doi: 10.1109/SSCI.2018.8628668.

# Comparison of Dart and JavaScript Programming Languages

I. Vecštejn, V. Ognjenović, E. Brtka, T. Milić, M. Gaborov

University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia

igor.vecstejn@gmail.com, visnjaognjenovic@gmail.com, eleonorabrtka@gmail.com, tamaraamilic021@gmail.com, majagaborov26@gmail.com

**Abstract** – This paper elaborates comparison of a new programming language Dart that has been developed by Google and JavaScript. Dart has great potential to dominate in future. Both programming languages will be displayed and defined with their history features, frameworks, properties, functions of client server request. Focus is going to be on investigating key comparison points, comparison of main characteristics, concepts of programming languages Dart and JavaScript.

## I. INTRODUCTION

Dart is an open-source, structured and flexible programming language developed by Google. It is primarily used for browser-based Web applications, but now with new versions, Dart is being used across a wide variety of compilation targets (Web applications, Mobile Applications, Desktop applications).

Google cares a lot about helping to make the Web great. They are writing a lot of Web Applications, many of them are refined, such as: Gmail, Google Calendar, etc. They want Web Applications to load quickly, run smoothly and present engaging and fun experiences to users [1]. Behind Dart are Lars Bak and Kasper Lund, two of the authors of the V8 JavaScript engine for Google Chrome. For that reason the performance and efficiency of the language were two factors that were very important at the moment of developing this language. Despite what it seems, Dart was not developed to replace JavaScript. It was developed to offer an additional, modern option for Application development with better performances and for big projects in which the maintenance process is complicated [2].

Dart is a general purpose programming language. It is a new language in program C tradition, thereby to all programming languages that inherit from it. It was designed with ease, familiarity to the vast majority of programmers and scalability in mind. It is fully object-oriented, class-based, optionally typed and single-threaded [2,3].

As it is said, Dart is an open-source, structured language used to create complex, browser-based Web Applications. Applications that were created in Dart could be run by using Web browser that directly support Dart code or by compiling Dart code to JavaScript. Dart has a familiar syntax, and it's class-based, it has a concurrency model called isolates that allows parallel execution [3,4]. In addition to running Dart code in Web

browser and converting it to JavaScript, user can also run Dart code on the command line, hosted in the Dart VM(Virtual Machine), allowing both the client and the server parts of users application to be coded in the same programming language. Dart has incorporated Future class into its SDK (Software Developer Kit). This class allows users to isolate in an object-oriented way asynchronous return results as built-in function. Future API makes user's code more cleaner, readable.

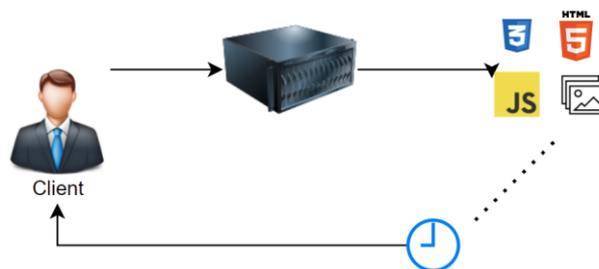


Figure 1. Standard Client-Server Request Scheme

The programming paradigm has changed a lot over past few years. The server was in charge of doing everything. On Figure 1. when user wants to visit Web Page, server makes database queries, makes access validation, builds the HTML document, loads external resources as CSS, JavaScript scripts, images, videos, etc. This is happening in Web Applications developed with PHP, Python, JS. Every link that user visits, it makes all processes run again and displays an HTML document rendered by the client browser. This makes every request very expensive with each time the client wants to view some resource on the server[2]. The emergence of new mobile devices and social networks, there are a millions of people trying to attempt all those requests. It is not possible for server to meet all the demands of users because every request is using a lot of resources and time. New HTML, CSS, JS versions makes it possible. Every time user's browsers access through the Internet, to a Web service, from login screen to social network profile, etc., server sends to browser few scripts that use browser. On Figure 2. its shown new asynchronous Client-Server request.

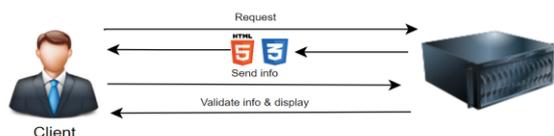


Figure 2. New Client-Server Requests

JavaScript was introduced in 1995 as a way to add programs to Web pages in the Netscape Navigator browser. The language has since been adopted by all other major graphical Web browsers. It has made modern Web applications possible – applications with which user can interact directly without doing a page reload for every single action. After its adoption outside of Netscape, a standard document was written to describe the way the JavaScript language should work so that the various pieces of software that claimed to support JS were actually talking about same language. This is called ECMAScript standard. In practice, the terms ECMAScript and JavaScript can be used for the same language[5]. The overwhelming majority of Websites use JavaScript, and all modern Web browsers on different devices such as desktop PC, tablets and phones include JavaScript interpreters, making JavaScript the most-deployed programming language in history[6]. JavaScript is high-level, dynamic, interpreted programming language that is well-suited to object-oriented and functional programming styles. JavaScript's variables are untyped. Its syntax is based on Java, but Java and JavaScript are unrelated. To be useful, every language must have a platform, or standard library for performing things like basic input and output. The core JavaScript language defines minimal API for working with numbers, text, arrays, etc. But does not include any input or output function. The main host environment for JavaScript was a Web browser and this is still most common execution for JavaScript code. Web browser allows code that is written in JavaScript to obtain input from the user's peripherals by making HTTP request, it also allows JavaScript code to display output to the user with HTML and CSS[6].

Since 2010, another environment has been available for JavaScript code. Instead of constraining JavaScript to work with the APIs that is provided by browser, Node JS gave JavaScript access to the entire OS, allowing JavaScript programs to read and write files, send and receive data over network.

## II. DART AND JAVASCRIPT: COMPARISON

Google is a real time interactive system application dealing with search, electronic mail, translation, play images, drive and many other applications. Hence, the company is in a constant search to develop programming languages that connect all of these and also future applications. Dart programming language has come to meet this need [7]. Dart is fairly new programming language for most of the developers that are outside of Google. Google has put a lot of effort into documenting the Dart programming language, it is still hard for developers to find solutions for specific problems.

JavaScript is on the height of its popularity, we all know that JavaScript is in the fields of cross-platform

mobile application development and server-side development. Node.js framework can be used for both frontend and backend development, thus JavaScript became highly popular among Web developers. The rise of React Native by Facebook, also mobile application developers moved towards JavaScript. This made JavaScript the most popular programming language to this day[8].

As an alternative to JavaScript or framework, React Native, Google has created framework for cross-platform mobile application development. Flutter is using Dart programming language. In contrast to React Native's applications, applications that are previously built with Flutter framework look much more native and slick. While writing Dart code, code is simultaneously compiling and executing in background. With this, user is getting real time results. JavaScript has been in industry for so long and its stable programming language.

JavaScript is very easy to use and has a lot of libraries and frameworks available online so developers can use existing code to develop their applications faster.

Table 1 shows comparison of both programming languages on different criteria and comparison points.

TABLE 1. COMPARISON OF DART AND JAVASCRIPT

Comparison points	Dart	JavaScript
Popularity	Before Google announced Flutter, Dart was nowhere to be found. Since the announcement of Flutter in Google, Dart has drastic attention among mobile developers. Big companies like Google, Alibaba adopted Flutter. The popularity has raised but it is not popular as JavaScript.	At very same moment, JavaScript is everywhere. There is no device in world that does not use JavaScript. There are many companies that are using JavaScript frameworks for developing applications. Companies like eBay, Slack, Instagram, Reddit use JavaScript.
Performance	Dart does not rely on any type of bridge communication, instead it's graphic core Skia, better known as 2D rendering engine that was written in C++, compiles user interface (UI) to cose as the user interacts with the application.	JavaScript's architecture requires a bridge that enables interaction between JavaScript and components of user's interface and some of the device's elements such as camera, sensors, bluetooth. Because of this bridge communication, JavaScript is slower than Dart
Productivity	Dart has awesome documentation but developers can get confused with Dart language features. Dart's syntax is very similar to Java's so developers who had programming previously with Java can pick syntax easily. There is a comparatively smaller community for Dart and Flutter.	Because JavaScript has many frameworks, it boosts developers productivity by offering countless libraries, packages and frameworks . It is easier for new developers to learn JavaScript because its faster, lightweight and dynamic language.
Learning Curve	Programming Language Dart can be challenging for begginers since it is not widely and commonly used programming language. There are only few courses and books that are available on Internet.	Learning JavaScript is not easy task. After learning fundamentals concepts in programming, JavaScript can be really easy for learning. There are enormous number of courses, tutorials and books on Internet.
Frontend vs Backend development	Dart is currently used with Flutter for developing frontend cross-platform mobile applications. It can also be used for developing Web applications. Right know Dart is not being used for backend development	JavaScript was in beginning used for frontend development with combination with HTML and CSS. Since the appearance of Node.js framework, JavaScript is now used even in server side and backend development
Ease of use	In terms of coding and syntax, Dart has similar syntax to Java, so developers with OOP knowledge can use Dart easily.	JavaScript is mature language, it is easy to use and understand, it is stable with enormous frameworks that are available for everyone.
Web vs Mobile development	Dart can be used for developing both Web and mobile applications. Dart became very popular with combination with Flutter for developing mobile applications. Dart can also develop Web applications but it is rarely used for that purpose.	JavaScript has dominated Web and mobile application development with different frameworks. With releasment of React and React Native from Facebook, JavaScript became glue for developing Web and mobile applications for business. Frameworks that are also being used for Web and mobile development are Angular, Vue.js, etc.
Safe-type	Dart allows strict prototyping. Since it is a compiled programming language, programming errors are discovered during compiled process. With this information, Dart is more safe-type than JavaScript	JavaScript is interpreted programming language that supports duck and dynamic typing. JavaScript is not safe-type since user can write any type of code. Errors can only be found during runtime

Debugging	In Dart, debugging is easier while using tools for support that enabled in Android Studio and in VS Code. One of many tools is Inspector. He checks visual layer of the application and questions why application is the way it looks.	In JavaScript, debugging is problematic, especially when a specific error originates in some part of user's application. Bug can be on side of programming language JavaScript, in JavaScript's framework or from third-party libraries. One of the tools that can help with errors is Flipper. Flipper can visualize, inspect and control applications from desktop interface.
-----------	--	---

### III. CONCLUSION

In comparison of Dart and JavaScript programming languages, both are great choices for developers for many reasons. Dart is a relatively new programming language compared to JavaScript. In this paper, these two programming languages were compared through popularity, performance, productivity, learning curve, frontend vs backend development, ease of use, Web vs Mobile development, safe-type and debugging. Right now, JavaScript is much more popular than Dart, but Dart's popularity is rasing with every month. In comparison point, performance, dart is now much faster since it is not using bridge that is using JavaScript, but instead it is using Skia, graphic engine. Both programming langugages are challenging for beginners. For Dart, there is much less books, tutorials and courses where can people learn to program and design from. In other hand, JavaScript has a lot of books, courses, tutorials on Internet. Many references, literatures, videos can boost people to learn this language faster. JavaScript is very popular for developing Mobile applications, Web applications, etc., but Dart, right now, is mostly meant for mobile development. With knowledge of OOP, both programming languges can facilitate learning. It has interesting features such as building complex applications using modern paradigms, has great support from Google and it has big potential to be competitor for JavaScript. Darts helps programmers to find easily mistakes they have made in code. JavaScript is going to be top programming language for this kind of development for long time.

### ACKNOWLEDGEMENT

This research is financially supported by Ministry of Education and Science of the Republic of Serbia under the project number TR32044 „The development of software tools for business process analysis and improvement“, 2011-2021.

### REFERENCES

- [1] Kathy Walarth, Seth Ladd, „Dart Up and Running“, 2013
- [2] Moises Belchin, Patricia Juberias, „Web Programming with Dart“, 2014
- [3] Gilad Bracha, „The Dart programming language“, 2015
- [4] Chris Buckett, „Dart in action“, 2013
- [5] Marijn Haverbeke, „Eloquent JavaScript A Modern Introduction to Programming“, 2018
- [6] David Flanagan, „JavaScript The Definitive Guide Master The Worlds Most Used Programming Language“, 2020
- [7] Afaf Mirghani Hassan, “Java and Dart programming languages: Conceptual comparison”, Indonesian Journal of Electrical Engineering and Computer Science”, 2019, pp. 845-849
- [8] Website: <https://enquero.com/javascript-vs-dart/>

# Mitigating Covid-19 Impact on Small Businesses and Startups Using Digital Technologies

Z. Kotevski and A. Shijakova Kotevski

St. Kliment Ohridski University, Bitola, North Macedonia  
zoran.kotevski@uklo.edu.mk, aneta.sijakova@uklo.edu.mk

**Abstract** – The COVID-19 pandemic struck the world with fierceness unprecedented in more than a hundred years. Almost immediately after the outbreak, digital technologies were employed to directly fight the pandemic in manner of tracking the spread of the disease, monitoring the infected people, as well as managing the vaccination process. But, besides the primary global health concerns, there was a secondary impact, i.e. impact on the economy where small businesses and startups appeared to be the most vulnerable. Having this in mind, in this paper, we explore the possibilities of using digital technologies to mitigate the pandemic's devastating impact on the most fragile economic entities. We compiled largely available digital systems that are immediately usable, affordable and appear to be quite adequate to help small businesses and startups survive the pandemic and even boost their income.

## I. INTRODUCTION

The sudden attack of SARS-CoV-2 virus at the end of 2019 took the world by surprise, and even the most advanced healthcare systems were unprepared for such highly infectious disease. It swiftly circled around the globe severely affecting public health. As a primary response, the medical institutions promptly reacted to establish medical procedures for treatment of COVID-19. But, knowing the reproduction number of the new infection, it was maybe of greater importance to fight the spread of the disease. For this purpose the main role was given to digital technologies since electronic devices, tools and systems can quickly generate, store and process data. In this manner, multitude of tools for monitoring, detection and prevention of COVID-19 were developed in order to remediate the outbreak. Some of the research that discusses such implementations is presented in the following lines.

Ting, Carin, Dzau and Wong [1] in early 2020 explored the potential application of four inter-related digital technologies, such as Internet of Things (IoT), Big-Data Analytics, Artificial Intelligence (AI) and Blockchain, to augment public-health strategies to prevent the spread of COVID-19 with better detection, monitoring and surveillance of existing and potential COVID-19 patients.

Kumar, Kumar Gupta and Srivastava [2] discuss the various aspects of modern technologies used to fight against COVID-19 crisis at different scales, including medical image processing, disease tracking, prediction outcomes, computational biology and medicines. Most

notably, they speak about using smartphones and Internet to aid in detection and mitigation of the disease impact, the recent advances in technology that can help in effective tracking and quarantine of COVID-19 cases and potential uses of AI in managing the pandemic.

Whitelaw, Mamas, Topol and Van Spall [3] share their viewpoint on the importance of the application of digital technologies in pandemic management and response, i.e. for early surveillance, testing, contact tracing, and strict quarantine. They also emphasize that countries that have adopted these technologies maintained low COVID-19 per-capita mortality rates.

He, Zhang and Li [4] examine the emerging technologies used to mitigate the threats of COVID-19 and provide insights and suggestions into how information systems and technology scholars can help fight the pandemic.

The aforementioned papers tackle the implementation of digital technologies to fight the direct health threats caused by the COVID-19 pandemic. But there was a secondary impact as well, i.e. impact on the economy, where the most vulnerable entities appeared to be small businesses and startups.

For example, Facebook's report from June 2020 [5] surveyed 86,000 owners, managers, and workers in companies across the US with fewer than 500 employees. They summarize that 31% of the small and medium businesses have shut down in the first 3 months of the outbreak in the country. Furthermore, the situation was particularly severe for personal business, among which even 52% reported shutting down.

Bartik et al. [6] in July 2020 published a similar research, reporting that 43% of the businesses in the US stopped working, while the most affected were food, art and sport industries, reducing about 50% of the employees.

In August 2020, the MIT Technology Review [7] published a research that explores the influence of the pandemic on the organizational plans and expectations of businesses during such outbreak. The research was based on a survey of 372 business leaders. Among the most notable conclusions were that 62% of the respondents expected 2020 company revenues to decrease, 25% of them expected decrease of more than 25%, but 90% of them expected accelerated implementation of new technologies in their businesses.

Another research on a similar subject was published in October 2020 by McKinsey & Company [8] that takes a rather different approach. They surveyed 899 executive directors and managers that work in different industries, regions and company sizes. Their research casts light on the acceleration of the planned implementation of digital technologies in their businesses in order to quickly respond to the market demand during such difficult times as the COVID-19 pandemic. The research revealed that digital technologies were implemented 43 times faster than planned, the organization of offerings of online products and services was accelerated 27 times with 20% immediate increase in actual online offers, and the implementation of digital data security increased 19 times faster than expected.

Regarding the startups specifically, Calvino, Criscuolo and Verlhac [9] from the Organization for Economic Co-operation and Development (OECD) in May 2020 published a report tackling the employment growth for startups, concluding that the employments are doubled for startups that implement digital technologies compared to ones that don't. They also report that business formation during COVID-19 declined significantly and has been as severe as 70% in April 2020 in Portugal compared to the same month of prior year, and 46%, 54%, and 58% in Hungary, France, and Turkey, respectively.

De Cuyper, Kucukkeles and Reuben [10] also published a research in June 2020, at the World Economic Forum, tackling pandemic induced problems of startups where they conclude that even more than 70% of startups have had to terminate full-time employee contracts since the start of the COVID-19 pandemic.

This secondary impact of the pandemic, i.e. the devastating consequences to the economy, appeared to be as important to mitigate as the threat to the public health, thus the world needed to react quickly and intensively to alleviate the economy shock. As we mentioned previously, the most affected entities by the COVID-19 outbreak are small businesses and startups. We have also disclosed that digital technologies were immediately applied to fight the spread of the COVID-19 disease. Hence, in this paper we are taking an approach to explore the possibilities to mitigate the devastating impact on small businesses and startups using digital technologies. We compile variety of digital systems and solutions that are available to the public, and seem to be especially useful to these vulnerable entities in such difficult times for economic prosperity.

The rest of this paper is organized as follows. In section 2 we present the digital technologies that are readily available for use, and third section concludes the paper with a brief summary of contribution.

## II. DIGITAL TECHNOLOGY TRENDS IN BUSINESS

Everyday more and more businesses implement digital technologies. As elaborated in [9], employment growth is twice as high in business entities that use digital technologies compared to the ones that decided to be non-digital. The COVID-19 pandemic further accelerated the digitization process, which turned out to be quite beneficial to businesses, especially the small entities and

startups. In this chapter we compile several digital technology trends that can help small businesses and startups survive the pandemic and even increase their revenues. The following suggestions do not target a specific industry branch. They are a compilation of available and affordable techniques, trends, applications and systems of digital (Internet) technology that can be easily implemented by vast number of business entities. We organize the implementation of digital technologies in four categories, i.e. (i) teaching courses online, (ii) building e-commerce strategy, (iii) strengthening online presence and (iv) building an online marketing strategy.

### A. Teaching courses online

For educational businesses or startups this is a vital step. As people mainly restrain to attend social or other gatherings in Covid-19 pandemic, offering the services online is the crucial step to realize. Even though the teaching as a service mostly implies in person, online teaching is an opportunity to offer educational services to interested parties in their homes and offered to much greater audiences. The possibilities to create online teaching courses include three distinct ways for delivery:

- Online course delivery through web platforms that offer online learning environment;
- Online course delivery through web platforms that offer course delivery by download;
- Online course delivery on own course platform.

Regarding the web platforms for delivery of online courses, some of the most popular ones, as well as most affordable solutions, are presented in the Table I.

TABLE I. COURSE DELIVERY PLATFORMS THAT OFFER ONLINE LEARNING ENVIRONMENT

Platform	Basic/Starter plan
Udemy.com	- No monthly subscriptions; - Undisclosed details before account creation, but usually in the range of 50-75% of the course sales.
Stillshare.com	- 8.25\$/month; - Undisclosed details before account creation.
Teachable.com	- 29\$/month + 5% per transaction; - Unlimited students, videos, courses and hosting.
LearnWorlds.com	- 24\$/month + 5\$ per transaction; - Unlimited courses.
Teachery.co	- 49\$/month; - Unlimited students and courses.
Kartra.com	- 99\$/month; - 50 videos, 20 products.
Podia.com	- 39\$/month.
Systeme.io	- 27\$/month.
Kajabi.com	- 119\$/month.
Pathwright.com	- 99\$/month; - Unlimited courses, 1000 members
Xperienify.com	- 49\$/month; - unlimited courses, unlimited students
Thinkific.com	- Free plan with limited courses; - Unlimited students, Quizzes & surveys, Content hosting, Instant access to your funds.

Among the many platforms for selling online courses, thinkific.com appears as most suitable to begin with. Not only that it is the only platform that offers free option that does not expire, the options included appear even better than many paid plans of other platforms. Some of the more interesting additional features of the free plan include: instant access to funds, full e-commerce solution, courses with any content type, unlimited number of students and detailed reporting & analytics.

To summarize, thinkific.com free option offers complete website builder with content creation tools and included e-commerce solution than can integrate 3rd party payment processors. The online entity will receive a free subdomain and with a paid option it can be replaced with custom domain. The only, partly-undisclosed, feature of the free account is the “limited courses” feature that does not explain what exactly the limitation applies to.

On the other hand, there is always the possibility to offer courses as digital downloads, even though it is not the most appropriate way for courses delivery. Regarding teaching courses, these platforms lack some of the essential features that dedicated course platforms offer, such as video preview of the course, student discussions and student surveys, to name a few.

However, for the sake of completeness, in the the following Table II we present the most popular platforms that offer sales of digital items by download.

TABLE II. COURSE DELIVERY PLATFORMS THAT OFFER COURSE DELIVERY BY DOWNLOAD

Platform	Basic/Starter plan
Shopify.com	- 29\$/month; - Additional 2.9% + 0.3\$ per transaction.
Etsy.com	- 0.2\$/item listing fee for 4 months; - 5% per transaction;
Sellfy.com	- 19\$/month; - 0\$ per transaction; - Unlimited products.
Sellwire.net	- 9\$/month; - Manage up to 10 files; - Unlimited file versions; - Unlimited orders.
Simplegoods.co	- 0\$/month; - 3% per transaction; - Up to 3 products; - 4BG per file; - Works only with Stripe accounts!
Podia.com	- 39\$/month; - Includes custom website
Getdpd.com	- 10\$/month; - 1GB storage space; - 20 products; - Unlimited bandwidth; - Unlimited sales.
Fetchapp.com	- 10\$/month; - 2GB storage space; - Unmetered bandwidth.
Sendowl.com	- 15\$/month; - 30 items; - 3GB storage space.
Payloadz.com	- 29\$/month; - 2.9% + 0.29\$ per transaction

Platform	Basic/Starter plan
Selz.com	- 26\$/month; - 2.9% + 0.3\$ per transaction; - Unlimited products; - Custom online store.

It appears that the best options of this type are Etsy and Simplegoods, but since Simplegoods works with Stripe only (which is not yet supported in many countries), it is not quite suitable at this moment. The next best options are Sellwire and Fetchapp. In comparison to the previous dedicated course delivery platforms, Thinkific appears to be a better solution for delivering online courses.

In essence, delivering any services, including online courses, via own custom platform is always the better solution regarding functionalities and customization. But, these solutions are expensive. From building to hosting, to processing payments etc. As such, these solutions are not recommended as starter solutions and should be employed only after a startup or small business has already built a sustainable community of users and has achieved a steady monthly income.

#### B. Building e-commerce strategy

E-commerce is the new business norm. It is the form of activity that many businesses cannot survive without. Building an e-commerce strategy is extremely valuable to any entity, and e-commerce solutions are readily available for any business today, regardless of its size or type. There are three different ways to incorporate e-commerce solution and we present them in ascending order of their investment requirements:

- Using well established online marketplaces;
- Using e-commerce platforms to build online store;
- Building a custom online store.

The easiest way to offer products online is by utilizing online marketplaces. According to UK Tech News [11] marketplaces in 2019 contributed 1.7 trillion USD to the world economy and they forecast that this financial impact will grow significantly in the next five years and it is expected to exceed 7 trillion USD in the next five years. This rise is driven by more and more companies embracing marketplaces as their choice to facilitate online sales, which emerged as a perfectly viable solution to expand sales quite substantially, in comparison to the traditional local sales.

Choosing the right online marketplace to sell own products is a challenging task. To be as profitable as possible on these platforms, first we need to identify the best online marketplaces for the brand and products. On the other hand, each online marketplace has its own unique requirements such as product categories, listing fees, subscription fees, commission rates etc. Furthermore, we need to analyze and understand the different strategies for selling on e-commerce marketplaces, to ensure which ones hold the most promise for our products and what are the first steps to undertake to a stable beginning.

In Table III we have compiled a list of nine online marketplaces that hold promise as being among the best solutions for small businesses and startups, among which Bonanza's offer appears as the most appealing.

TABLE III. ONLINE MARKETPLACES

Marketplace	Commission
Etsy.com	5% per sale and 4% + 0.30 EUR per transaction (EU)
Happeno.com	10% per sale and 3.4% + 0.35 EUR per transaction
Bonanza.com	Zero fees for direct sales made through the free referral link. OR 3.5% per sale (min. 0.50 USD)
Ebid.net	5% per sale
Ecrater.com	Zero fees for buyers you bring to the store, OR 2.9% per sale
Asos.com	20% per sale
Poshmark.com	20% per sale
Depop.com	10% per sale

On the other hand, e-commerce platforms offer tools to build online stores. More specifically these platforms offer Software as a Service (SaaS) for quick building of online shops. The majority of such platforms will host the newly built store on their servers as well. In Table IV we present the most popular ones, accompanied with some relevant information.

TABLE IV. E-COMMERCE PLATFORMS

Platform	Basic features	Base pricing
Shopify.com	Online store with included e-commerce website and blog Unlimited products	29 USD / Month
Bigcommerce.com	Online store Unlimited products, storage and bandwidth	29,99 USD / Month
Ecwid.com	Online store 10 products (free) 100 products (\$15/m)	Free 15 USD / Month
Wix.com	Online store Unlimited products and bandwidth	17 USD / Month
Volusion.com	Online store 100 products Unlimited bandwidth	29 USD / Month
Weebly.com	Online store Free domain Unlimited storage	12 USD / month

E-commerce platforms compared to the previous online marketplaces offer e-commerce services that are somewhat more expensive, but using these platforms has some advantages, such as building a standalone store.

Custom e-commerce solutions are the most expensive options and are best suitable for larger companies.

To summarize, online marketplaces presented in Table III are chosen among many other primarily because of

their commission rates. The majority of them charge less than 10% commission, which is an acceptable rate since no other costs, such as for e-commerce website development and hosting, are required to launch and online store. Since different online marketplaces usually have different customers, a good practice is to be present in more than one online marketplace, to ensure higher visitor rates and sales conversions. Using e-commerce platforms is mostly recommended for businesses that are not new and already have some steady revenues. They are a good option for boosting sales or expanding the store in general. Custom e-commerce solution should be employed when the business has grown enough to support the expenses that come along with it, but it is certainly recommended when the time to build such solution is right.

### C. Strengthening online presence

Projecting an image of a prosperous business is crucial for its success. Thus, one of the main features that a business should possess is a corporate website. A website is not a means that would attract internet traffic by itself, but it is a vital part of the corporate image that describes the business quality. Having a website and online presence strategy allows marketing the business online. A website is also important because it helps to establish credibility as a business. Another important aspect is that the business can be adequately represented to the customers, and to build trust. A good website gives credibility to the business, but it also helps to give a positive impression that the company is bigger and highly successful. One of the great things about the Internet is that the size of the company does not really matter. Therefore, maybe the website is more important for small businesses than to the big ones. Websites are always available and accessible 24 hours a day. As a result, the existing and potential customers can visit the website and acquire information about the new and upcoming products or services whenever it is convenient for them.

The secondary benefit that should come along with building a business website is the possibility to build a database of email contacts for the users who wish to subscribe to newsletters or other periodicals. An email list enables the business to engage with current and potential customers on a daily, weekly or monthly basis. Some of the tactics to gather new users is to create a gated content that would be visible only to registered users. Additionally Call-To-Action (CTA) can be used on the website and social media pages to promote the email newsletter.

Another good practice is to register the business with web search engines such as Google, Yahoo and Bing. The registration can be easily performed, while registering the business with search engines allows it to be found more easily on searches and it shows up on Google Maps. All that needs to be done is to fill out the registration form and verify the business through their confirmation process, which can be done either with a phone call or by standard mail.

Finally, maybe the most important option to raise online presence are social media. Social media started as a tool for getting exposed, but over a short time it has grown

to become a necessary investment in time for every business to make. Today it is easy to achieve placing ads and offers on a Facebook page or have a direct channel with the customers on Twitter. Similarly, networking on LinkedIn, both at personal and company level, can be another way to help startup businesses promote. Besides standard social networking, a good practice is to employ a media channel of short videos (on YouTube or similar platforms) about the services offered that will act as a teaser advertisement. Depending on the business, other social media can be employed as well.

#### D. Building an online marketing strategy

There are multiple online tools for marketing support. Depending on their functions, online marketing tools can be classified into several groups, such as:

- E-mail marketing
- Social media marketing
- Search engine marketing
- Online marketplaces marketing

E-mail marketing tools are tools used by marketers to create, send, test, optimize, and report on their email campaigns. Many of these tools are offered in free option, while other, more feature full options are not much expensive. Some of the best e-mail marketing tools are presented in Table V.

TABLE V. E-MAIL MARKETING TOOLS

Tool	Basic features	Base pricing
Mailchimp.com	2000 contacts	Free
Mailerlite.com	1000 subscribers 12000 emails / month	Free
Sendinblue.com	Unlimited contacts 300 emails / day	Free
Convertkit.com	1000 subscribers	Free

Even though these tools offer adequate features in their free plans, they are not necessary and with a little bit of effort e-mail marketing can be easily built when the business has a website and offers subscription possibilities to the users.

Social media marketing is the use of social media platforms to connect with the audience or potential users to increase webstore traffic and increase sales. Advertising on social media comes with a certain costs. In Table VI we present the advertisement prices for the most popular social media platforms.

TABLE VI. SOCIAL MEDIA MARKETING TOOLS

Social media	Base pricing
Facebook.com	About 0.2 – 2.0 USD per click (depending on the industry)
Twitter.com	About 0.2 – 2.0 USD per click
LinkedIn.com	Min 2 USD per click
Instagram.com	About 0.5 – 1 USD per click

Search engine marketing is the use of web search engines to advertise products or services over Internet. In Table VII we present the advertisement prices for the most popular web search engines.

TABLE VII. SEARCH ENGINE MARKETING TOOLS

Search engine	Base pricing
Google.com	Average of 1 to 2 USD per click
Bing.com	
Yahoo.com	

However, we need to bear in mind the conversion rate, i.e. the percentage of clicks that have resulted in purchases. For example, the average conversion rate of Google Ads is about 3,75%, and for social networks is about 2,5%.

Similarly to search engine and social media marketing, online marketing and advertising can be performed directly in the marketplaces presented in Table III. Here we present in detail the marketing and advertising possibilities and expenses for two marketplaces (Etsy and Bonanza), as their conditions and charges appear to be the most suitable for a startup or small business.

Etsy Ads are based on an auction for ad space on Etsy's properties. Etsy enters "bids" for an ad space on behalf of the business based on a daily budget set by the administrator of the business. Thus, the actual costs incurred will vary depending on the number of participants in the auction and their bid amounts. A bid is the maximum amount of money the business will be charged if a buyer clicks on the Etsy Ad. Etsy's calculation of bids is based on information the business administrator provides, such as the daily budget or the listings the business wishes to include for Etsy Ads. Etsy sets the minimum daily budget and each business can set its daily maximum budget for Etsy Ads. To summarize, Etsy Ads prices are variable and therefore are undisclosed in general information. To see the current Ads prices the administrator must go to the Advertising Dashboard in the Shop Manager section of the business user account. The average conversion rate using Etsy Ads appears to be between 1% and 5%, but is varies largely among different industries. Therefore, to calculate whether Etsy Ads will yield more profit, a conversion rate of 2,5% should be considered.

Bonanza's advertising program functions quite differently compared to Etsy's, and appears to offer a program that is the most suitable to small businesses, especially for startup businesses. The main reason for this suitability is because Bonanza advertising is payable only after a sale is made, so there are no prior expenses for advertisements. Bonanza offers their marketing plans in four advertising levels, which consist of a range of fees and benefits:

- Basic level - 9% maximum fee per sale
- Standard level - 13% maximum fee per sale
- Superior level - 19% maximum fee per sale
- Elite level - 30% maximum fee per sale

Higher advertising levels unlock access to additional shopping channels and include a higher ad spend to drive more buyer traffic to the listings. When we combine the “Zero fees for direct sales made through the free referral link” then Bonanza emerges as the primary choice to open an online store.

### III. CONCLUSION

Besides the public health issues caused by the COVID-19 pandemic there was a devastating impact to the world economy as well, where the most vulnerable entities were small businesses and startups. To fight the pandemic, digital technologies were immediately deployed, not only to prevent the spread of the disease, but to help vulnerable businesses survive and even boost their revenues. We have compiled four groups of digital technologies, quite appropriate for supporting businesses, i.e. teaching courses online, building e-commerce strategy, strengthening online presence and building an online marketing strategy. We reviewed the most popular systems from each group and we presented the most adequate options for mitigating the impact of the COVID-19 pandemic on the small businesses and startups.

### REFERENCES

- [1] D. S. W. Ting, L. Carin, V. Dzau and T. Y. Wong, “Digital technology and COVID-19”. *Nat Med* 26, pp. 459–461, 2020.
- [2] A. Kumar, P. Kumar Gupta and A. Srivastava, “A review of modern technologies for tackling COVID-19 pandemic”, *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14 (4), pp. 569-573, 2020.
- [3] S. Whitelaw, M. A. Mamas, E. Topol, H. G. C. Van Spall, “Applications of digital technology in COVID-19 pandemic planning and response”, *The Lancet Digital Health*, vol. 2(8), pp. e435-e440, 2020.
- [4] W. He, Z. (J.) Zhang and W. Li, “Information technology solutions, challenges, and suggestions for tackling the COVID-19 pandemic”, *International Journal of Information Management*, vol. 57, 2021.
- [5] Facebook & Small Business Roundtable, “State of Small Business Report”, <https://about.fb.com/wp-content/uploads/2020/05/SMBReport.pdf>, 2020.
- [6] A. W. Bartik, M. Bertrand, Z. Cullen, E. L. Glaeser, M. Luca and C. Stanton, “The impact of COVID-19 on small business outcomes and expectations”, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117(30), pp. 17656-17666, 2020.
- [7] MIT Technology Review, “Amid the covid-19 pandemic, shifting business priorities ”, <https://www.technologyreview.com/2020/08/31/1007723/>, 2020.
- [8] L. LaBerge, C. O’Toole, J. Schneider and K. Smaje, “How COVID-19 has pushed companies over the technology tipping point—and transformed business forever”, *McKinsey & Company*, 2020.
- [9] F. Calvino, C. Criscuolo and R. Verlhac, “Start-ups in the time of COVID-19: Facing the challenges, seizing the opportunities”, *OECD Policy Responses to Coronavirus (COVID-19)*, 2020.
- [10] L. De Cuyper, B. Kucukkeles and R. Reuben, “Discovering the real impact of COVID-19 on entrepreneurship”, *World Economic Forum*, 2020.
- [11] UK Tech News, <https://www.uktech.news/news/new-analysis-shows-online-marketplaces-set-to-exceed-%247-trillion-by-2024-20190929>, 2019.

# Conceptual Data Model Design for Adaptable Web-Based Museum Information System

Ljubica Kazi\*, Dragica Radosav\*, Dejan Masliković\*\*, Natalija Vulikić\*\*\*, Tijana Stanković Pešterac\*\*\*,  
Zoltan Kazi\*, Biljana Radulović\* and Ivana Berković\*

\* University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia

\*\* Ministry of Culture and Information, Belgrade, Serbia

\*\*\* Muzej Vojvodine, Novi Sad, Serbia

ljubica.kazi@tfzr.rs, dragica.radosav@tfzr.rs, dejan.maslikovic@kultura.gov.rs,  
natalija.vulikic@muzejvojvodine.org.rs, tijana.pesterac@muzejvojvodine.org.rs,  
zoltan.kazi@tfzr.rs, biljana.radulovic@tfzr.rs, ivana.berkovic@tfzr.rs

**Abstract - This paper presents part of results in the project of setting up the unified information system for museums in Serbia in the web environment. Aim of this project was to propose a solution for a web application that could be used by all museums in Republic of Serbia. Adaptability of web application was identified as one of key features, because of changes in museum data recording standards to be applied, as well as diversity of museum collections and their specific data to be registered. Special emphasize in this paper has been put on the proposed conceptual data model in this project, that should support such a web application.**

## I. INTRODUCTION

According to UNESCO and International Council of Museums (ICOM) organizations, members of the international heritage community have acknowledged the lack of adequate human, technical, and financial resources to protect museums and their collections from a variety of threats. [1] The importance of the application of information technologies in the formation of data registers on museum collections is especially emphasized. In the aim to improve their organization, museum information systems are developed to support all business activities.

International organizations have established data models that support obtaining and recording of data related to museum items, i.e. cultural heritage objects. Cultural heritage is defined as “any thing or concept considered of aesthetic, historical, scientific or spiritual significance” [1,2]. The structure of these standard data models have changed during time, as well as their scope and targeting objects. There are some international standards [3, 4], but also some national standards receive international recognition, such as UK [5]. At Serbian national level, one of the most comprehensive documents was established in 2017 (entitled “Guidelines for digitalization of cultural heritage in Republic of Serbia”) [6]. That document defines data structures needed to describe digitalized items of cultural heritage and it has been written according to international standards.

Museum-related software design and development face difficulties of having consistent requirements with all different standards implemented. In constantly changing

environment, one possible solution to software design is to have it directed towards adaptable (or even adaptive) solution. This way, the software could be adjusted to new standards and user requirements. Adaptable solution could be modified by a user or automatically, by importing new standards metadata. Other important issue is related to the need for extending and differentiation of data according to specific needs and types of cultural heritage objects. Aim of this paper is to present a conceptual data model that could support museum information system adaptability, particularly to support standards implementation, specialization to diversity of museum items categories and scalability to support any amount and type of data.

The rest of the paper is organized in following sections – section two describes standards for museum data records formats and exchange, section three presents existing solutions of web –based museum software, section four explains the problem with short summary of the current state, while section five contributes with the proposed conceptual data model, that will support adaptable web portal for museum information system. Final section describes conclusions and future work.

## II. STANDARDS FOR MUSEUM DATA RECORDS FORMATS AND DATA INTERCHANGE

### A. Standards Organizations and Summary of Relevant Standards

Several internationally recognized institutions have defined standards in the field of documentation related to museums:

- UNESCO - United Nations Educational, Scientific and Cultural Organization [7],
- ISO - International Standards Organization [8],
- ICOM - International Council of Museums [9].

There are particular organizational units within aforementioned institutions and separate organization that are responsible for museum documentation and information interchange:

- UNESCO World Heritage Center [10],
- CIMI - Consortium for Interchange of Museum Information [11] [12],
- International Committee for Documentation - CIDOC – organizational unit within the International Council of Museums. [13].

Globally acknowledged and respected are some particular countries institutions, that established standards for museum documentation, such as UK MDA (Museum Documentation Association) [5] and its' successor organization Collection Trust [14].

Summary of some of the most relevant standards related to museum data are presented in Table 1.

TABLE I. OVERVIEW OF MUSEUM-DATA RELATED STANDARDS

YEAR	Standard	Level
1995	CIDOC standards	International
1997	Object ID	International
2002	SPECTRUM	UK
2014, 2020	ISO 21127:2014 Ontology-based interchange of cultural heritage data	ISO

### B. Object ID Standard

ObjectID [15] is an international standard, established and maintained by ICOM organization. This standard identifies and describes an object of cultural goods, i.e. heritage, with archaeological, artistic or other origin. "The Object ID standard was launched in 1997. It has been promoted by major law enforcement agencies, including the FBI, Scotland Yard, Interpol; organizations including the WCO and UNESCO; and by museums, cultural heritage organizations, art trade and art appraisal organizations, and insurance companies. Having established a descriptive standard, Object ID now helps combat the illegal appropriation of cultural objects by facilitating documentation of items from collections and by bringing together organizations from around the world to encourage its implementation." [15] The aim of the application of this standard is to describe the key characteristics of the cultural heritage object, so that the data would be available for international identification and recognition, particularly in the event of theft, when Interpol can be hired on that basis. "The Object ID standard defines nine categories of information as well as four steps to fulfill the procedure. The categories are:

- Type of object,
- Materials and techniques,
- Measurement,
- Inscriptions and markings,
- Distinguishing features,
- Title,
- Subject,
- Date or period,
- Maker.

The four steps are divided as follows:

- Taking photographs of the object,
- Identifying the above mentioned categories,
- Writing a short description, including additional information,
- Keeping the constituted documentation in a secure place."[15]

### C. CIDOC Standards

International Council of Museums [9] is an international institution whose main focus is standardization and institutional work on cooperation and improvement of museums. The standards it adopts refer to specific activities of keeping and publishing museum objects, including loans. Some of the ICOM standards related to documentation are [16]:

- Statement of Principles of Museum Documentation (CIDOC, 2012),
- The CIDOC Conceptual Reference Model (CIDOC, 2011),
- Lightweight Information Describing Objects (CIDOC, 2010),
- The CIDOC Conceptual Reference Model (CIDOC, 2001),
- International Guidelines for Museum Object Information: the CIDOC Information Categories (CIDOC, 1995),
- Recommendations for Identity Photography (CIDOC Fact Sheet 3, 2010),
- Labeling and Marking Objects (CIDOC Fact Sheet 2, 1993),
- Registration Step by Step: When an Object Enters the Museum (CIDOC Fact Sheet 1, 1993).

International Committee for Documentation – CIDOC [13] is an organizational part of International Council of Museums. In 1995, this institution formed the document: "International Guidelines for Museum Object Information: The CIDOC Information Categories" [17]. This document provides guidance on the objectives of documentation in museums, as well as explains groups of relevant data and their formats. The most important data groups refer to the data of the museum object related to: acquisition, condition of the object, description, images of the object, institutions related to the object, location, marking, material and production technique, dimensional measurements and other measurements, association and collections objects, data on the entrance of the object, the name of the object, origin-production, parts, references, reproduction rights, and more.

### D. SPECTRUM Standard

Within the United Kingdom (UK), an Information Retrieval Group of the Museums Association (IRGMA) was formed in the 1970s. In 1977, the Association changed its name to the Museum Documentation Association (MDA) [5]. At that time, a system was formed where each museum was given an identification mark (MDA code). Beginning in 2008, the Collections Trust [14] continues to operate and take over activities previously implemented by the MDA. The MDA standard refers to the structure of data that needs to be recorded within the core business of the museum. The key feature of this standard is that, although it describes the data of individual segments of business processes, it is not process-oriented, because there is no defined procedure for working with data and, in general, business processes within the museum. Some standards developed within the MDA include:

- MDA Data Standard, 1991,
- SPECTRUM procedural and data standard, 1994.

The Consortium for the Interchange of Museum Information (CIMI) [11, 12] develops relevant standards and approaches to the exchange of digital information, through various projects and testing within organizations belonging to the consortium. Projects in 2002. concerned metadata, the description of museum collections, and the structure of a standard XML schema that will represent museum objects. “On behalf of the international museum community, CIMI has developed an XML Schema for the description of museum objects. This is based on MDA’s international museum standard for describing objects called ‘SPECTRUM’; a well-known and widely-used standard in the museum world. The CIMI XML Schema will enable museums to encode rich descriptive information relating to museum objects, including associated information about people, places and events surrounding the history of museum objects, as well as information about their management and use within museums. The CIMI XML Schema will be useful for migrating data, the sharing of information between applications, and as an interchange format of OAI (Open Archives Initiative) metadata harvesting.” [12]

“Spectrum is the UK collection management standard that is also used around the world” [18]. Current version 5 is promoted at the UK Collection Trust web site.

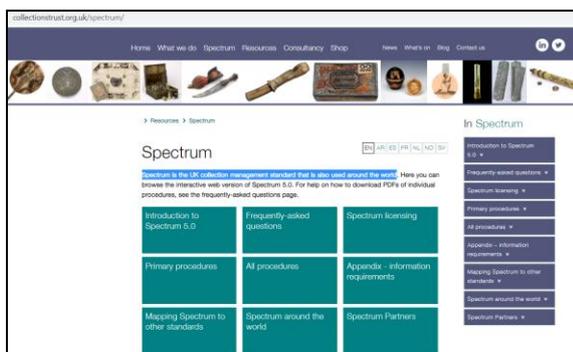


Figure 1. Web site for UK SPECTRUM standard [18]

### E. ISO standards

International Standards Organization ISO has established some general standards that are applicable to multimedia information that is related to digitalized data about museum items. These are:

- ISO/IEC JTC 1/SC 29 – describes coding of multimedia digital presentation. It includes specific aspects, such as synchronization, presenting, recording and transport of separate media or their combination, their security, privacy, quality of experience, evaluation and performance metrics;
- ISO 2709:2008 - defines general requirements for exchange format of recorded data suitable for bibliographic description. This standard describes generalized structure, i.e framework for communication between systems that perform data

processing, but it does not define internal processing and details of interoperability.

There are ISO standards that are focused on museum-related information:

- ISO 21127:2014 “Information and documentation — A reference ontology for the interchange of cultural heritage information” – this standard defines guidelines for data exchange between cultural heritage institutions. It particularly describes information that is under control of museums, libraries and archives. Main focus of this standard is to define exchange and integration of heterogeneous documentation that is related to museum collections. Last version of this standard is made in 2020;
- ISO TC46 – defines standard activities that are related to libraries, documentation, information centers, publishing, archives, records control, museum documentation, indexing and abstracting services.

### III. EXISTING SOLUTIONS OF WEB-BASED MUSEUM SOFTWARE

#### A. Example of Estonian Museum Web Portal

Museum information systems are often developed at particular institution level, but there are many countries that cultural heritage is treated with adequate attention at national level and inter-institutional information systems allow sharing of data across organizational boundaries.

One of illustrative example is the central web portal which unites all museum objects in Estonia [19, 20]. It is based on a database that allows the exchange of data between different museums and the presentation of whole cultural heritage objects of the country, by selecting items of each museum that is listed at the web portal. This integrated web-based information system of the Estonian museum has been in use since 2008 and it has two modules: the part used by employees and the part that is for public use. The portal is implemented through support for three languages: Estonian, Russian and English. It was finalized in 2010. By 2014 it has integrated the work of 61 museums and about 600,000 digital images. It allows visitors to search for museum objects, regardless of the location-specific museum.

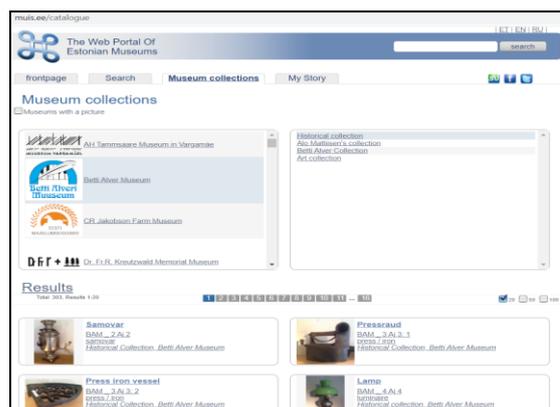


Figure 2. Estonian central web portal for all museums [19]

### B. Serbian Web Portals for Cultural Heritage

In Serbia, much effort has been made to create centralized web-based information system, which includes digitalized relevant documents and multimedia records about all cultural heritage objects. Some of existing or previously developed solutions include:

- Web portal of Republic Institute for the Protection of Cultural Monuments [21];
- Web-based interactive map with presentation of Serbian national heritage buildings and monuments [22];
- “Eternitas” web application for museum information system of Serbia [23];
- Serbian Cultural Heritage Browser [24], created according to [6].



Figure 3. Web portal of Republic Institute for the Protection of Cultural Monuments [21]

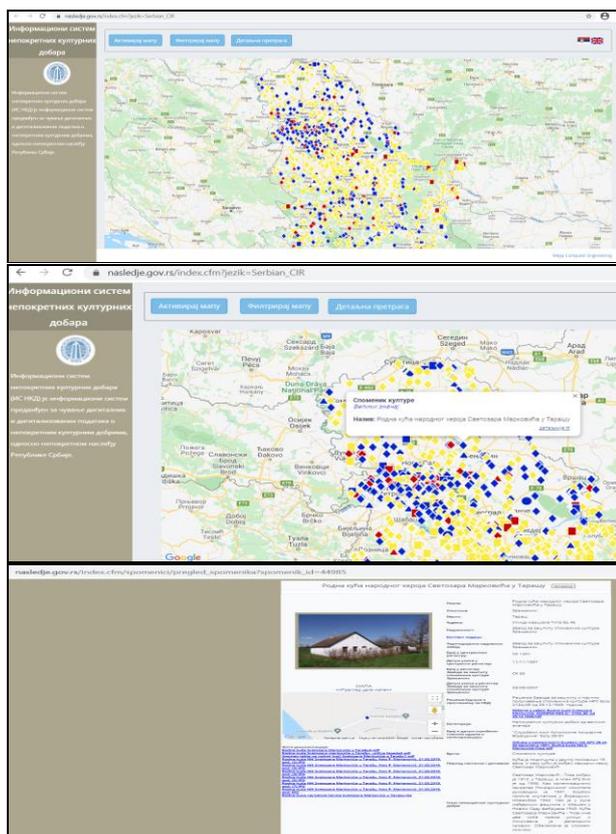


Figure 4. Web-based interactive map with Serbian national heritage buildings and monuments [22]

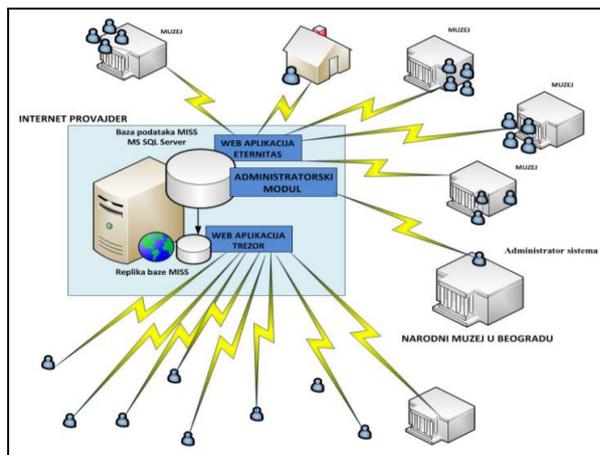


Figure 5. Deployment schema of Eternitas solution [23]

“Eternitas” web application was developed as ASP.NET / MS SQL Server application, based on previously used Klio, CR and MISS solutions, in cooperation of National Museum of Serbia and external partner – company “Softver informacioni sistemi”. Development history behind this solution has been described in [23]. Main aim of previous solutions was to support Central registry of museum fund, i.e. movable cultural goods. Some of important steps towards centralized information systems of museums in Serbia are [23]:

1. Clio program, developed in the National Museum of Serbia in 1993 and in use until 2009;
2. CR program, installed in 1996 at 46 museums in Serbia (Clipper for DOS) and in use until 2009;
3. MISS (unified Museum Information System of Serbia) development started in 1995, as a strategic study. Practical implementation started in 2008. with implementation of central MS SQL Server database, primarily for Central registry of museum fund;
4. Eternitas was the first software solution based on the new MISS MS SQL Server database. This application was officially available to all museums in Serbia from 2011 at [www.app.eternitas.rs](http://www.app.eternitas.rs);

Currently, [www.eternitas.rs](http://www.eternitas.rs) web portal is not used.

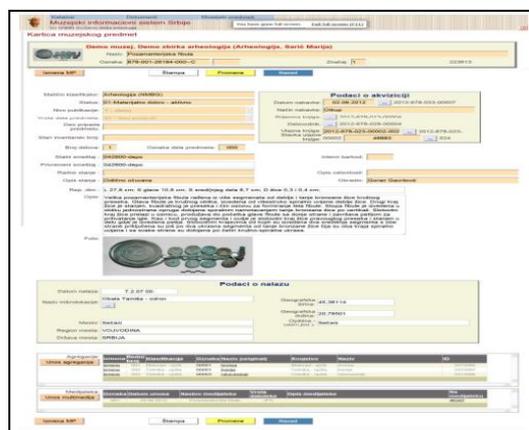


Figure 6. Museum object card in Eternitas system [23]

The IMUS system has been developed by Museum of Vojvodina and Historical Museum of Serbia and it has been used since 2014. [25] Development of new integral

information systems for Serbian museums started in 2018, where History Museum of Serbia received the primary role in conducting the project [26].

Web-application for browser of Serbian cultural heritage [24] has been developed in aim to integrate data and digitalized documents and other items from Serbian archives, libraries, museums, galleries, Institute for protection of cultural monuments and other sources. It has been developed with data structures according to “Guidelines for digitalization of Cultural Heritage in Republic of Serbia” [6], within a project that was conducted with support of Ministry of culture Republic of Serbia, Vojvodina Museum and technically supported by Microsoft Serbia and other IT companies [27].



Figure 7. Serbian Cultural Heritage Browser [24]

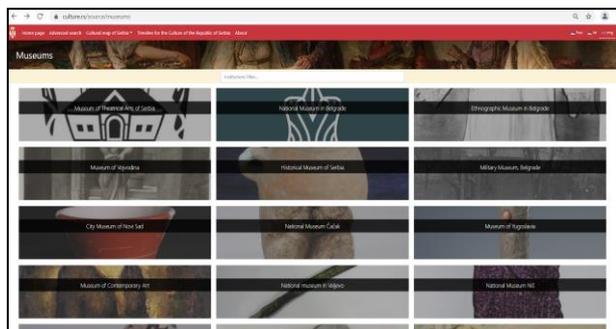


Figure 8. List of Serbian museums at culture.rs portal [24]

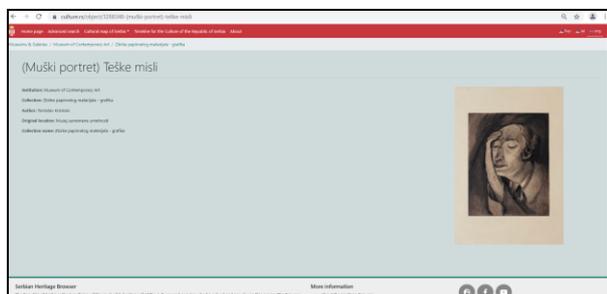


Figure 9. Details about museum item at one of Serbian Museums – Museum of Contemporary Art [24]

#### IV. PROBLEM STATEMENT AND CONCEPTUAL APPROACH FOR NEW ADAPTABLE MUSEUM WEB PORTAL

By analyzing current status of Museum information system in Serbia, it could be concluded that there are several web-based solutions currently in use and they have been developed by different institutions and with support of Ministry of culture and information of

Republic of Serbia. Regarding web support to museums in Serbia, currently there is “culture.rs” portal with integral data from museums, available from data browsing, while there is IMUS system (client-server solution) used in museums in Serbia for data entry. IMUS system is deployed at each museum by installing the software at local museum server and periodically sending copy of relevant data (not all data) to the central database for all museums, located at Ministry of culture and information of Republic of Serbia.

Issues that arise from current IMUS solutions are:

- Software maintenance, i.e. changes in client/server software require all museums to receive an update installation, which is very time-consuming;
- Central database in Ministry of culture and information of Republic of Serbia does not have all data that were entered at local databases of each museum;
- There is need for more easy adjustments of software application to changes of international standards related to museum data acquisition, recording, sharing and transport;
- Data exchange between museums is not supported;
- Currently not all different types of museum collections are supported with appropriate specific data. There is a need to make heterogeneous modules within the solution, to support different types of data for each specific museum collection type;
- The existing software solution is too complex for users, since they are able to see parts of software that are not assigned to their role.

Motivation for creating a new integral information system for museums in Serbia is to establish system that will:

- Enable easy adjustments to constant changes in standards related to museum data acquisition, recording, sharing and transport;
- Centralize the maintenance;
- Increase availability of data to the Ministry of culture;
- Improve the data exchange between museums;
- Simplify the user interface for users from museums;
- Support diversification for presentation of details that are related to the particular user type;
- Support diversity of data needed for a particular museum items category.

Conceptual approach to design of new information system is presented as follows:

- Web portal will centralize data records and enable availability of data for all museums and to the Ministry, with integration of data and easier maintenance of the system;
- “Event-to-Document” related data records in the database, with clear assignment of data entry and maintenance responsibilities of the system users. The approach is: an event produces documents, the event is conducted by a registered user, a document is signed by the registered user and the data from the

document could be stored in the database records. Data in documents describe events and characteristics of related museum items (objects);

- Support to data records related to different types of multimedia digital representation of museum items;
- Adaptable system:
  - Scalable to any amount of data items (features) related to any type of museum item (i.e. museum items categories),
  - Adaptable to changes in museum documentation standards,
  - Specialization supporting – adjustable to fit into data records for any particular type of museum object.

### V. THE PROPOSED CONCEPTUAL DATA MODEL DESIGN

Data model represents the essence of any information system. In this particular case of museum information system, the proposed conceptual data model is designed to

support aforementioned conceptual design approach. Therefore, there are entities in the data model that will enable records on:

- Activities (events) with types of activities (such as procurement, restoration, exhibition) with attributes of activities (characteristics). Each activity produces one or many multimedia results;
- Assigned employer is responsible for the activity conduction;
- Multimedia result describes the event i.e. activity, such as document, photograph, 3D model, video etc;
- Multimedia result consists of one or many items, i.e. particular museum items (objects) listed or included in the multimedia result;
- Each museum item (object) is described with set of characteristics of certain type that has appropriate value.

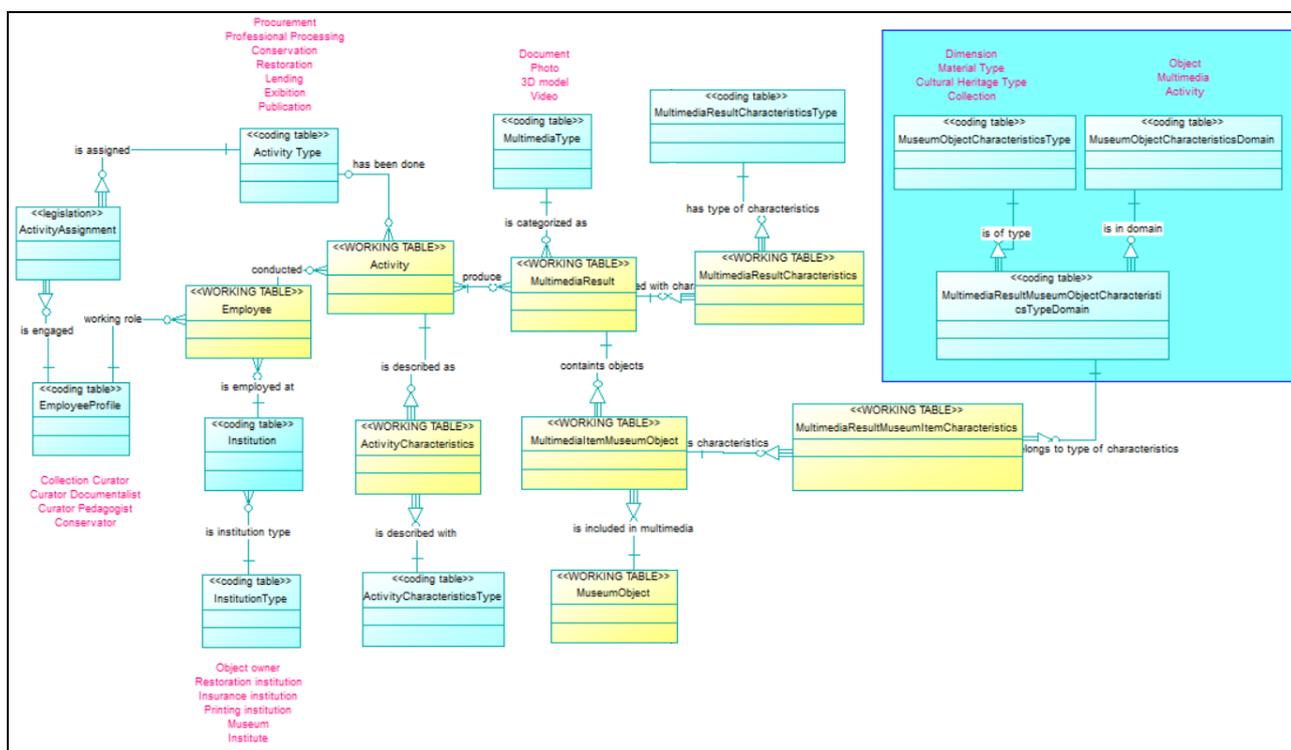


Figure 10. The proposed conceptual data model for adjustable museum web portal

### VI. CONCLUSION

This paper presents part of results in the conceptualization of creating new web-based museum information system in Republic of Serbia. The focus of this paper is in conceptual data model that will enable adaptability of the solution.

To make an introduction and represent the current status in the field related to the topic of this paper, several most important internationally recognized standards were analyzed. These standards are related to data acquisition, recording and transport between cultural heritage

institutions, with particular emphasize on museums. The most important standards come from ISO, UNESCO and other international museum-related institutions, but there is also significant role of UK institutions, which are globally acknowledged with their standards for museum-related documentation and information systems.

As an important part of this paper there is also a brief overview of existing web-based solutions applied in cultural heritage institutions, particularly related to museums in Serbia. It has been shown that there are several web portals in use, created by different Serbian institutions.

Issues of currently used system IMUS, as well as diversity of international standards for data recording and exchange, induced the idea of creating adaptable web portal for museums. This paper presents the conceptual data model that could support adaptability, scalability and specialization of data regarding particular museum categories of items.

By implementing web portal based on the proposed data model, it would be possible to improve the maintainability of the system and to enhance the user experience.

#### ACKNOWLEDGMENTS

This paper presents part of results from the project “Application of digital photogrammetry within creating digital 3D objects, at the example of Antique collection of Vojvodina Museum”, financed by Ministry of culture and information of Republic of Serbia.

#### REFERENCES

- [1] UNESCO Expert Meeting on the protection and promotion of museums and collections, Rio de Janeiro, Brazil, 11-14 July 2012, <http://www.unesco.org/culture/pdf/eo/en.pdf>
- [2] ICOM Code of Ethics for Museums, revised, 2004.
- [3] Object ID international standard for documentation related to identifying and recording data about cultural goods, <https://icom.museum/en/resources/standards-guidelines/objectid/>
- [4] Standards within the International Council of Museums, <https://icom.museum/en/resources/standards-guidelines/standards/>
- [5] Museum Documentation Association – MDA - <http://www.mda.org.uk/>
- [6] Guidelines for Digitalization of Cultural Heritage in Republic of Serbia, <https://www.kultura.gov.rs/extfile/sr/205/smernice-za-digitalizaciju-kulturnog-nasledja-u-republici-srbiji.pdf>
- [7] Unesco organization, <https://en.unesco.org/>
- [8] International Organization for Standardization, <https://www.iso.org/home.html>
- [9] International Council of Museums, <https://icom.museum/en/>
- [10] UNESCO World Heritage Center, <https://whc.unesco.org/en/>
- [11] David Bearman and John Perkins: “Standards Framework for the Computer Interchange of Museum Information”, 1983, <https://cool.culturalheritage.org/byorg/cimi/cimifram.html>
- [12] Consortium for the Interchange of Museum Information <http://xml.coverpages.org/cimi.html>.
- [13] International Committee for Documentation CIDOC, International Council of Museums, <https://icom.museum/en/committee/international-committee-for-documentation/>
- [14] Collection Trust - <https://collectiontrust.org.uk/>
- [15] Object ID international standard for documentation related to identifying and recording data about cultural goods, <https://icom.museum/en/resources/standards-guidelines/objectid/>
- [16] Standards within the International Council of Museums, <https://icom.museum/en/resources/standards-guidelines/standards/>
- [17] “International Guidelines for Museum Object Information: The CIDOC Information Categories”, International Committee for Documentation CIDOC, International Council of Museums, <https://icom.museum/en/ressource/international-guidelines-for-museum-object-information-the-cidoc-information-categories/>
- [18] SPECTRUM standard for museums, <https://collectiontrust.org.uk/spectrum/>
- [19] Estonian Museum Portal, [www.muis.ee](http://www.muis.ee)
- [20] Estonian Museum Portal, description: <https://medisoft.ee/en/projects/development-projects/museum-information-system/>
- [21] Republic Institute for the Protection of Cultural Monuments (Republički zavod za zastitu spomenika kulture), <http://www.heritage.gov.rs/cirilica/index.php>
- [22] Map of Historical Monuments in Serbia, <https://nasledje.gov.rs/index.cfm>
- [23] Goran Gavrilovic, “The Eternitas web application – implementation step of the museum information system of Serbia”, Pregled NCD 27 (2015), 1-12. <http://elib.mi.sanu.ac.rs/files/journals/ncd/27/ncdn27p1-12.pdf>
- [24] Serbian Cultural Heritage Browser, <https://culture.rs/>
- [25] IMUS system catalogue, <http://esnaf-m.com/wp-content/uploads/2016/01/katalog-2-VERZIJA-.pdf>
- [26] Start of new integral information systems of museums in Serbia, <http://imus.org.rs/pocetak-projekta-implementacije/>
- [27] Details about project of creating Serbian Cultural Heritage Browser, <https://culture.rs/about>

# Clean Code Quality Attributes and Measurements: an Initial Review

Lj. Kazi\*, S. Mihajlović\* and M. Bhatt\*\*

\* University of Novi Sad, Technical Faculty “Mihajlo Pupin”, Zrenjanin, Serbia

\*\* Pragat Consultancy Services – Mumbai, India

ljubica.kazi@gmail.com, mihajlovicsinisa25@gmail.com, mmbhatt@gmail.com

**Abstract - This paper analyzes the concept of clean code, in context of software structural quality. It presents a review of quality attributes of a clean code, related to ISO 25010 standard for quality of a software product and attributes of structural aspect of software quality. It also provides an overview of clean code effects experiments and clean code attributes measurements methods. This paper also addresses methods, practices and effects of refactoring existing code to clean code.**

## I. INTRODUCTION

According to [1], code quality is important because there is estimation that 80% of all software work is maintenance and reading code takes a large portion of coding time. Bad code is difficult to understand, extend, maintain and test, which lead to delays in releases.

“The true nature of clean code rely on the fact that other practitioners should be able to read and maintain the code. Clean code is not something a machine or script can recognize (so far) but rather something that professionals can decide”. [2] Robert Martin describes that [3] writing a clean code is similar to creating art, but it requires discipline in applying set of techniques to reach the clean code goal. Having such an importance in programming practice [4] [5], it is crucial to emphasize the relevance of clean code within the educational environment and prepare students for requirements of code quality [6].

Many programming communities and companies have established informal rules that direct creating code with a specific technology – coding conventions [7]. These rules are related to naming conventions, code formatting and layout, commenting etc. By following coding conventions, it enables [7] a consistent presentation of the code, so that readers can focus on the code meaning, not layout. This enables readers to understand the code more quickly and facilitates maintaining the code.

Aim of this paper is to outline the concept and attributes of the clean code, as well as possible measurements methods and techniques to enable evaluation of a code regarding its’ quality, particularly cleanness aspect. [8] Particular focus is put on different programming languages and comparison of their rules in this context. Finally, the concept of refactoring the code is related to activities of establishing a code with a better quality, i.e. to have the code improved regarding the clean code principles.

## II. STRUCTURAL ASPECT OF SOFTWARE QUALITY AND CLEAN CODE CONCEPT

ISO (International Standards Organization) standard for software quality (ISO/IEC 25010:2011) defines quality attributes of a software product, which are presented at Figure 1. These attributes include functional suitability, reliability, performance efficacy, operability, security, compatibility, maintainability, portability.

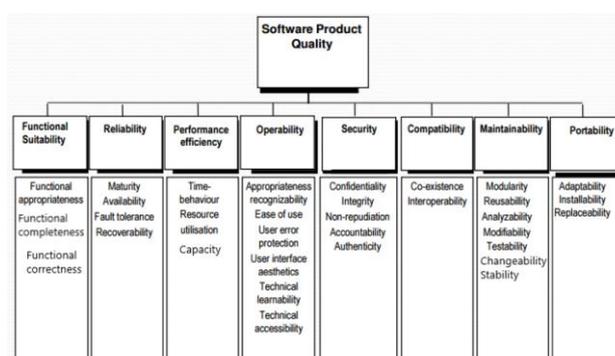


Figure 1. Quality attributes of software as a product (according to [9] [10])

According to Chappel [11], structural aspect of software quality “means that the code itself is well structured” and includes characteristics such as:

- “Code testability. Is the code organized in a way that makes testing easy?”
- Code maintainability. How easy is it to add new code or change existing code without introducing bugs?
- Code understandability. Is the code readable? Is it more complex than it needs to be?
- Code efficiency, especially in resource-constrained situations.
- Code security. Does the software allow common attacks such as buffer overruns and SQL injection?” [11]

“Clean code advocates for writing readable code, so that other people can know the code intent almost directly“ [3] [5]. Code readability leads to better understanding and increase of maintainability. The concept of clean code is tightly related to concepts of code

quality and code smells, but it is associated to naming, structuring, formatting, refactoring and testing of code.[5].

Clean code could be defined with multiple definitions and described via different attributes, and some of them have been provided by different authors in [3]:

- *Bjarne Stroustrup*, inventor of C++ programming language, when describing quality code, uses word elegant and efficient, logic straightforward, dependencies minimal, error handling complete and performance close to optimal.
- *Grady Booch*, author of the book “Object-Oriented Analysis and Design with Applications” [10], explains the concept of clean code as simple and direct and highlights the code readability and clearly presents the designers intent (i.e. the purpose of the code).
- *Dave Thomas*, founder of Object Technology International which produced the basis for Eclipse IDE, also points out the readability and ease of making improvements, but performed by other people, not original author of code. He also requires unit and acceptance tests to be enclosed with the clean code, while having minimal dependencies explicitly defined.
- *Ron Jeffries*, author of books related to Extreme Programming, underlines simplicity of code - runs all tests successfully, contains no duplication, it is expressive with meaningful names and articulate all the design ideas that are included in the system, minimizes the number of elements (such as classes, methods, functions), follows rules of SOLID principles (like single responsibility and using abstractions early in building and application) etc.

### III. QUALITY ATTRIBUTES AND STRATEGIES IN BUILDING CLEAN CODE

Robert C. Martin has presented the most important quality attributes of a clean code, in the handbook that could be used as a reference in coding practice [3]. These important clean code elements and appropriate attributes are presented in Table I.

TABLE I. CLEAN CODE ELEMENTS AND ATTRIBUTES, ACCORDING TO ROBERT C. MARTIN [3]

CODE ELEMENTS	CLEAN CODE ATTRIBUTES
Names	Meaningful, intention-revealing, searchable, pronounceable, avoid encodings (type-related names), using problem domain names, avoid mental mapping names (too short names, such as single letter)
Functions	Small, single focus, one level of abstraction per function, no side effects, command query separation, function arguments,
Comments	Clear and complete, Informative, explanation of intent, warning of consequences, amplification the importance, non redundant, non misleading

Formatting	Vertical formatting, horizontal formatting, openness between concepts, density, distance, ordering, indentation, alignment, scopes
Classes	Encapsulation, small classes (single responsibility, cohesion), organizing for change
Objects	Data transfer objects, active records
Data Structures	Data abstractions
Error Handling	exceptions with try/catch blocks (instead of returning error codes), no returning null
External modules	Boundaries to third-party code, having dependency to external services wrapped into separate classes “having very few places in the code that refer to them” [3].

There are some strategies, i.e. activities related to code creating decisions and testing, which support clean code building, such as those presented in Table II.

TABLE II. CLEAN CODE STRATEGIES, ACTIVITIES AND ASPECTS

CLEAN CODE STRATEGIES, ACTIVITIES AND ASPECTS	DETAILS
Unit tests	Modularization of software leads to unit tests, but also integration tests
Aspect-oriented approach	Cross-cutting concerns, Separation of concerns
Standards	Using standards wisely, but be focused on value to bring with the system
Using frameworks	Semi-final modules, configuration, industrial software development
Test driven development	Test-based specifications
Dependency injection	Abstraction first, creating interfaces, then specific classes
No duplication	No duplicate code, but abstraction to a single class
Refactoring	Improving structure while preserving user interface and functionality
SOLID principles	Single responsibility, Open-close, Liskov substitution, Interface segregation, Dependency inversion
Synchronization	Concurrency of code
Dependencies	Among modules, class libraries
Code smells heuristics	Typical errors and refactoring to solve issues

By analysis of the previously presented, it could be concluded that the concept of clean code is closely related to many ISO standard-defined quality attributes for software as a product and, also, to quality attributes of structural aspect of software quality. The concept of clean code, related to certain attributes that are included in both ISO standards and particular area of structural aspect of quality, has been presented at Table 3.

TABLE III. RELATING THE CLEAN CODE ATTRIBUTES WITH ISO STANDARD FOR SOFTWARE AS A PRODUCT QUALITY ATTRIBUTES AND STRUCTURAL QUALITY ATTRIBUTES

IMPLEMENTED WITH				CLEAN CODE			
FUNCTIONALITY	STRUCTURE	READABILITY	TESTABILITY	RELIABILITY	PERFORMANCE	MAINTAINABILITY	SECURITY
ISO QUALITY AS PRODUCT	STRUCTURAL QUALITY	STRUCTURAL QUALITY	STRUCTURAL QUALITY	ISO QUALITY AS PRODUCT	ISO QUALITY AS PRODUCT	ISO QUALITY AS PRODUCT	ISO QUALITY AS PRODUCT
Logic straightforward	Object-oriented structure	Naming		Testing	Efficiency	Easy to change	
	Modular structure	Formatting				Easy to understand - Clear intent, Expressive	
	Dependencies minimal, maximal cohesion, external modules dependency wrapping	Comments					
	No duplication						
	Minimal elements						
	SOLID principles						
	Organization of functions, classes, objects, data structures						
	Error handling						

According to Table III, there are certain quality attributes that are not included in ISO-defined quality attributes, but are included in structural quality attributes and clean code attributes as well, such as: structure, readability, testability. Since testability is closely related to structure, i.e. it is defined as a structure of software suitable for testing, therefore it could be concluded that there are two most important aspects of clean code (as well as structural aspect of software quality), different from general ISO 25010 standard:

- Software structure – organization of code, functions, classes, objects, dependencies, application of standards, frameworks, SOLID principles and error handling;
- Software readability – names, formatting and comments, which leads to better expressiveness, understandability and clear intent presentation.

#### IV. CLEAN CODE EFFECTS – A SHORT REVIEW OF EXPERIMENTAL RESULTS

Academic research and education has been challenged with the clean code aspects, in aim to improve approaches, methods and tools for evaluation, but also to improve professional [5] and teaching practice [6] in making clean code.

Some research and educational efforts are related to empirical analysis of impact of clean code to variety of software attributes and software development process, such as understandability. In research [13], experiment was conducted with legacy code transformed into a clean code version. Two groups of developers were given assignments upon legacy code version and clean code version of software and it has been proven that clean code version group was more efficient (faster) in solving given tasks.

Particular issue arises when new developers enter a software project and there is a large amount of existing code to understand. Study [14] presents results of experiment with software project newcomers, that were

exposed to previously unknown code. They were separated to two groups – one having the refactored code (made to to adhere to clean code guidelines) and other having introductory document to gain familiarity with existing (non-refactored and non clean) code. In this experiment, both groups were given tasks to implement within appropriate given version of software. It was concluded that the group with refactored (i.e. clean) code was more productive than the group having the introductory document about non-clean version of software.

Research [15] examines the impact of clean code to unit testing. One of key hypotheses in this research is that applying clean code principles could lead to better code coverage in unit testing. Experimental research was conducted with two groups of developers that were given the task of write unit tests – one group for ad-hoc written code and the second for clean code. It was observed that the clean code achieves a higher code coverage in unit testing, compared to ad-hoc written code.

#### V. CLEAN CODE MEASUREMENT METHODS

As previously highlighted, the essence of clean code concept is readability and understandability. There are some recent research efforts in establishing methods for measurements related to clean code or clean code attributes.

Research [16] is conducted in aim to address issues related to source code readability. This feature could be used as a metric of the extent of source code comprehension. “The better the code is readable, the easier it is for code readers to comprehend the system based on the source code” [16]. This research proposed an enhanced source code readability metric which is a quantitative measure of the extent of a code readability and two-way linear regression analysis was performed with data collected by a survey.

“Reading code is an essential activity in software maintenance and evolution” [17]. There are several studies that investigated how different factors, such as naming conventions and employed programming constructs, influence code readability. Research [17] addresses the concepts of code readability (ease of reading and apprehending of code) and code legibility (ease of identifying elements of a program). This separation directs the research towards linguistics. This research contributes with comprehensive literature review and an approach to relate program readability with program comprehension as a human-centric learning activity and appropriate learning taxonomy and competencies.

Classification of code from the aspect of readability has been considered in research [18]. This classification is related to referring a source code segment as being either readable or unreadable. This paper contributes with automated source code readability classification, based on the use of Convolutional Neural Networks and Deep Learning Techniques. The proposed approach has been examined with experiment and it prove the efficacy, comparing to manual feature engineering.

## VI. REFACTORING TO CLEAN CODE

“Code refactoring is a process used in a software development approach that involves editing and cleaning previously written software code without any change in code function. The main purpose of code refactoring is to make code more efficient and maintainable.” [19]

Architectural refactoring and clean up upon previous code was addressed in research [20]. In this work, experiences were reported in this large migration from one to another software architecture, while preserving the functionality. The migration was performed with the use of reverse engineering upon industrial Java application. Together with architectural migration, the code clean up has also been performed. This paper contributes with experimental results of applying various quality metrics that describe the migration effects.

There are various approaches and methods to code refactoring, with special focus on improving cleanliness of code in target solution. One of such approaches use databases as a tool for shifting a software solution from one form to another [21]. In this research, databases are used instead of syntax trees and the authors elaborate advantages of this approach.

In research [22], special attention has been put on readability and comprehensibility of the source code made with functional programming languages. Particularly, this paper presents new methods of refactoring imperatively styled source code into a functional-styled code. Empirical findings (performed with real-life examples and data) contribute with conclusions that refactorings of this type lead to shortenings in length of source code, but also affect the complexity and readability.

Technical debt grows as software systems evolve over time, but it could be reduced by using refactoring or using clean code approach in creating new code, where technical

debt is limited. Research [23] addresses both cases, but focuses on the second approach. It investigates technical debt density in new code comparing to existing code. Experimental research has been conducted with 27 open-source software projects by the Apache Software Foundation, with more than 66000 classes and more than 50000 commits. The results enable making conclusions that writing clean new code can be efficient strategy for reducing technical debt density and preventing software decay over time. Explicit policy regarding quality improvements and team meetings discussions influenced higher frequency of cleaner new code commits. Therefore, it is emphasized that there should be a process established upon technical debt density monitoring in new code, but also to control clean code practices.

## VII. CONCLUSION

Clean coding has become inevitable in modern programming practice. Importance of this approach has been proven in multiple experiments, where it has been shown that writing a clean code could lead to better readability, maintainability and lower density of technical debt.

There is a need to address clean coding in educational environment and some efforts have been made in this particular direction. Academic environment also recognizes this area as an important research topic, which was addressed in several previous years in research results.

Still, there is a large research space open, when clean code is in the focus, starting with the concept definition, key attributes, evaluation methods, experiments in practical use of clean code and evaluation methods in professional and educational environment. Professional challenges of writing clean code could be directed towards development of software frameworks that incorporate principles of clean code and the software tools for automated evaluation of particular clean code attributes.

## REFERENCES

- [1] “Clean Code or: How to care for code”, <https://rti.etf.bg.ac.rs/rti/13S112POOP/materijali/Clean%20Code.pdf> (accessed 28.09.2021.)
- [2] M. Anaya, “Clean Code in Python: Refactor Your Legacy Code Base”, Packt Publishing Limited (Verlag), August 2018. Discussion on: <https://hub.packtpub.com/clean-coding-in-python-with-mariano-anaya/> (accessed 28.09.2021.)
- [3] Robert C. Martin: “Clean Code, A Handbook of Agile Software Craftmanship”, Pearson Education, 2009.
- [4] R. Rachow, S. Schoder, M. Riebisch, “Missing Clean Code Acceptance and Support in Practice – An Empirical Study”, 25<sup>th</sup> IEEE Australian conference on Software Engineering, 2018.
- [5] K. Kjung, “Clean Code in Practice, Developers’ Perception of Clean Code”, Bachelor Thesis, Blekinge Institute of Technology, Faculty of Computing, Karlskrona, Sweden, May 2021.
- [6] L.W. Dietz, J. Manner, S. Harrer, J. Nenhard, “Teaching Clean Code”, Software Engineering, 2018.
- [7] Microsoft, “Coding conventions in C#”, <https://docs.microsoft.com/en-us/dotnet/csharp/fundamentals/coding-style/coding-conventions#> (accessed 28.9.2021).

- [8] Sealights, “Code quality metrics: Is your code any good?”, <https://www.sealights.io/code-quality/code-quality-metrics-is-your-code-any-good/> (accessed 28.09.2021)
- [9] Desharnais J-M: —Software Measurement - Analysis of ISO/IEC 9126 and 25010, <http://www.testingnotes.eu/wpcontent/uploads/9126-vs-25010.pdf>
- [10] Hwang S.M: Process Quality Levels of ISO/IEC 15504, CMMI and K-model, International Journal of Software Engineering and Its Applications, Vol 3, No 1, January 2009.
- [11] David Chappel: The Three aspects of software quality – functional, structural and process, David Chappel and Associates, [http://www.davidchappell.com/writing/white\\_papers/The\\_Three\\_Aspects\\_of\\_Software\\_Quality\\_v1.0-Chappell.pdf](http://www.davidchappell.com/writing/white_papers/The_Three_Aspects_of_Software_Quality_v1.0-Chappell.pdf) (accessed 28.09.2021)
- [12] G. Booch, R. A. Maksimchuk, M.W. Engle, B.J.Young, J. Conallen, K. A. Houston, “Object-Oriented Analysis and Design with Applications”, Addison-Wesley, 2007.
- [13] H. G. Koller, “Effects of Clean Code on Understandability”, University of Oslo, Department of Informatics, Master’s Thesis, Spring 2016.
- [14] R. Bottema, “Investigating Whether Clean Code Helps Developers in Understanding a New Project”, Master ‘s Thesis, Delft University of Technology, Faculty EEMCS, Department of Software Technology, Delft, Netherlands, 2019.
- [15] T. Laneva, “Influence of Clean Code on Unit Testing”, Bachelor Thesis, Haaga-Helia University of Applied Sciences, Helsinki, Finland, 2020.
- [16] S. Chou, S. Kim, J-H Lee, J. A. Kim, J-Y. Choi, “Measuring the Extent of Source Code Readability Using Regression Analysis”, International Conference on Computational Science and Its Applications, ICCSA 2018, Lecture Notes in Computer Science LNCS, vol 10963, pp. 410-421.
- [17] D. Oliveira, R. Bruno, F. Madeiral, F. Castor, “Evaluating Code Readability and Legibility: An Examination of Human-centric Studies”, IEEE International Conference on Software Maintenance and Evolution (ICSME), 2020.
- [18] Q. Mi, J. Keung, Y. Xiao, So. Menash, Y. Gao, “Improving Code Readability Classification Using Convolutional Neural Networks”, Information and Software Technology, Elsevier, Vol. 104, December 2018, pp. 60-71.
- [19] Altexsoft, “Refactoring best practices”, <https://www.altexsoft.com/blog/engineering/code-refactoring-best-practices-when-and-when-not-to-do-it> (accessed 28.09.2021.)
- [20] R. Koschke, “Industrial Experience on Code Clean-up Using Architectural Conformance Checking”, Proceedings of the 12th European Conference on Software Architecture ECSA ’18, September 24–28, 2018, Madrid, Spain, pp.1-7.
- [21] P. Divianszky, R. Szabo-Nacs, Z. Horvath, “A Framework for Refactoring Clean Programs”, Proceedings of 6<sup>th</sup> International Conference on Applied Informatics ICAI 2004.
- [22] B. Poor, M. Toth, I. Bozo, “Transformations Towards Clean Functional Code”, Proceedings of the 19<sup>th</sup> ACM SIGPLAN International Workshop on Erlang ’20, August 23, 2020, USA, pp. 24-30.
- [23] G. Digkas, A. Chatzigeorgiou, A. Ampatzoglou, P. Avgeriou, “Can Clean New Code Reduce Technical Debt Density”, Journal of Latex Class Files, Vol 14, No 8, August 015.

# Synthetic media (Deepfake) generation and detection methods and challenges

Sasa Arsovski\*, Angely Sim Jia Wun\*, Branko Markoski\*\*, Aleksandar Sofic\*\*\*, Velibor Premcevski\*\*

\*Raffles University, Malaysia

\*\* University of Novi Sad, Technical faculty "Mihajlo Pupin" Zrenjanin, Republic of Serbia

\*\*\* University of Novi Sad, Faculty of Technical Sciences, Republic of Serbia

arsovskisasa@raffles-university.edu.my, angely@raffles-university.edu.my, markoni@uns.ac.rs,  
velci.aspire@gmail.com, sofic.ftn@gmail.com

**Abstract** - Synthetic media (Deepfake) are digital manipulations of an image/video in which a person and/or content of the media has been altered. Artificial Intelligence (AI) driven Deepfake can lead towards online disinformation, reduce trust in social media, and challenge online culture. Deepfakes can easily change viewer perception of truth in business or politics, so it is very important to develop models and strategies to improve their detection. In this paper authors will analyze latest state-of-the-art Deepfake generation and detection methods. The aim of the research that is presented in this paper is to detect and analyze disparity gap between Deepfake detection methods - machine learning and deep learning approaches used for binary classification and generative adversarial approaches used for Deepfake generation. From the results of our experiment, we can conclude that binary classifiers used for Deepfake detection are not reliable enough to properly generalize real-world data, thus Deepfake media still represent a high threat in online society, and new technological solutions are indispensable to counter this challenge.

**Key Words:** *Synthetic media, Deep Neural Networks, Autoencoders, Machine Learning*

## I. INTRODUCTION

Digital manipulations of an image/video recently become a threat to online media, human society, and democracy [1],[2]. According to the authors in [3] synthetic media refers to news style and media content that is fabricated to deceive the public. Fake information spreads through social media, and it can impact millions of users [4]. Today, most of the Internet users get their news via YouTube, Facebook [5]. It is hard to know what to trust, and what effect fake information can induce decision making [6]. Digital manipulations and disinformation warfare led by different actors tends to manipulate public opinion [7].

There are few different approaches to the Deepfakes generation. Face swap, lip synchronization, face synthesis and audio Deepfake. Generating videos by animating object in still images has various application across different fields such as movie production and e-commerce [8]. Deep learning approaches as facial recognition algorithm and

variational auto-encoder (VAE) are used swap one person's face with another and to make a deepfake video [9]. VAEs' role is to encode images into low-dimensional representation before decoding those representations back into images. To transform any video into a deepfake, two auto-encoders are needed. One auto-encoder is trained on the images of the target person's face while the other trained on images of a wide diversity of faces. The images of faces used for both training sets can be curated by applying a facial recognition algorithm to video frames in order to capture different possess and lighting conditions that naturally occur. After the training is done, the encoder that was trained on the diverse faces will be combined with the decoder trained on the target's face, this would result the target's face to be on someone else's body. The implementation of the Deep learning techniques and voice cloning into deepfakes generation become a unique challenge for deepfake detection. In the research [10] authors present deepfakes detection approaches. In the research [11] deepfake detection approaches are briefly analyzed.

In this paper, authors will discuss the probabilistic perspective of the typical machine learning models based on predictive modeling and Autoencoder deep learning architectures which can be formulated as self-supervised models. Currently, Deepfake detection methods are defined as frame-based binary classification problems, and the quality and accessibility of Deepfake datasets present the biggest challenge of Deepfake detection methods.

To evaluate accuracy of the available Deepfake detection methods, we generate a Deepfake video using the latest, state-of-the-art Deep learning model for Deepfake generation based on generative neural network architecture and a first-order motion model [9]. Generated Deepfake video is used to test accuracy of available Deepfake detection online solutions.

## II. EXPERIMENT

### Generation of the Deepfake content

To create Deepfake video using First order motion model described in [9] and Google-Colab platform [12]. Google-Colab is a product of Google Research that allows users to execute Python code through the browser while providing free access to computing resources including GPUs [12]. In our experiment (Figure 1.) we implemented the research presented in [9]. Authors [9] train Neural network model to reconstruct the training videos using combination of the single frame and representation of the motion in the video. Model learns to encode motion as a combination of motion key-point and local transformations.

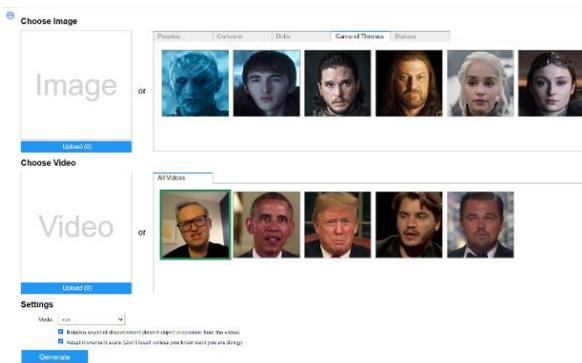


Figure 1. Deepfake Generator

As is shown on Figure 2., Authors use source image  $S$  and a frame of a driving video (frame  $D$ ) as inputs. The unsupervised key-point detector extracts first order motion representation consisting of sparse key-points and local affine transformations with respect to the reference frame  $R$ . The dense motion network uses the motion representation to generate dense optical flow  $T_{S \leftarrow D}$  from  $D$  to  $S$  and occlusion map  $O_{S \leftarrow D}$ . The source image and the outputs of the dense motion network are used by the generator to render the target image (Figure 2.).

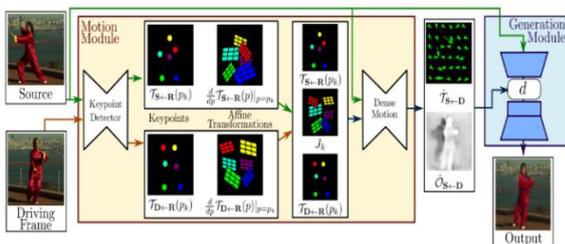


Figure 2. First order motion model from [9]

In our experiment we purposely choose to animate Leonardo da Vinci. Mona Lisa portrait shown on Figure 3. As driving video, we use author video as is shown on Figure 4.



Figure 3. Mona Lisa portrait by Leonardo da Vinci

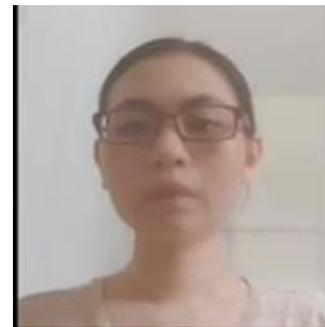


Figure 4. Driver video

For the sound generation, service from naturalreaders.com is used, as the website comes with different sound generator that could be chosen by the user.

The result of the generation successfully caused the target image to move and follow the movement as in the video. However, it wasn't very realistic and one of the factors that could improve this is having video editing skills that could adjust the glitching and some unnatural angle at some spot in the video. Figure 5. shows the pictures of the output video generated from the still image.



Figure 5. Frames of Output Video

### Detection of the Deepfake content

For detection of our generated Deepfake video we use several free available Internet tools. First used Deepfake detection tool is designed to recognize the latest AI-based media manipulation and synthesis techniques, including fake human faces in social media profiles, and realistic face swaps in videos. This platform can detect only face swap, and face reenactment models as presentation or replay attacks. For our Deepfake video platform return result showed on the Figure 6.



Figure 6. First Deepfake detection platform results

The second Deepfake detection tool is developed as AI-based antivirus engine. Company develops and opensource Deepfake scanner. This platform also could not recognize our Deepfake video, as is shown on Figure 7.

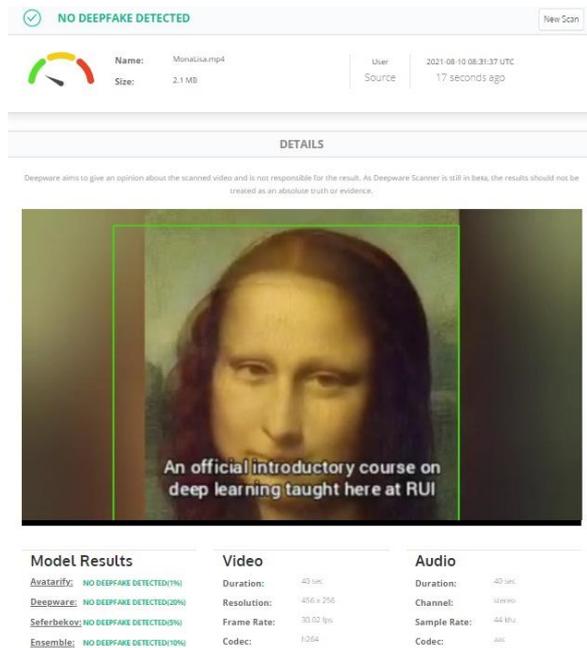


Figure 7. Second Deepfake detection platform results

### III. DISCUSSION

Authors in [13] extend Google’s dataset with 1000 videos from YouTube, from which they extracted images by applying techniques presented in Face2Face [14], FaceSwap [15], DeepFakes [16] and NeuralTextures [17]. The authors [13] fine train an Xception net, pre-trained on ImageNet, to detect real and Deepfake videos. Presented results in the paper [13] suggest a state-of-the-art Deepfake detection mechanism tailored to face manipulation techniques. In the research [18] authors tested proposed state of the art model for the Deepfake detection proposed in [13] and conclude that both the training and test data are drawn from the same distribution and this distribution doesn’t represent examples of the Deepfake videos found in real life. In other words, train and test data don’t represent real-world to achieve generalize to unseen data. This means that the model, when developed using such test data, usually overfit to the data samples specific to the paper from [13]. General conclusion is that there is no broad enough data set to enable generalization of the trained model.

### IV. CONCLUSION

In the recent future, we can expect potentially destructive use of Deepfake media, particularly as phishing attacks, and frauds. Currently, the vast majority of deepfakes are deepfakes of politicians. With the new generation of the Deepfake tools that become publicly accessible, and the Deepfake media spreading across the social media platforms, damage made by Deepfake may be irreparable. From the results of our experiment, we can conclude that binary classifiers used for Deepfake detection are not reliable enough to properly generalize real-world data, thus Deepfake media still represent a high threat in online society, and new technological solutions are indispensable to counter this challenge.

### REFERENCES

- [1] Borges, L., Martins, B., & Calado, P. 2019. Combining Similarity Features and Deep Representation Learning for Stance Detection in the Context of Checking Fake News. *Journal of Data and Information Quality*, 11(3): Article No. 14. <https://doi.org/10.1145/3287763>
- [2] Qayyum, A., Qadir, J., Janjua, M. U. & Sher, F. 2019. Using Blockchain to Rein in the New Post-Truth World and Check the Spread of Fake News. *IT Professional*, 21(4): 16–24. <https://doi.org/10.1109/MITP.2019.2910503>
- [3] Aldwairi, M., & Alwahedi, A. 2018. Detecting Fake News in Social Media Networks. *Procedia Computer Science*, 141: 215–222. <https://doi.org/10.1016/j.procs.2018.10.171>
- [4] Figueira, A., & Oliveira, L. 2017. The current state of fake news: challenges and opportunities. *Procedia Computer Science*, 121: 817–825. <https://doi.org/10.1016/j.procs.2017.11.106>
- [5] Anderson, K. E. 2018. Getting acquainted with social networks and apps: combating fake news on social media. *Library HiTech News*, 35(3): 1–6
- [6] Borges, L., Martins, B., & Calado, P. 2019. Combining Similarity Features and Deep Representation Learning for Stance Detection in the Context of Checking Fake News. *Journal of Data and Information Quality*, 11(3): Article No. 14. <https://doi.org/10.1145/3287763>
- [7] Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. 2019. The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *Journal of Data and Information Quality*, 1(3): Article No. 10. <https://doi.org/10.1145/3309699>
- [8] A. Sally. (2020). What Are Deepfakes and How Are They Created?. Retrieved from <https://spectrum.ieee.org/tech-talk/computing/software/what-are-deepfakes-how-are-they-created>
- [9] S, Aliaksandr., L, Stephane., T, Sergey., R, Elisa., S, Nicu. First Order Motion Model for Image Animation. Retrieved from <https://arxiv.org/pdf/2003.00196.pdf>
- [10] T. T. Nguyen, C. M. Nguyen, D. T. Nguyen, D. T. Nguyen, and S. Nahavandi, "Deep Learning for Deepfakes Creation and Detection," *arXiv preprint arXiv:1909.11573*, 2019
- [11] Y. Mirsky and W. Lee, "The Creation and Detection of Deepfakes: A Survey," *arXiv preprint arXiv:2004.11138*, 2020.
- [12] Collaboratory – Google Frequently Asked Questions. (n.d.). Google Research. Retrieved from <https://research.google.com/collaboratory/faq.html>
- [13] A. Rössler, D. Cozzolino, L. Verdoliva, C., J. Thies, M. Nießner, "FaceForensics++: Learning to Detect Manipulated Facial Images", *arXiv preprint arXiv:1901.08971v3*, 2019
- [14] Justus Thies, Michael Zollhofer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. Face2Face: Real-Time Face Capture and Reenactment of RGB Videos. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2387–2395, June 2016.

- [15] Faceswap. [https://github.com/ MarekKowalski/FaceSwap/](https://github.com/MarekKowalski/FaceSwap/), Accessed: 2021.10.02.
- [16] Deepfakes github. [https://github.com/ deepfakes/faceswap](https://github.com/deepfakes/faceswap). Accessed: 20210-01-29.
- [17] Justus Thies, Michael Zollhofer, and Matthias Nießner. De-ferred neural rendering: Image synthesis using neural textures. ACM Transactions on Graphics 2019 (TOG), 2019
- [18] Rayhane Mama, Sam Shi, “Towards Deepfake Detection That Actually Works”, 2020 Square, Inc. ATLAS, DESSA

# Review of software architecture patterns in traffic systems

Z. Stojanov\*, G. Jotanovic\*\*, G. Jausevac\*\* and D. Perakovic\*\*\*

\* University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia

\*\* University of East Sarajevo, Faculty of transport and traffic engineering, Doboj, Republic of Srpska,  
Bosnia and Herzegovina

\*\*\* University of Zagreb, Faculty of Transport and Traffic Sciences, Zagreb, Croatia

zeljko.stojanov@uns.ac.rs, gordana.jotanovic@sf.ues.rs.ba, goran.jausevac@sf.ues.rs.ba, dperakovic@fpz.unizg.hr

**Abstract - Technical systems used in traffic are generally complex socio-technical systems, with variety of technical, organizational, and human issues to be considered during system design. Software components are key parts of these systems that provide data collecting, processing and storage, as well as the presentation of relevant information to stakeholders. During the design of traffic systems, the most suitable software architecture pattern should be selected to provide the best performance for the required functionalities. This paper presents a review of studies related to use of software architecture patterns in traffic systems published in period from 2010 to 2021. Results indicate the use of variety architecture patterns that are constrained with the used hardware and other equipment, as well as adjustment of common patterns to the needs of the system. At the other side, the major use of common software patterns is in the field of monitoring and controlling various traffic conditions. Insight into existing solutions in traffic systems may be helpful for design of new solutions and improvement of the existing ones.**

## I. INTRODUCTION

Traffic systems can be classified as complex socio-technical system, viewed as a collection of social and technical elements integrated to achieve specific behaviour. According to Sommerville [1], socio-technical systems can be viewed as a socio-technical stack with a layered structure in which software engineering part includes operating system, communication and data management, software applications and business processes. Understanding, design, operation, and maintenance of these systems require a broad knowledge of many technical and social disciplines [2]. A pragmatic approach to the engineering of socio-technical systems assumes the gradual introduction of socio-technical considerations into procurement and development processes of software parts of the systems [3]. Before designing a software solution as a part of complex traffic system, it is necessary to fully understand the domain and the stated problem, which will lead to the most appropriate design and the selection of the most suitable software architecture. Since traffic systems can introduce problems of resource scarcity, climate change and environmental degradation, the selection of the most efficient technological solution for both hardware and

software components becomes even more important for relieving these problems [4].

Complex socio-technical systems use variety of software architecture patterns, such as layered architecture, event-driven architecture, microkernel architecture, service-oriented architecture, microservices architecture, domain-driven architecture, resource-oriented architecture, or space-based architecture [5][6][7]. Determining the most suitable software architecture requires understanding the functioning of each component and module in the system. Without a formal architecture specification, designed software systems are tightly coupled, brittle, and difficult to change [6]. Adoption of architecture patterns enables more reliable and accurate defining characteristics and behaviour of software applications, which ensures better response to the proposed business requirements and constraints [8], as well as more efficient maintenance and support for controlled evolution [9][10]. Russo and Ciancarini [11] suggested development of antifragile software systems based on fine grained architecture resilient on faults of the system. Taibi et al. [12] presented a systematic mapping study on recent research trends on microservice software architecture patterns, while Pozdniakova and Mazeika [13] presented a systematic literature review of contemporary cloud-ready software architectures.

Traffic systems are built from variety of technical components, such as sensors, electrical and mechanical devices, communication devices, hardware components and software components. All these components are usually distributed in a larger area of interest for the observed traffic issues. This complex distribution of system components requires complex pattern for organizing software elements in the systems. The recent trend is design of smart traffic systems (e.g., traffic congestion management, smart parking systems, smart transportation, etc.) as integral parts of more complex smart city systems [14], which are based on contemporary technologies such as cloud-based systems, big data, IoT, sensing technologies, and communication technologies. Software components are distributed in all segments of these systems, and their organization within well-established architecture patterns is essential for overall

system performance. Based on a survey with researchers and experts from industry, Banijamali et al. [15] identified that event-driven and service-oriented architectures are dominant cloud-based architecture patterns in automotive domain, while multi-layered architecture, client-server architecture, pipeline and microservice patterns are less attractive for implementations. Literature reviews were performed on smart traffic light management systems [16], smart vehicles and elderly drivers [17], or intelligent transport systems for urban smart mobility [18]. However, based on our literature search, we have not detected any literature review on software architecture patterns in traffic systems, which motivated our research.

Based on the previous observations, and our experience with complex traffic systems and software architectures [19][20][21][22][23], the objective of this paper is to provide review of the use of software architecture patterns in traffic systems. The following research questions are proposed for the literature review:

RQ1: *Which architecture patterns are used in traffic systems?*

RQ2: *In which types of traffic systems are implemented software architecture patterns?*

The rest of the paper is structured as follows. The second section presents a literature review study with details on the method and results. The third section presents discussion of benefits and limitations of the study, while the last section contains conclusions and remarks for further research.

## II. LITERATURE REVIEW STUDY

With the aim to conduct a planned systematization of research evidence on using software architecture patterns

in traffic systems, literature review guideline for systematic literature review (SLR) [24][25] was selected and simplified for this study. Since SLRs show better effectiveness and performances than informal literature reviews in all phases of literature review [26], the decision in designing the review was to make it as much formal as possible. Since it is a part of a larger study aimed at investigating the use of software architecture patterns in complex socio-technical systems, this study is a pilot study with simplified protocol for conducting literature review and systematization of findings based on the proposed research question. The simplification relates to constraining the search only to Google Scholar, while search in other widely used scientific publishing databases (ScienceDirect, Springer, IEEE, ACM, Wiley) was left for future systematic literature reviews.

### A. Method

The review method roughly follows a general review approach proposed in [24], with the following main phases: (1) planning – aimed at defining a protocol for conducting review, (2) conducting – implementation of a proposed protocol to get results, and (3) reporting – production of a document with a technical report on the review.

The review protocol contains description of used keywords and construction of search string, selection of databases for search, selection of the period for search, and description of inclusion/exclusion criteria for the studies to be analyzed. Keywords for searching were selected based on the proposed research questions without considering possible synonyms that can enhance search and provide more reliable results. The composed search string is:

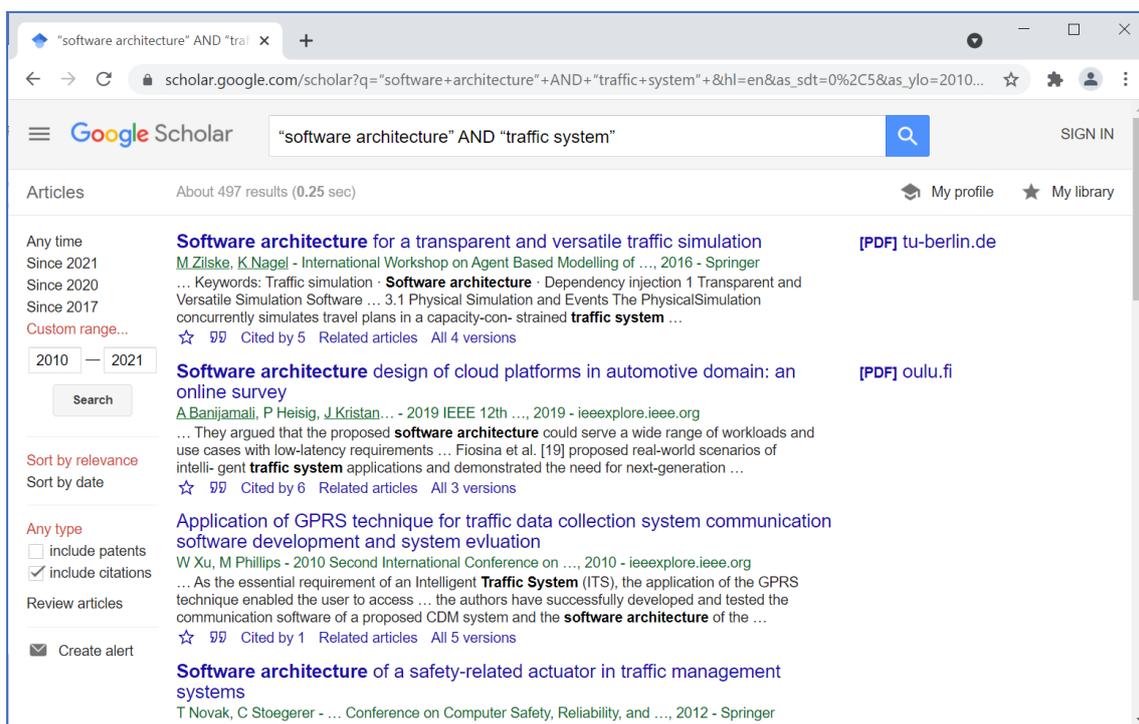


Figure 1. Search results for the constructed search string

“software architecture” AND “traffic system”

The search was performed on Google Scholar. The inclusion criteria for selecting the primary studies (PS) for further analysis are: (1) PS reports real experience within a case study, (2) PS was published in the period from 2010 to 2021, (3) PS is written in English, and (4) PS is available for downloading. Exclusion criteria for studies are: (1) conceptual and review papers are excluded, and (2) simulation studies are excluded.

The search that returns 497 results in Google Scholar is presented in Fig. 1. The assessment of the search results included the first 70 titles, since after the 50th title, majority of studies do not relate to the research topic.

The first and quick review of search results was based on reviewing titles, abstracts, and keywords. During this first review, 23 papers (PS candidates) were selected for the detailed analysis of the text. The reasons for exclusions of the papers were: study relates to simulation and there is no real implementation, study topic does not match research objectives (studies related to computer networks traffic), study is literature review, and the paper is not available for download.

For the selected 23 potential PSs, the second-round review included detailed reading of the whole paper. After inclusion/exclusion assessment 19 PSs were identified. The included PSs are listed in Tab. 1.

TABLE I. PRIMARY STUDIES RELATED TO THE RESEARCH OBJECTIVES

PS01	Xu, W., & Phillips, M. (2010, April). Application of GPRS technique for traffic data collection system communication software development and system evaluation. In 2010 Second International Conference on Computer and Network Technology (pp. 402-406). IEEE.
PS02	Aziz, M. W., Musharaf, U., & Sayyed, A. (2021). Towards a Software Architecture for Internet of Things based System of Systems. <i>International Journal of Emerging Trends in Engineering Research</i> , 9(3), 231-240.
PS03	Ippolito, C. A., Krishnakumar, K., Stepanyan, V., Chakrabarty, A., & Baculi, J. (2019). An Autonomy Architecture for High-Density Operations of Small UAS in Low-Altitude Urban Environments. In 2019 AIAA Modeling and Simulation Technologies Conference. San Diego, CA. Jan (Vol. 2109).
PS04	Chih-Ju, C., Sheng-Hao, S., Kuo-Hsiung, T., & To-Cheng, L. (2015, May). A novel SCADA system design and application for intelligent traffic control. In The 27th Chinese Control and Decision Conference (2015 CCDC) (pp. 726-730). IEEE.
PS05	Zeng, P., Wang, X., Li, H., Jiang, F., & Doss, R. (2020). A Scheme of Intelligent Traffic Light System Based on Distributed Security Architecture of Blockchain Technology. <i>IEEE Access</i> , 8, 33644-33657.
PS06	Consiglio, M., Duffy, B. J., Balachandran, S., Glaab, L., & Munoz, C. (2019). Sense and avoid characterization of the independent configurable architecture for reliable operations of unmanned systems. Thirteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2019)
PS07	Wang, S., Hou, Y., Gao, F., & Ji, X. (2016, December). A novel IoT access architecture for vehicle monitoring system. In 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT) (pp. 639-642). IEEE.
PS08	Serrano, D., Baldassarre, T., & Stroulia, E. (2016, December). Real-time traffic-based routing, based on open data and open-source software. In 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT) (pp. 661-665). IEEE.
PS09	Nugra, H., Abad, A., Fuertes, W., Galarraga, F., Aules, H., Villacis, C., & Toulkeridis, T. (2016, September). A low-cost IoT application for the urban traffic of vehicles, based on wireless sensors using GSM technology. In 2016 IEEE/ACM 20th International Symposium on Distributed Simulation and Real Time Applications (DS-RT) (pp. 161-169). IEEE.
PS10	Gong, Y., Rimba, P., & Sinnott, R. (2017, December). A big data architecture for near real-time traffic analytics. In <i>Companion Proceedings of the 10th International Conference on Utility and Cloud Computing</i> (pp. 157-162).
PS11	Horcas, J. M., Monteil, J., Bouroche, M., Pinto, M., Fuentes, L., & Clarke, S. (2018). Context-dependent reconfiguration of autonomous vehicles in mixed traffic. <i>Journal of Software: Evolution and Process</i> , 30(4), e1926.
PS12	Misrak, D. (2015). Web Based Transport Monitoring And Traffic System For Adama Transport Agency (Doctoral dissertation, ASTU).
PS13	Sinnott, R. O., Morandini, L., & Wu, S. (2015, December). SMASH: A cloud-based architecture for big data processing and visualization of traffic data. In 2015 IEEE International Conference on Data Science and Data Intensive Systems (pp. 53-60). IEEE.
PS14	Shakil, A., & Juric, R. (2020). Semantic Management of Urban Traffic Congestion. In <i>Proceeding of the 53rd Hawaii International Conference on System Sciences (HICSS 2020)</i> .
PS15	Pendleton, S., Chong, Z. J., Qin, B., Liu, W., Uthaicharoenpong, T., Shen, X., ... & Frazzoli, E. (2014, October). Multi-class driverless vehicle cooperation for mobility-on-demand. In <i>Intelligent Transportation Systems World Congress (ITSWC)</i> .
PS16	Asensio, A., Trasviña-Moreno, C., Blasco, R., Marco, Á., Casas, R. (2015) Wireless Sensor Networks in traffic management systems. <i>Proceedings of the 6th International Conference on Applied Informatics and Computing Theory (AICT '15)</i> , pp. 60-68. Salerno, Italy. 2015.
PS17	Kar, A., Syamal, S., Chatterjee, S., Basu, A., Saha, H. N., & Choudhuri, S. (2019, October). SPEED CONTROLLING & TRAFFIC MANAGEMENT SYSTEM (SCTMS). In 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 1058-1063). IEEE.
PS18	Shafie, A. A., Ali, M. H., Hafiz, F., & Ali, R. M. (2011). Smart video surveillance system for vehicle detection and traffic flow control. <i>Journal of Engineering Science and Technology</i> , 6(4), 469-480.
PS19	Njama-Abang, O., & Edim, A. E. (2017) Design Of A Road Traffic Monitoring And Report System. <i>Journal of Multidisciplinary Engineering Science Studies</i> , 3(6), 1811-1820.

**B. Results**

The primary studies were read in detail to get answers on the proposed research questions, which is presented in the following subsections.

**RQ1: Which architecture patterns are used in traffic systems?**

Insight into the sections related to design of the systems or implementation details, revealed the use of variety of software architecture patterns: client server [PS01,PS12], microservices [PS02], monolithic [PS04], distributed components [PS05,PS11,PS18], distributed publish-subscribe software [PS06], Master-slave [PS07], layered [PS08,PS09,PS10,PS13,PS14,PS15,PS17,PS19], and service-oriented architectures [PS06,PS16]. It is evident that layered structure is dominant pattern for organizing software parts in traffic systems, while service-oriented architectures gain attention recently.

In study [PS14], the authors reported combination of layered and component-based architecture.

The authors of study [PS03] did not report any architecture pattern. The architecture of the software system is described, but there is no formal statement about the used architecture pattern, nor it is possible to draw any conclusion about an architectural pattern based on the text and images in the paper.

**RQ2: In which types of traffic systems are implemented software architecture patterns?**

Based on the description of implementation details of PSs, the types of traffic systems in which software architecture patterns were implemented are presented in Tab. 2. Implementation details column provides some details on the specific use for each type of traffic system. Most implementations were in traffic control systems. Another important type of implementation relates to vehicles in transport systems, including vehicles with driver and without driver. Traffic systems that include unmanned vehicles gained popularity in recent years, which includes road and air vehicles.

**III. DISCUSSION**

Results related to software architecture patterns indicate that system designers selected patterns that are the most suitable for the system purpose, which resulted in the use of a large variety of architecture patterns. It is evident that in systems related to monitoring and controlling

traffic the dominant pattern is layered architecture which enables distribution of software components in different layers (sensing layer, middleware layer with servers, user or application layer).

**A. Benefits and implications**

This review study provides short and precise guidelines for conducting preliminary literature review as a pilot study before engaging with more detailed and comprehensive SLRs. This study, therefore, may be of interest for PhD students and researchers that plan to systematize literature for specific topic based on the proposed research questions.

For researchers in the field of software architectures in traffic systems, this study can be used as starting point, from which some promising directions, such as selection of architecture quality attributes or reengineering of old systems to service oriented architectures, can be pursued.

In addition, experience reported in reviewed studies can help system engineers and designers in avoiding some typical obstacles and constraints in selecting the optimal architecture solution for software part of complex traffic systems.

**B. Limitations and validity**

Despite the reported benefits, this study has some limitations that treats validity of the study findings, which should be discussed to increase the study validity [27]. Internal validity was increased through detailed description of all research phases and steps, which enable checking the reported study findings.

The first limitation relates to constraining the search only to Google Scholar, which listed limited set of literature sources. For example, seven PSs, and majority of listed sources are published in IEEE Xplore digital library, which strongly indicates that future literature reviews should include this library. In addition, other digital libraries, such as ACM, Springer or ScienceDirect, with large number of journals and conference series should be also considered.

The next limitation relates to constraining the search string only to combination of two phrases “software architecture” and “traffic system”, which certainly reduces the number of studies that can be included in the analysis. Therefore, identifying synonyms for selected keywords (e.g., “software structure” or “architecture pattern” for

TABLE II. TYPES OF TRAFFIC SYSTEMS FOR THE IDENTIFIED SOFTWARE ARCHITECTURE PATTERNS

Type of traffic system	Primary studies	Implementation details
Vehicles in traffic	PS01, PS04, PS07, PS11, PS17	Intelligent control and adaptive control of vehicles, and speed control in electrical vehicles.
Info for citizens	PS02, PS12	Traffic route info, vehicle registration and licensing, driver registration and licensing.
Control of unmanned vehicles	PS03, PS06, PS15	Operation in urban areas, sensing and avoiding obstacles, cooperation of driverless vehicle fleet.
Traffic control	PS05, PS08, PS09, PS10, PS13, PS14, PS16, PS18, PS19	Intelligent lights, transportation, traffic flow, visualization of traffic data, urban traffic congestions, occupation of spaces in cities.

”software architecture”) is necessary for a more detailed and systematic review of the literature.

These limitations are signs that should guide future systematic literature reviews that can provide more comprehensive and reliable findings.

#### IV. CONCLUSION

Presented review study provides insights into the recent research (period from 2010 to 2021) related to use of software architecture patterns in traffic systems. The findings are very helpful for the authors of the paper since their research focus is on analyzing, modelling, and designing traffic systems as complex socio-technical systems. Benefits of this study for other researchers, especially PhD students, and practitioners from economy are discussed in the third section.

Future work will be directed towards modelling software architectures for specific traffic systems, such as control of unmanned aerial vehicles (UAVs), air pollution caused by the traffic, and monitoring of traffic congestions and parking spaces in urban settings. Future work will also include development and implementation of design principles that are based on considering software architecture quality attributes that will increase system performance and maintainability.

#### ACKNOWLEDGMENT

Ministry of Education, Science and Technological Development, Republic of Serbia, supports this research under the project "The development of software tools for business process analysis and improvement", project number TR32044.

#### REFERENCES

- [1] I. Sommerville, *Software Engineering* 9th ed. Boston, MA, USA: Addison Wesley, 2011.
- [2] Z. Stojanov, J. Stojanov, G. Jotanovic, D. Dobrilovic, "Weighted networks in socio-technical systems: concepts and challenges". In Proceedings of the 2nd International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2020), pp. 265-276. Irkutsk, Russia. 2020. CEUR Workshop Proceedings, Vol-2638.
- [3] G. Baxter, I. Sommerville, "Socio-technical systems: From design methods to systems engineering". *Interacting with Computers*, vol. 23, no. 1, pp. 4-17. 2011. doi: 10.1016/j.intcom.2010.07.003.
- [4] L. Fuenschilding, B. Truffer, "The structuration of socio-technical regimes: Conceptual foundations from institutional theory". *Research Policy*, vol. 43, no. 4, pp. 772-791, 2014. doi: 10.1016/j.respol.2013.10.010.
- [5] D. Duggan, *Enterprises software architecture and design: entities, services, and resources applications*. Hoboken, NJ, USA: John Wiley & Sons, 2012.
- [6] M. Richards, *Software Architecture Patterns: Understanding Common Architecture Patterns and When to Use Them*, 1st ed. Sebastopol, CA, USA: O'Reilly Media, 2015.
- [7] J. Ingenu, *Software Architect's Handbook*. Birmingham, UK: Packt Publishing, 2018.
- [8] L. Bass, P. Clements, R. Kazman, *Software architecture in practice*, 4th ed., SEI series in software engineering. Upper Saddle River, NJ, USA: Addison-Wesley Professional, 2021.
- [9] PO. Bengtsson, N. Lassing, J. Bosch, H. van Vliet, "Architecture-level modifiability analysis (ALMA)", *Journal of Systems and Software*, vol. 69, no. 1-2, 1 pp. 129-147. 2004. doi: 10.1016/S0164-1212(03)00080-3.
- [10] H. Unphon, Y. Dittrich, "Software architecture awareness in long-term software product evolution". *Journal of Systems and Software*, vol. 83, no. 11, pp. 2211 - 2226. 2010. doi: 10.1016/j.jss.2010.06.043.
- [11] D. Russo, P. Ciancarini, "Towards Antifragile Software Architecture". *Procedia Computer Science*, vol. 109, pp. 929-934. 2017. doi: 10.1016/j.procs.2017.05.426.
- [12] D. Taibi, V. Lenarduzzi, C. Pahl, "Architectural Patterns for Microservices: A Systematic Mapping Study". In Proceedings of the 8th International Conference on Cloud Computing and Services Science - CLOSER, pp. 221-232. 2018. doi: 10.5220/0006798302210232.
- [13] O. Pozdniakova and D. Mazeika, *Systematic Literature Review of the Cloud-ready Software Architecture*. *Baltic Journal of Modern Computing*, vol. 5, no. 1, pp.124-135, 2017. doi: 10.22364/bjmc.2017.5.1.08.
- [14] B. N. Silva, M. Khan. K. Han, "Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities". *Sustainable Cities and Society*, vol. 38, pp. 697-713. 2018. doi: 10.1016/j.scs.2018.01.053.
- [15] A. Banijamali, P. Heisig, J. Kristan, P. Kuvaja, M. Oivo, "Software architecture design of cloud platforms in automotive domain: an online survey". In Proceedings of the IEEE 12th Conference on Service-Oriented Computing and Applications (SOCA), pp. 168-175. Kaohsiung, Taiwan. 2019. doi: 10.1109/SOCA.2019.00032.
- [16] A. A. Magableh, M. A. Almahadmeh, N. Alsrehin, A. F. Klaib, "Smart Traffic Light Management Systems: A Systematic Literature Review". *International Journal of Technology Diffusion (IJTD)*, vol. 11, no. 3, pp. 22-47. Doi: 10.4018/IJTD.2020070102
- [17] I. Rhiu, S. Kwon, S. Bahn, M. H. Yun, Y. Wooyeon, "Research Issues in Smart Vehicles and Elderly Drivers: A Literature Review", *International Journal of Human-Computer Interaction*, vol. 31, no. 10, pp. 635-666, 2015. doi: 10.1080/10447318.2015.1070540.
- [18] R. Mangiaracina, A. Perego, G. Salvadori, A. Tumino, "A comprehensive view of intelligent transport systems for urban smart mobility", *International Journal of Logistics Research and Applications*, vol. 20, no. 1, pp. 39-52, 2017. doi: 10.1080/13675567.2016.1241220.
- [19] G. Jotanovic, J. Stojanov, D. Perakovic, Z. Stojanov, G. Jausevac, M. Stojicic, "Internet of Things Technology Purposeful For Monitoring Road Traffic Air Pollution". In Proceedings of the 10th International conference on Applied Internet and Information Technologies (AIIT2020), pp. 53-57. Zrenjanin, Serbia. 2020.
- [20] D. Dobrilović, V. Brtko, G. Jotanović, Z. Stojanov, G. Jauševac, M. Malić, "Architecture of IoT system for smart monitoring and management of traffic noise". In Proceedings of the 5th EAI International Conference on Management of Manufacturing Systems (EAI MMS 2020), pp. 251-266. 2020. DOI: 10.1007/978-3-030-67241-6\_21.
- [21] D. Dobrilovic, V. Brtko, G. Jotanović, Z. Stojanov, G. Jauševac, M. Malić, "The urban traffic noise monitoring system based on LoRaWAN technology". *Wireless Networks*, 2021. DOI: 10.1007/s11276-021-02586-2.
- [22] G. Jausevac, D. Dobrilovic, V. Brtko, G. Jotanovic, D. Perakovic, Z. Stojanov, "Smart UAV Monitoring System for Parking Supervision". In Proceedings of International Conference on Future Access Enablers of Ubiquitous and Intelligent Infrastructures, pp. 240-253. 5th EAI International Conference, FABULOUS 2021, Virtual Event, 2021. Springer, Cham. Doi: 10.1007/978-3-030-78459-1\_18.
- [23] Z. Stojanov, D. Dobrilovic, "Software Architecture Quality Attributes of a Layered Sensor-Based IoT System". In Proceedings of the 4th Scientific-practical Workshop Information Technologies: Algorithms, Models, Systems (ITAMS 2021), pp. 66-74. Irkutsk, Russia, 2021.
- [24] B.A. Kitchenham, *Procedures for Undertaking Systematic Reviews*, Joint Technical Report, Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd. (0400011T.1). Keele, UK, 2004.
- [25] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, S. Linkman, "Systematic literature reviews in software

- engineering - A systematic literature review". *Information and Software Technology*, vol. 51, no. 1, pp. 7-15. doi: 10.1016/j.infsof.2008.09.009.
- [26] M. Niazi, Do Systematic Literature Reviews Outperform Informal Literature Reviews in the Software Engineering Domain? An Initial Case Study. *Arabian Journal for Science & Engineering*, vol. 40, no. 3, pp. 845-855. doi: 10.1007/s13369-015-1586-0.
- [27] H. K. Wright, M. Kim, D. E. Perry, Dewayne, "Validity Concerns in Software Engineering Research". In proceedings of the FSE/SDP Workshop on Future of Software Engineering Research - FoSER '10, pp. 411-414. Santa Fe, New Mexico, USA. 2010. doi: 10.1145/1882362.1882446.

# The Application of Semi-Linguistic Summaries in Traffic Data Analysis

V. Brtko<sup>\*</sup>, M. Sisak<sup>\*\*</sup>, V. Makitan<sup>\*</sup>, G. Jotanović<sup>\*\*\*</sup>, G. Jauševac<sup>\*\*\*</sup>

<sup>\*</sup>University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia

<sup>\*\*</sup>West Regional Development Agency, Timisoara, Romania

<sup>\*\*\*</sup>Faculty of Transport and Traffic Engineering, Dobož, Republic of Srpska, Bosnia and Herzegovina  
vbrtko@tfzr.uns.ac.rs, mirasisak26@gmail.com, vesna@tfzr.uns.ac.rs, gordana.jotanovic@sf.ues.rs.ba,  
goran.jausevac@sf.ues.rs.ba

**Abstract - The paper deals with the qualitative analysis of real-life traffic data. The method of synthesis of semi-linguistic summaries is described, and the conducted experiment illustrates the application of this method on a real-life data sample. It is shown how the method enables the detection of hidden patterns in the data, and the conclusions refer to the direction of vehicle movement. Furthermore, the prediction of the direction of movement leads to the possible detection of traffic jams, problems with parking space and taking appropriate action to eliminate them.**

## I. INTRODUCTION

A significant increase in the number of motor road vehicles entails a number of negative effects on the environment and people. We are witnessing almost daily traffic jams, traffic accidents and various situations that are not regular or are very unpleasant. Traffic jams have a negative impact on resource consumption, especially time and fuel, but there are also many other resources whose consumption is more difficult or impossible to measure. Therefore, traffic management problems and traffic congestions are problems that are being solved, more or less successfully in many urban areas, especially in multi-million cities.

There are differences in the way these problems are addressed, at the level of cities, regions or countries, but in many cases traffic problems are solved through modern information technologies (IT). New IT includes the application of various types of sensors and sensor devices, drones, computer networks, cameras and other devices that significantly contribute to reducing traffic problems. The IT technologies used, in addition to devices, include algorithms for data analysis and decision making. These algorithms are often based on Artificial Intelligence (AI), and in many cases Machine Learning (ML) techniques are used.

The importance of implementing new IT approaches should, above all, make traffic safer, reduce the number of injuries and fatalities, but also provide less time/energy consumption and reduce traffic congestion. The number of vehicles, vehicle types, time of day, various weather readings, etc. are often used as input parameters of algorithms. By data analysis, classification and prediction, the effects of improving the safety of the traffic system,

better management of the parking space of vehicles when they do not participate in traffic, less pollution, etc. can be achieved.

The paper deals with the analysis of traffic data, such as vehicle count by type, the direction of vehicle movement, at several locations within and near the urban area. A qualitative type of analysis, the so-called semi-linguistic summaries, was used, with the help of which it is possible to observe patterns in the data.

Real-life data originating from the city of Helsinki, Finland were used. The method is described primarily from a theoretical point of view, with a few practical details. The analysis was performed, and the paper describes only a part of the conclusions of the analysis with the aim of presenting this type of qualitative analysis to readers. The main goal of the described research is to consider the possibilities provided by this type of qualitative analysis of traffic data.

The rest of the paper is structured as follows: Section two provides an insight into selected previous research. Section three contains a description of the data and methods used in this research. Section four presents the research results and the discussion of the results. Finally, the conclusion summarizes the results, highlights their importance and identifies some of the possible directions for future research.

## II. PREVIOUS RESEARCH

Having in mind the importance of transport, of course the number of conducted researches is impressive, and the material in the form of written works is extremely extensive. In this paper, only some of the available sources are analyzed, namely those that correspond in form and content to the research described in this paper. The aim is to gain insight into some of the techniques of analyze traffic data, as well as to gain insight into the importance of analysis and possible applications.

Center research issues in [1] are traffic violations and factors associated with them. The Automated Enforcement System (AES) was tested as an additional way to control traffic. The probability of violations and predictions were made by the random forest algorithm. On the basis of this, a model was formed that includes a number of parameters. Factors influencing traffic

violations are treated, with special reference to regional differences, i.e. the fact that traffic violations are not treated or defined the same in all locations (countries or regions) is taken into account. Intersections were especially considered, and it was determined that the most common traffic violations are: traffic sign or marking violation, stopping beyond the stop line and red-light running. The research conducted in China is described, and the AES system was in operation for 85 days, during which it generated data. All in all, 12 types of possible traffic violations have been identified, and the distribution of traffic violations is given in the paper. The logic model contains the probability equation for traffic violence and the Python sklearn package was used to estimate the model coefficients [1, 2]. There are four factor groups: time, space, traffic and weather. Every group consists of multiple elements. The random forest algorithm is described by a pseudo-code, while performance measures are usual: Accuracy, recall, specificity and precision. F1-score and G-mean were also used. The results showed that the random forest algorithm outperforms logistic regression algorithm.

In [3], traffic accident detection from the aspect of social networking data was investigated. The goal of reducing injuries and fatalities can be obtained using an advanced data classification that uses traffic data. The model was formed under the assumption that the data collected by the sensors are not sufficient, and that it is necessary to use social network-based data. The proposed model reportedly outperforms existing models, achieving an accuracy of 97%. A new smart framework was proposed on order to analyze social networking data including sentiment analysis. The system architecture for accident event detection and analysis is proposed, system uses twitter and facebook data. Deep Learning (DL) models were used in order to classify data. However, this approach is limited by the rights of social network users, and the issue of personal security can be raised, as well as a number of privacy and end user rights issues. Nevertheless, it is planned to further develop the system and to use more refined techniques.

In addition to the ones shown, there are a number of studies that use manual class labeling methods, various word embedding models and classifiers [4, 5, 6, 7].

### III. METHODS

The conclusions of previous research assume that the data collected by sensors are not always sufficient in the prediction or analysis of traffic accidents or violations. The use of social networking data can be useful, but it raises the issue of violating the privacy of users, as well as many other issues related to the rights of users of social networks. Therefore, it was decided to use existing data collected from sensors in this research, which will be analyzed by a qualitative method, which can overcome the barrier of quantitative analysis and various statistical methods. The mentioned barrier often refers to insufficiently descriptive results of the analysis, and the absence of "deeper" conclusions and analysis of situations that occur less frequently.

#### A. Data

The data on which the method of analysis was applied are publicly available via Internet (<https://hri.fi/data/>), and refer to traffic data collected at several key points in the city of Helsinki, Finland. The data includes the following fields:

- piste: ID of the counting point
- nimi: name / location of the counting point
- x\_gk25 x coordinate in ETRS-GK25 coordinate system
- y\_gk25 y coordinate in ETRS-GK25 coordinate system
- suunta: direction (direction 1 is towards city centre; D1-D13 towards west)
- aika: time (hour; at 6am-9am and 3pm-6pm 15 minutes)
- vuosi: year
- ha (henkilöautot): passenger cars
- pa (pakettiautot): vans
- ka (kuorma-autot): trucks
- ra (rekka-autot): lorries
- mp (moottoripyörät): motorbikes
- rv (raitiovaunut): trams
- autot: cars

ID was used in this research, while location name and GPS coordinates were not used. The data were written in a csv format that is readable from any text editor or using MS Excel. The influence of the mentioned parameters on the direction was observed, which can be: towards the city center (1) or towards the west (2). Data have been collected since 2010, using the methodology described on the aforementioned web-site. There are 47,880 records in total, but due to the lack of resources only 5,284 were actually used. Namely, some of the algorithms required for data processing or data preparation are not feasible on an optimal or sub-optimal computer configuration. One example of this algorithm is the Boolean Reasoning Algorithm (BRA), which serves to discretize continuous values and which cannot be executed because it is extremely demanding in terms of memory. Discretization was performed with the Equal Frequency Binning (EFB) algorithm, which is much less demanding from the point of view of time and memory consumption. This means that continuous (real) attribute values were presented by ranges.

#### B. Algorithms

As mentioned earlier, the technique of qualitative analysis was used in order to determine the influence of 10 parameters (attributes) to direction parameter. The question is how do the values of 10 attributes affect the value of the direction attribute? Technique based on linguistic summarization was applied in order to find pre-defined patterns in data and verbalize those patterns. Linguistic Summaries (LS) were proposed by R. R. Yager [8] as a mean for getting an insight to deeper meaning of dataset. LSs were revised by L. A. Zadeh [9] who introduced term "protoform". The concepts of LSs and protoforms have been upgraded and used in different domains by several authors, (see e.g. [10, 11]). The process is broken down into five steps shown in Fig. 1.

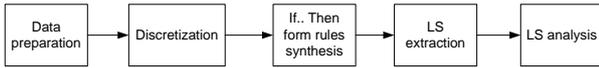


Figure 1. LS extraction and analysis process

Data preparation involves dealing with missing values, changing the format of data records, etc. Discretization, as already stated, converts continuous values into ranges. Here it is important to choose an algorithm that can be executed on a hardware/software platform, due to the high consumption of resources.

After data preparation step and data discretization, it is necessary to generate the rules in If ... Then form. The indiscernibility relation that is the basis of Rough Set Theory [12, 13] was used. In addition to indiscernibility relation, the process of generating rules in the If ... Then form was executed by application of indiscernibility graph [14, 15]. Let  $E=\{e_1, e_2, \dots, e_j\}$  be a finite set of elements while  $A=\{a_1, a_2, \dots, a_n\}$  is a finite set of parameters or attributes so that each element from  $E$  is characterized by attributes from  $A$ . If there are two sub-sets  $C$  and  $D$  so that  $C \subset A$ ,  $D \subset A$  and  $C \cap D = \emptyset$ , then attributes from  $D$  are called decision attributes while attributes from  $C$  are called condition attributes. In this research, there is a single attribute in  $D$ , that is direction attribute, while set  $C$  consists of all other attributes. Generally, the goal is to find dependence in form:  $C \Rightarrow D$ . This is a set of rules that describes dependences of direction attribute from set  $D$  and all other attributes from set  $C$ . Attributes from  $C$  are connected by AND logical operator. For a dataset an indiscernibility tree graph is formed. The attributes from set  $C$  (condition attributes) are arranged in an array, so that the most important attribute is at the beginning, and the remaining attributes are arranged according to the estimated significance. The significance of an attribute can be determined by subjective and objective methods, or a combination of the two, but it is usually the price of measuring the value of the attribute. In this study, the leading attribute is the ID of the location where the measurement is performed. By changing the meaning of the attributes, i.e. by changing the order of the attributes in the array, it is possible to generate different sets of If ... Then rules. An algorithm developed in the Java programming language classifies elements from the set  $E$  using indiscernibility relation and forms a graph. In the further procedure, the graph arcs form the If ... Then rules. Graph nodes are sets of elements from set  $E$  that satisfy the indiscernibility relation. Because the number of graph nodes is large, the Depth First Search (DFS) algorithm is used to generate the graph. Semi-linguistic summaries are formed according to tree-graph structure, and their form is:

If elements from set  $S \subset E$  have condition  $p$  then  $q$  of them have property  $s$ .

Condition  $p$  is described by the values of condition attributes connected by AND logical operator, while  $s$  is the value of decision attribute. Parameter  $q$  is a linguistic fuzzy quantifier. Given that there is only one linguistic quantifier, the form is semi-linguistic.

### C. System Tuning

A system that uses elements of RST, indiscernibility graph, and semi-linguistic summaries, is dependent on multiple parameters. Fuzzy quantifiers are same as in [11, 15], and are defined as shown in Fig. 2.

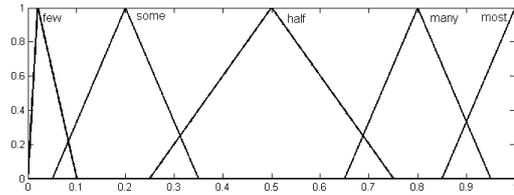


Figure 2. Definition of Fuzzy Quantifiers

Measures of confidence, coverage and, consequently truth value for each summary are calculated, as usual. Coverage value  $c$  is altered by a parametric S-shaped function  $f_{a,b}(c)$ , so small values are normalized, where  $a=0.02$  and  $b=0.15$ . Five features of semi-linguistic summaries are calculated: validity (number of records that are used to derivate summary), generality (how many data support a summary), usefulness (how useful is a summary), novelty (is there something unexpected) and simplicity (how many conditional attributes were used). It is possible to fine-tune the system by selecting fuzzy quantifiers, the values of parameters  $a$  and  $b$ , as well as determining the minimum values of validity, generality, usefulness, novelty and simplicity.

## IV. RESULTS AND DISCUSSION

By applying the indiscernibility graph, for a given order of conditional attributes: [id, time, year, passenger-cars, vans, trucks, lorries, motorbikes, trams, cars], a total of 2630 If ... Then rules were generated. Based on these rules, a total of 242 semi-linguistic summaries were generated for:  $\text{truth} \geq 0.75$ ,  $\text{validity} > 0.75$  and  $\text{simplicity} > 0.001$ . This setting of values that determine the system configuration leads to the summaries that are not general ( $\text{generality} = 0$ ), while the validity and novelty values are large. In other words, the semi-linguistic summaries that have been singled out describe interesting cases that are not supported by a large number of records. It is not possible to describe the analysis and conclusions of the analysis of all semi-linguistic summaries, but the description of the part of the  $\text{id}=\text{F01}$  location analysis for  $\text{year}=\text{2010}$  is given below. Seven semi-linguistic summaries were identified and presented in Tab. I.

TABLE I. SEMI-LINGUISTIC SUMMARIES

	time	vn	lo	mb	tra	ca	dir
1	AN	[8-26)	[4-*)	[1-4)	[*-1)	-	<b>A</b>
2	AN	[8-26)	[4-*)	[1-4)	[*-1)	[97-306)	<b>A</b>
3	BN	[26-*)	[4-*)	[1-4)	-	-	<b>A</b>
4	BN	[26-*)	[4-*)	[1-4)	[*-1)	-	<b>A</b>
5	BN	[8-26)	-	-	-	-	<b>B</b>
6	BN	[8-26)	[4-*)	-	-	-	<b>B</b>
7	BN	[8-26)	[4-*)	[1-4)	-	-	<b>C</b>

The following abbreviations were used: vn-vans, lorries, mb-motorbikes, tra-trams, ca-cars. Time of the day (time) have values: BN-Before noon [7:45h–15:45h) and AN-Afternoon [15:45h–24:00h). Direction (dir) is represented by: A-Many vehicles are going towards city center, B-Many vehicles are going towards west, C-Most vehicles are going towards west. The number of passenger-cars ranges from 79 to 250 for all cases, and the number of trucks is always greater than six, so these two attributes are not listed in the Table.

There are several ways of analysis, and the procedure itself does not have to be strictly defined. By analysis of Tab. I, the before-noon and afternoon traffic summaries at this location are clearly visible. The afternoon traffic regime is described using summaries 1 and 2, and it can be seen under what conditions many vehicles are going towards the city center. The before-noon traffic regime is described using summaries 3–7, the difference in the direction is, again clearly visible. Namely, the number of vans exceeds 25 in the case when many vehicles are moving towards the city center. The analysis also indicates when and which attributes are not significant for the assessment of vehicle's direction.

The analysis process can be continued, and in this sense, the presented results serve as an illustration of the possibility of applying semi-linguistic summaries.

## V. CONCLUSION

The previously described research shows how to use qualitative traffic data analysis. Specifically, semi-linguistic summaries based on real-life traffic data originating from the city of Helsinki, Finland were used.

This type of analysis enables the synthesis of summary conclusions that enable the observation of patterns hidden in the data, and in that sense these are the techniques used in the domains of data analysis and data mining. Specifically in the case of the presented research, it was determined that the number of vehicles per type at one of the locations of the city of Helsinki, in before-noon and afternoon hours, affects the direction in which many vehicles move. The values of the attributes according to which it is possible to determine whether many or most vehicles are moving towards the city center or towards the west are clearly distinguished. It is possible to single out the necessary attributes and attributes that are redundant. In this case, a "van" type vehicle count can be an indicator of the direction of movement.

The presented research results refer to one location, with a clear intention that a similar analysis can be performed at multiple locations within the city or in its surroundings. A prerequisite for performing this procedure is certainly the existence of sensors/cameras in various locations. The direction of vehicle movement indicates the creation of traffic jams, problems with the lack of parking spaces, etc. Analysis of this type enables a timely response in terms of traffic re-direction, changes in traffic signaling, etc.

Future directions of research will include the use of a better computer platform and different system settings.

## ACKNOWLEDGMENT

Ministry of Education, Science and Technological Development, Republic of Serbia financially supported this research, under the project number TR32044: "The development of software tools for business process analysis and improvement", 2011-2021.

## REFERENCES

- [1] Yunxuan Li, Meng Li, Jinghui Yuan, Jian Lu, Mohamed Abdel-Aty, "Analysis and prediction of intersection traffic violations using automated enforcement system data", *Accident Analysis and Prevention* 162 (2021) 106422.
- [2] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: Machine learning in Python. *The Journal of machine Learning research* 12, 2825–2830.
- [3] Farman Ali, Amjad Ali, Muhammad Imran, Rizwan Ali Naqvi, Muhammad Hameed Siddiqi, Kyung-Sup Kwak, Traffic accident detection and condition analysis based on social networking data, *Accident Analysis and Prevention* 151 (2021) 105973.
- [4] Cao, J., Zeng, K., Wang, H., Cheng, J., Qiao, F., Wen, D., Gao, Y., 2014. Web-based traffic sentiment analysis: methods and applications. *IEEE trans. Intell. Transp. Syst.* 15 (2), 844–853. <https://doi.org/10.1109/TITS.2013.2291241>.
- [5] Cao, D., Wang, S., Lin, D., 2018. Chinese microblog users' sentiment-based traffic condition analysis. *Soft Comput.* 22 (21), 7005–7014. <https://doi.org/10.1007/s00500-018-3293-8>.
- [6] Ali, F., El-Sappagh, S., Kwak, D., 2019a. Fuzzy ontology and LSTM-based text mining: a transportation network monitoring system for assisting travel. *Sensors (Switzerland)* 19, 2. <https://doi.org/10.3390/s19020234>.
- [7] Ali, F., Kwak, D., Khan, P., Islam, S.M.R., Kim, K.H., Kwak, K.S., 2017. Fuzzy ontologybased sentiment analysis of transportation and city feature reviews for safe traveling. *Transp. Res., Part C, Emerg. Technol.*, 77, <https://doi.org/10.1016/j.trc.2017.01.014>.
- [8] Yager, R. R. (1982). A new approach to the summarization of data. *Information Sciences*, vol. 28, pp. 69-86.
- [9] Zadeh, L. A. (2002). A prototype-centered approach to adding deduction capabilities to search engines—the concept of a protoform. *BISC Seminar, Berkeley, California, USA, Feb. 8*.
- [10] Kacprzyk, J., Zadrozny, S. (2009). Protoforms of Linguistic Database Summaries as a Human Consistent Tool for Using Natural Language in Data Mining. *Int. J. of Software Science and Computational Intelligence*, vol. 1(1), pp. 1-11.
- [11] Wu, D., Mendel, J. M., Joo, J. (2010). Linguistic summarization using IF-THEN rules. *IEEE International Conference on Fuzzy Systems, (FUZZ) 2010, Barcelona, Spain, 18-23, July*.
- [12] Pawlak, Z., Grzymala-Busse, J., Slowinski, R., Ziarko, W. (1995). *Rough Sets. Communications of the ACM*, vol. 38, pp. 88-95.
- [13] Pawlak, Z. (2007). Skowron, A. Rudiments of rough sets. *Journal of Information Sciences*, vol. 177, pp. 3–27.
- [14] Ognjenovic V., Brtko V., Jovanovic M., Brtko E., Berkovic I., "The representation of indiscernibility relation by graph", *IEEE 9th International Symposium on Intelligent Systems and Informatics, SISY*, pp. 91-94, Sept. 8-10, 2011, *IEEE Catalog Number: CFP1184C-CDR*, ISBN: 978-1-4577-1973-8, 8-10, doi: 10.1109/SISY.2011.6034298.
- [15] Brtko V., Makitan V., Brtko E., Dobrilovic D., Berkovic I., (2020), *LP-WAN Performance Analysis by Semi-Linguistic Summaries, AD HOC & SENSOR WIRELESS NETWORKS*, vol. 48, no. 1-2, pp. 145-165.

# Recommender systems for carer guidance

Trần Đức Thế<sup>1</sup>, Viktoria Kopylova<sup>1</sup>, Evgeny Cherkashin<sup>1,2,3</sup>, Nikita Lukyanov<sup>1</sup>

<sup>1</sup> Institute for Informational Technologies and Data Analysis,

National Research Irkutsk State Technical University, Irkutsk, Russia

<sup>2</sup> Laboratory of Complex Informational Systems,

Matrosov Institute for System Dynamics and Control Theory, Siberian Branch of Russian Academy of Sciences, Irkutsk, Russia

<sup>3</sup> Chair for Informational Technologies,

Institute for Mathematics and Information Technologies, Irkutsk State University, Irkutsk, Russia

eugeneai@icc.ru

**Abstract**—The research of the paper is devoted to present of an R&D of two recommender systems in carer guidance for Việt Nam middle school students and one for testing a paper [1] technique for data obtained in National research Irkutsk state technical university. The first system is built on the base of John Holland's six personality types theory and collecting university students' opinions on their study conditions, and the second one relates data from a social network and faculty enrollment documents. Both systems describe new user (an entrant or a pupil) properties by distances to student groups and then presents top- $N$  universities/departments as a recommendation. Techniques of overcoming "cold start" problem and other ones in this domain are considered.

**Index Terms**—recommender system, cold start, carer guidance, user group description, machine learning

## I. INTRODUCTION

In the countries with predominantly young population, such as Việt Nam, the demand of young specialists is high. While still a high school student, one have to make a decision about a career. Career choice depends on many factors such as personal interests, academic ability and employment opportunities available after the graduation. Therefore, choosing the right carer is not an easy task. In fact, quite often graduates remain unemployed for a long time or do not work in their specialties mastered in an educational institution. This leads to inefficient use of state funds, their irrational distribution. That's why the choice of an university after receiving a certificate of maturity is very important. Choosing a suitable university results in an increase students' study productively, as they work and practice with enthusiasm. Supporting the decision activities will give a better chance of achieving students' goals in the future.

In Tables 1 and 2, we present the result of interest assessment to career guidance and relevance of improving of the guidance system. Tables show that there is a real problem of carer guidance and middle school students are very interested in improving the corresponding activities.

At present, with the development of the Internet, the search for information about universities is sufficiently efficient. However, having a huge amount of information, selecting substantive information is a difficult task. Recommender systems (RS) have emerged as a decision support tool, providing

TABLE I  
INTEREST ASSESSMENT TO THE CARER GUIDANCE IN VIỆT NAM, HÀ TĨNH CITY

Interest level	Count of votes	Ratio, %
Very interested	218	51.9
Relatively interested	155	36.9
Less interested	36	8.6
No interest	11	2.6
<b>Total</b>	<b>420</b>	<b>100</b>

TABLE II  
QUESTIONARY RESULTS OF MIDDLE SCHOOL STUDENTS ON NEED FOR IMPROVEMENT OF THE CARER GUIDANCE SYSTEM

Answer option	Quantity	Ration, %
very necessary	272	64.8
necessary	145	34.5
no necessity	3	0.7
<b>Total</b>	<b>420</b>	<b>100</b>

users with the most useful and personalized variants for goods and services. The functioning of RS is based on filtering information with respect to a set of known properties of objects and users. For example, RSs are used to assess the users' preferences for goods and services (songs, films, video clips, books, articles, etc.), which have been not previously given ratings by the user trying to make a choice.

RSs are also quite successfully used in many business spheres, such as entertainment: offering songs to listeners (e.g., the LastFM system – www.last.fm), offering films (the Netflix system – www.netflix.com), recommended videos (the YouTube system – www.youtube.com); in education and training (learning resources, books, articles, site addresses), in intelligent systems of teacher assistants (predicting students' learning capabilities). RSs are the field of active IT research since 2007. Our literature and technologies review is in [2].

Teaching experience of the first author of this paper in secondary school shows the relevance of the problem of choice of a profession. This research, being the results of a master thesis, is dedicated to the development of an RS that allows students (applicants) to receive recommendations when taking career decisions. The aim of the research is development of a RS for supporting students in their carer decision making.

The resulting RS prototype has two options for producing recommendations. The first one is realized on the base of John Holland theory [?], according to which most people have one of six personality types. After determining the type, a set of corresponding universities are produced. The second one is an aggregation of questioning data about concrete universities obtained from students already being taught there. The recommendations, then, are generated by means of collaborative filtration (CF) based on users' profile comparison. Thus, the RS prototype combines two approaches: using expert knowledge and CF.

## II. TECHNIQUE OF CONSTRUCTION FOR THE RS HELPING WITH THE CARER CHOICE

*Recommender systems* [4] are decision support information systems designed to assess the user's level of interest in a particular product or service (object) based on available information about user and object. The RS development industry began to actively develop with the emergence of online sales services, and now it is one of the active areas of development of decision support systems, a direction of artificial intelligence, focused primarily on commercial use, as well as on solving problems of increasing the productivity of searching for relevant information. A profession is the object of RS recommendation production. Let us consider the development technique for RS construction from the point of view of solving the standard set problems and challenges.

Development of the RSs is aimed at solving the following set of problems [3]:

- 1) Increasing the sales of a product
  - a number of commodities sold,
  - organizing wider range product sales;
- 2) Increasing user satisfaction and/or loyalty;
- 3) Better understanding user needs;
- 4) Better products offers with respect to the user needs;
- 5) Selection of sets of products for users with a common properties
  - "good" ones, and
  - product groups, having a common usability properties;
- 6) Mining the classes of products, structuring the RS product domain, and
- 7) Generate a continuity of recommendations using the classification;
- 8) Rectifying user profile, *e.g.* with targeted questioning;
- 9) For the users having no goal to make choice, but searching for an expert opinions,
  - Analysis of other users impact on a choice,
  - Formalizing opinions,
  - Recommending opinions.

In this context, the RS is aimed at solving 4–7-th problems. This set defines the methods of RS proposal generation and algorithms, which have to be implemented. As we can see, the universities and their departments can be distributed on various carer directions by two ways: with the use of expert knowledge, when experts analyze a departments' properties

and relate them to a direction, and with use of unions of the students (users who already "consume" a product). In the second case, a good opinion of a user on a department (product) creates or solidifies a corresponding relation. In the same time, a group of students positively characterized a department form a group of "similar" users.

The main problem solved during the R&D is "*cold-start*" *problem*. For RS based on CF, the similarity of user profiles in terms of sets of specified characteristics are performed. Content filtering RS compare products. In both cases, RS cannot generate recommendations if it does not have enough information about users or products. The cold-start problem arises at the first stages of RS functioning and when a new user's behavior had not been observed sufficient time, *i.e.*, the RS have no data about his preferences, or user's profile contains no useful information for a comparison. The same situations arises for a new product, *e.g.* there is no user bought the product, and no responses were given. In this system explicit information collection has been realized for solving cold-start problem in CF, and John Holland's theory was used as initial profile data filling in, including the input data for the Holland's questionnaire.

The simplest technique for user profile data collection and usage is so called *impersonalized acquisition*. Recommendations are generated for an "average" user related to a nearest group of users. An RS gives top-*N* products for the group, efficiently ignoring personal preferences. The approach can be developed to support personal data with construction of subgroups of the groups relating them to a set of personal characteristics. This may lead to unwillingness giving personal data of users and impossibility to generate proposals when a user cardinally changes his/her preferences. In this system, as our intention is a detection of a direction of a carer, impersonalized approach seems to give sufficient quality result.

The *scalability problem*. Scalability is a property of software systems to be able to cope with the load prescribed by design, as well as with increased load if additional information–communication and computing resources are available. For a simple RS with several users and products, there is no need to maintain scalability, and algorithms such as "nearest-neighbor" are widely used for assessing interests as the combinatorial space is not large.

Thus, to construct carer choice RS prototype, we have chosen to collect impersonalized data, structure users with respect to their university/department preference and user profiles by means of collaborative filtering, use questionnaire and expert approach for initial user profile filling in.

### A. Holland's six types TODO

Realistic, Investigative, Artistic, Social, Enterprising, Conventional.

In order to determine the type one must answer 54 questions, the questions are evenly distributed on this six types. For each question, a 0-4 scores are to be set depending on the degree of manifestation. The score 0 corresponds to "no manifestation" was mentioned of a kind, and score 4

denotes “always manifestative”. For each type the scores of the corresponding nine questions are collected, and the type is figured out as a type with the maximum score sum. Student profiles store all answers and estimates, just for a further system development.

### III. PRELIMINARY DATA PROCESSING

In realty RS, two program agents are implemented, which are executed on an event occurrence. The first one starts by timer of web browser and implements issuing user positive evaluation, if user spends some time viewing a page of a realty object. The second one, located at the server side, is activated if the first one raised event of a positive evaluation. Server agent receives the evaluation and may start recalculation of estimates for recommendations if there are enough computational resources.

All data are stored in a database, which structure is represented as objects in Fig. ?? . The central entity is IOffer, representing the offer of a realty object. The information on an object, which is independent of an offer, is stored as IObject-object. The object structure corresponds to standard Yandex and Google form of real estate representation.

### IV. DATABASE STRUCTURE

#### V. IMPLEMENTATION

The RS is a web-application implement in Django MVP<sup>1</sup> framework, MySQL has been chosen for data storage. Python programming language is the main glue subsystem, gathering web-application and recommender unit in one system. For carrying out experiments with data, a Jupiter subsystem was included in the system as well.

User roles are administrators, anonymous and regular user. Administrator can manipulate user accounts and any raw data. Anonymous users can register, view information on universities, taught courses and professions. Regular registered user can obtain recommendations, rate universities, add an university to his/her preference list. The use-case diagram is shown in Figure 1.

```
from django.db import models
from django.contrib.auth.models import AbstractUser

# User
class NGUOIDUNG(models.Model):
    IdUser = models.AutoField(primary_key=True)
    TenNguoiDung = models.CharField(max_length=20, null=False)
    HoTen = models.CharField(max_length=200, null=False)
    MatKhau = models.CharField(max_length=200, null=False)
    Email = models.CharField(max_length=200, null=False)
    Quyen = models.IntegerField(null=False)
    GioiTinh = models.BooleanField(default=False)
    SDT = models.CharField(max_length=15, null=True)
    NgaySinh = models.DateTimeField(null=False,
        default="1999-01-01")

# Interests
class TINHCACH(models.Model):
    IdTinhCach = models.AutoField(primary_key=True)
    Loai = models.IntegerField(null = False)
    ChiTiet = models.CharField(max_length=1000, null=False)
```

<sup>1</sup>Abbreviation of Model-View-Controller programming pattern

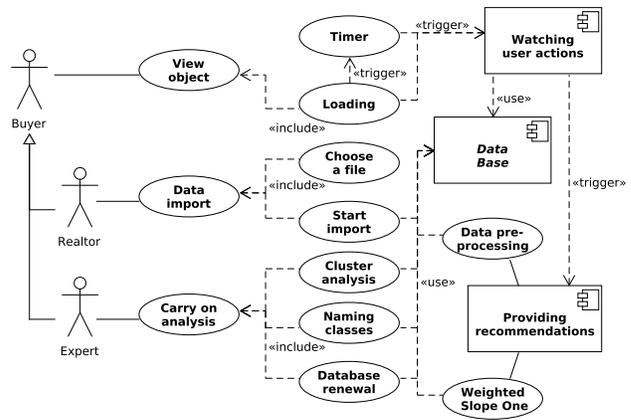


Fig. 1. Use case diagram for real estate RS [TODO: UPDATE to CDT’s thesis]

```
# List university
class TRUONG(models.Model):
    IdTruong = models.AutoField(primary_key=True)
    TenTruong = models.CharField(max_length=1000, null=False, default='')
    ChiTiet = models.CharField(max_length=1000, null=True)
    ChiTietFull = models.CharField(max_length=1000, null=True)

# Vote for an university
class VOTE(models.Model):
    IdVote = models.AutoField(primary_key=True)
    IdTruong = models.IntegerField(null=False)
    IdUser = models.IntegerField(null=False)
    Point = models.IntegerField(default=0)

# List of favorite universities
class FAVORITE(models.Model):
    IdFavo = models.AutoField(primary_key=True)
    IdTruong = models.IntegerField(null=False)
    IdUser = models.IntegerField(null=False)

from django.db import models
# Create your models here.
class ListUni(models.Model):
    name = models.CharField(max_length=100)
    description = models.TextField()
    excerpt = models.TextField(max_length=300)
    pictrure = models.ImageField(upload_to='Uni_picture')
    group = models.IntegerField(default=0)
    def __str__(self):
        return self.name
class list1(models.Model):
    name = models.CharField(max_length=100)
    address = models.CharField(max_length=200)
    description = models.TextField()
    excerpt = models.TextField(max_length=300)
    pictrure = models.ImageField(upload_to='Uni_picture')
    group = models.IntegerField(default=0)
    def __str__(self):
        return self.name
```

TODO Activity diagram

### VI. RECOMMENDATIONS GENERATION

## VII. WEB APPLICATIONS

## VIII. DISCUSSION

Comparing these two approaches to the RS R&D ....

The first approach ... allowed collecting data in the runtime of the RS and also used J. Holland's theory, which has been used for decades for carer guidance, ...

The second approach ... is more useful for solving problems 1, 6,7,9 of the list of Section II.

## IX. CONCLUSION

In this brief narration of the results of two master theses, information recommender systems (RS) were developed to support the choice of a profession for students in Viet Nam and Irkutsk, Russia. The following tasks have been solved:

1. The subject area of RS, assistance in choosing a profession (carer guidance), have been analyzed; typical problems were recognized and their solution method were proposed.

2. The theory of John Holland in carer guidance was applied for high school students in Ki Lam High School, Ha Tinh, Viet Nam.

3. High school students of last courses tested the universities rating setting playing role of university students.

4. A methodology for calculating recommendations for carer guidance, based on the method of collaborative filtering and expert evaluation, is proposed. The technique has been implemented with Python modules.

5. Implemented an RS subsystems in the Python programming environment based on the proposed techniques and data structures; This Rs was tested by younger students of the Viet Nam school.

The resulting RS MVP<sup>2</sup>, extended with automatic university data crawling and more efficient algorithms, is to be deployed in the school.

The source code is located at Github.com URL: [https://github.com/tranducthe/diplom\\_CollaborativeFiltering](https://github.com/tranducthe/diplom_CollaborativeFiltering).

## ACKNOWLEDGEMENTS

The results were obtained within the state assignment of the Ministry of Education and Science of Russia, the project "Methods and technologies of a cloud-based service-oriented digital platform for collecting, storing and processing large volumes of multi-format interdisciplinary data and knowledge based on the use of artificial intelligence, a model-driven approach and machine learning", No. FWEW-2021-0005 (State registration No. 121030500071-2).

## REFERENCES

- [1] Tomsk hypothesis.
- [2] AIIT'2020 Cherkashin, et al. Two recommender ...
- [3] Ricci.... types of RSs.
- [4] D. Jannach, M. Zanker, A. Felfernig, G. Friedrich. Recommender Systems: An Introduction. Cambridge University Press, 2010.
- [5] E. Charkashin, B. Shevchenko. Source code of a prototype real estate market recommender system. URL:<https://github.com/eugeneai/RSPO-CSharp>. (access date: 21-09-2020)

- [6] J. Beel, B. Gripp, S. Langer, C. Breiteringer. Research-paper recommender systems: a literature survey. *International Journal on Digital Libraries* (2016) 17: 305. DOI:10.1007/s00799-015-0156-0
- [7] S.P. Balandina, M.V. Bautin, A.V. Maiyorov, E.R. Smirnova. Social networks as means of education and educational participants interactions. *Pedagogicheskie i informacionnye tehnologii v obraxovanii*. No. 15 (2016). (in Russian) URL:<https://journals.susu.ru/pit-edu/issue/view/31>
- [8] K.S. Veber, A.A. Pimenova. Comparative analysis of social networks. *Vestnik Tambovskogo universiteta*. Series: Natural and technical sciences. Vol. 19, No. 2, 2014. p. 636–643. (in Russian)
- [9] A.V. Rozhkova. Self-expression (self-presentation) in social Internet networks as a phenomenon of human cyber-socialization. *Homo Cyberus*. 2017. No. 2(3). (in Russian) URL:[http://journal.homocyberus.ru/Samovyrazhenie\\_v\\_socialnyh\\_internet-setjah\\_kak\\_fenomen\\_kibersocializacii](http://journal.homocyberus.ru/Samovyrazhenie_v_socialnyh_internet-setjah_kak_fenomen_kibersocializacii)
- [10] A.V. Fescshenko. Social networks in education: experience analysis and development perspectives. *Gumanitarnaya informatika*, Tomsk, Issue 6. 2012. p. 124–134. (in Russian)
- [11] V.V. Matsuta, P.B. Kiselyev, A.V. Fescshenko, V.L. Goyko. Investigation of social networks potential for revealing talented students. *Psychologiya i psihotekhnika*. No. 4. 2017. p. 104–121. (in Russian)
- [12] Ch.C. Aggarwal *Recommender Systems: The Textbook*. Springer. 2016. ISBN 9783319296579.
- [13] P. Brusilovsky. *The Adaptive Web*. 2007. p. 325. ISBN 978-3-540-72078-2.
- [14] D.H. Wang, Y.C. Liang, D.Xu, X.Y. Feng, R.C. Guan. "A content-based recommender system for computer science publications", *Knowledge-Based Systems*. 2018. 157: 1-9
- [15] S. Blanda. *Online Recommender Systems – How Does a Website Know What I Want?*. American Mathematical Society. 2015.
- [16] J. Albahari, B. Albahari. *C# 6.0 in a Nutshell: The Definitive Reference 6th Edition*, 2015. 1138 p. ISBN-13: 978-1491927069
- [17] Entity Framework Tutorial. URL:<https://www.entityframeworktutorial.net/> (access date: 05.07.2020).
- [18] VKontakte API documentation. URL: <https://vk.com/dev/manuals> (access date: 05.07.2020).
- [19] D. Lemire, A. Maclachlan. Slope One Predictors for Online Rating-Based Collaborative. URL: [http://cogprints.org/4031/1/lemiremaclachlan\\_sdm05.pdf](http://cogprints.org/4031/1/lemiremaclachlan_sdm05.pdf) (access date: 10.05.2018)
- [20] A. Freeman. *Pro ASP.NET MVC 5*. 5th ed. Apress. 2013. 832 p. DOI:10.1007/978-1-4302-6530-6
- [21] J. Chadwick, T. Snyder, H. Panda. *Programming ASP.NET MVC 4: Developing Real-World Web Applications with ASP.NET MVC*. O'Reilly Media; 1st Edition. 2012. 492 p. ISBN-13: 978-1449320317
- [22] E. M. Alrawhani, H. Basirona, Z. Sa'ayaa. Real estate recommender system using case-based reasoning approach. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*. Vol. 8 No. 2. p. 177-182.
- [23] E. V. Britina. Segmenting recommender system with client-server connection based on programmable configured networks and usage protocol with fast-jumping IP-address. // *Sovremennye problemy nauki i obrazovaniya*. No 6. 2015. URL:<https://www.science-education.ru/ru/article/view?id=16875> (in Russian)
- [24] X. Yuan, J.-H. Lee, S.-J. Kim, Y.-H. Kim. Toward a user-oriented recommendation system for real estate websites. *Information Systems*, 38 (2013). p. 231–243.
- [25] T. Ginevičius, A. Kaklauskas, P. Kazokaitis, J. Alchimovienė. Recommender system for real estate management. *Verslas: Teorija ir praktika (Business: Theory and Practice)*. 2011 12(3). p. 258–267 DOI:10.3846/btp.2011.26
- [26] E. A. Belotskiy, A. V. Suetin, Construction of a recommendersystem for choosing higher education institutions for entrants, *Vestnik S.-Petersburg Univ. Ser. 10. Prikl. Mat. Inform. Prots.Upr.*, 2016, Issue 1, 66–77. URL:<http://www.mathnet.ru/links/1e01f91d57c8dc6f050c7e6d3fed3604/vspui277.pdf> (in Russian)
- [27] B. Bakhshinategh, G. Spanakis, O. Zaiane, S. ElAtia. A Course Recommender System based on Graduating Attributes. In *Proceedings of the 9th International Conference on Computer Supported Education*, Volume 1: CSEDU, ISBN 978-989-758-239-4, 2017, pp. 347-354. DOI:10.5220/0006318803470354
- [28] A. Al-Badarenah, J. Alsakran. "An Automated Recommender System for Course Selection" *International Journal of Advanced Computer Science and Applications(Ijacs)*, 7(3), 2016. DOI:10.14569/IJACSA.2016.070323

<sup>2</sup>Abbreviation of Minimal Valuable Product

- [29] J. Naren, M. Z. Banu, S. Lohavani. Recommendation System for Students'Course Selection. A. K. Somani et al. (eds.), Smart Systems and IoT: Innovationsin Computing, Smart Innovation, Systems and Technologies 141. 2020. pp. 825–833. DOI:10.1007/978-981-13-8406-6\_77
- [30] B. R. Avhadeev, L. I. Voronova, E. P. Ohapkina. Development of recommender system on the base of profile in social network "VKontakte". Vestnik Nizhnevartovskogo gosudarstvennogo universiteta. Issue No 3. 2014. (in Russian)
- [31] A. S. Amel'kin, D. M. Ponizovkin. Mathematical model of top-N problem for content recommender systems. Izvestiya MGTU "MAMI" No 3(17), 2013, Vol.2. p. 26-31. (in Russian)

# Machine Learning Techniques for Smart Digital Technologies

H. M. Said\*

\* Department of Business Management/ Information system, College of Science and Humanities at Alghat, Majmaah University, Al-Majmaah 11952, Saudi Arabia

\* Faculty of Computing Science & IT, Cairo, Egypt, Ahram Canadian University, Cairo, Egypt  
Email: hanaa.e@mu.edu.sa, hanaamoh@hotmail.com

**Abstract - Machine learning (ML) is a new revolution of the Internet of things (IoT) applications. ML in the industry is a rapidly developing area. Numerous ML research and application projects have been done by universities or in joint industry-university consortia in recent years. However, an important question to be further addressed is on value creation by ML industry applications. Machine learning applications in the sense of this paper are solutions using ML techniques to improve industrial manufacturing processes, enable new and efficient ways to operate production plants, create new service or supervision means for industrial installations, offer an optimized infrastructure, reduce operational cost or improve human safety in industrial areas. On the other side, smart digital systems require the integration of Machine Learning, Information mining algorithms, and artificial intelligence paradigms into specific application domains. The main issues are to design, develop and implement effective smart digital learning Systems that improve currently existing applications. The businesses are then able to leverage the extracted knowledge into more clients, more sales, and greater profits. The present paper brings together experts from academia, research, and industry offering a view on the ML application in the industrial environment, the challenges, the expected evolution of Machine learning technology, and its use in future factories, on connected and holistic processes. The paper is intended to contribute to a Machine learning supported paradigm change in manufacturing, industrial service, and over life sustainable industrial activities.**

## I. INTRODUCTION

Machine Learning Techniques, data mining, and Intelligent Data Analysis aim to extract useful information and discover some hidden patterns from the huge amount of databases, Smart digital Technologies which statistical approaches cannot discover [1][2]. It is a multidisciplinary field of research that includes machine learning techniques, statistics, expert systems, databases, high-performance computing, visualization, Rough sets, Fuzzy Logic, Neural Networks, Genetic algorithm, Case-based reasoning, SVM, Swarm intelligence and knowledge representation, etc [3][4].

Acting on the discovered knowledge by using the knowledge directly, incorporating the knowledge into another system for further actions, or simply documenting it and reporting it to interested parties. This process includes checking for and resolving potential conflicts [5][6][7]. Machine learning such as data mining of security data (e.g., related to police operations) could enable the discovery of systemic inefficiency in connection to security response, crime incidence analysis, and prevention efforts.

Data mining could help in providing an explanation of crime and terror [8][9]. The data mining techniques however could mine the historical data and extract hidden predictive information which could prove the initial assumption that it will face numerous problems. However, data mining can be a valuable tool in the hands of the decision-makers aiming to predict the internal and external environment, adjust accordingly, and hence make more rational decisions [10]. Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. Data mining consists of more than collecting and managing data; it also includes analysis and prediction [11], [12][13].

This paper may be a useful tool that enables manufacturing, industrial service, and over life sustainable industrial activities. to find the major points for managing effective Smart Digital Technologies.

This paper focused on a variety of techniques, approaches, and different areas of the research which are helpful, and Smart digital Technologies as the important field of data mining technologies.

This paper consists of 3 sections; the first section is the introduction about machine learning and e-technology, the second section includes a comparison between different machine learning approaches for manufacturing, industrial service, and over life sustainable industrial activities. Section 3 includes the conclusion and the future work.

## II. MACHINE LEARNING TECHNIQUES

Machine learning Techniques is one of the most interesting research areas in Computing and more specifically with great impact in Computer Science, Information Technology, and Information Systems. E-Technology systems require the integration of Machine Learning, Information Retrieval, Data mining algorithms into specific application domains. The main issues are to design, develop and implement effective Machine learning Systems that improve currently existing applications. Machine Learning (ML) is the branch of artificial intelligence (AI), concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data. As intelligence requires knowledge, computers must acquire knowledge shown in Figure 1.

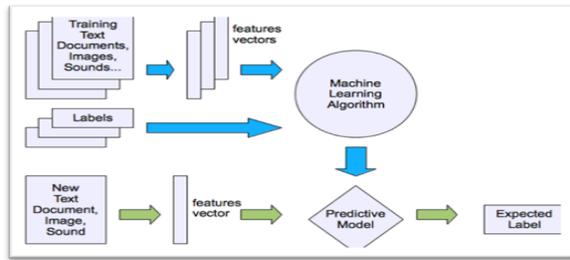


Figure 1: Machine learning structure

“Artificial Intelligence is a study of complex information processing problems that often have their roots in some aspect of biological information processing. The goal of the subject is to identify solvable and interesting information processing problems, and solve them.” -- David Marr.

“Artificial Intelligence is the design, study, and construction of computer programs that behave intelligently.” -- Tom Dean.

“Artificial Intelligence is the enterprise of constructing a physical symbol system that can reliably pass the Turing test.” -- Matt Ginsberg.

This paper overviews some techniques and algorithms in machine learning. Furthermore, there are more and more techniques that apply machine learning as a solution. In the future, machine learning will play an important role in our daily life.

### III. SMART DIGITAL TECHNOLOGIES

Smart Digital Technologies (SDT) has emerged as a powerful paradigm in e-Science, providing researchers with an immense volume of intelligent computing techniques and algorithms. SDT provides knowledge engineers to develop robust techniques and intelligent tools for e-government applications and tasks [10].

Smart Digital Technologies (SDT) is a sub-branch of AI. This computing branch is concerned with the study of adaptive mechanisms to enable or facilitate intelligent behaviour in complex and changing environments. These mechanisms include those AI paradigms that exhibit an ability to learn or to adapt to new situations, to generalize, abstract, discover, and associate. This computing area includes the following paradigms: fuzzy systems (FS), artificial neural networks (ANN), rough sets (RS), support vector machines (SVM), evolutionary computing (EC), swarm intelligence (SI), and genetic algorithms (GA). Each of the Smart Digital Technologies (SDT) paradigms has its origins in biological systems. ANNs model biological neural systems, EC models natural evolution (including genetic and behavioural evolution), SI models the social behaviour of organisms living in swarms or colonies, and FS originated from studies of how organisms interact with their environment and human thinking processes. On the other hand, soft computing, a term coined by, Zadeh, is a different grouping of paradigms, which usually refers to the collective set of CI paradigms and probabilistic methods. This study discusses exploiting

some of the computational intelligence paradigms in the applications of e-technologies and activities, e.g. e-learning, e-economy, e-commerce, e-court, e-business, e-health [17].

TABLE 1. ARTIFICIAL NEURAL NETWORKS AS MACHINE LEARNING TECHNIQUES IN E- TECHNOLOGIES

Application	Knowledge Types and Tasks
E-Business [14]	Developing Knowledge Bases Classification, Clustering Early Warning and Proactive, Control Systems.
E-Business [15,16]	Classification Early Warning and Proactive Control Systems
E-Health [17]	is used to design neural network With joint weight. In this algorithm testing of the neural network, the controller is done by the classic boost power transformer
E-Economy [18]	For clustering of large database Using the exemplar. Power system analysis computing, Power system measurements.
E-Commerce [19]	data mining classification and clustering tasks
E-Health [20]	Medical data mining
E-Economy [21]	Analysis of web mining applications
E-Commerce [22]	Mining the data over the www using various data through. E-government systems

TABLE 2. FUZZY LOGIC PARADIGM IN E- TECHNOLOGIES

Application	Knowledge Types and Tasks
E-Business [23]	Assignment of genes to templates is based on fuzzy membership function. A multi-objective evolutionary algorithm is used to determine compact clusters with a varying number of templates.
E-Business [24]	Historical flood data are clustered into several partitions by applying a fuzzy clustering iteration algorithm and the optimal number of clusters is obtained by using cluster validity criterion, and a real-time flood Hydrograph partition is recognized by the fuzzy pattern recognition model.
E-Market [25]	Knowledge Discovery
E-Health [26]	Analysis of transformation of the probed data to a linguistic level using fuzzy set theory
E-Business [27]	Focusing on rule-based fuzzy controllers. It is argued that the dominant position of analytic control theory prevented fuzzy control from being taken seriously until its increasing application in Japan
E-Market [28]	Knowledge Discovery using fuzzy logic
E-Learning [29]	Fuzzy Intrusion Detection
E-Commerce [30]	Analysis of web mining applications.

TABLE 3. GENETIC ALGORITHM (GA) TECHNIQUES IN E- TECHNOLOGIES

<i>applications</i>	<i>Knowledge Types and tasks</i>
E-Learning [31]	The application of GA techniques to the problem of overloaded arrays, in which the number of transmitted narrowband signals is greater than the number of receiver array elements, is explored.
E-Business [32]	The difference is defined using the contours of the fuzzy system of each image. The search for the optimal translation is considered a minimization process.
E-Business [33]	A stabilizing compensator was designed using genetic optimization combined with the Hinf design. The proposed controller can be implemented digitally with synchronous sampling.
E-Business [34]	A method is proposed to automatically extract numerical control rules from the sensor data without the help of experts through a GA.
E-Business [35]	It adopts GA to give information pheromone to distribute. And, it makes use of the ant colony algorithm to get several solutions through information pheromone accumulation and renewal.
E-Business [36]	GA is added to ant colony algorithms every generation in the proposed algorithm. Making use of the GAs advantage of whole quick convergence, the ant colony algorithm's convergence speed is quickened.
E-Business [37]	this paper analyzes the characteristics and shortcomings of simple GA, simulated annealing genetic algorithm as well as immune algorithm respectively

TABLE 4. CASE-BASED REASONING PARADIGM IN E- TECHNOLOGIES

<i>Application</i>	<i>Knowledge Types and tasks</i>
E-Business [38]	Ontology-based distributed case-based reasoning to effectively support knowledge retrieval within the virtual enterprise environment.
E-Business [39]	proposed a concept of software realization of a hybrid expert system
E-Business [40]	Induction learning technique.
E-Business [44]	The retrieval mechanism of case-based reasoning in fixture design is analyzed.
E-Commerce [41]	To identify the similarity between the cases and to develop a single and consistent rule set with their help.
E-Commerce [42]	Operates on a software process that includes various sub-processes and utilizes experience by CBR technique to increase its effectiveness.
E-Business [43]	Creation of customized management model through CBR technique to satisfy the specific management requirements.
E-Business [44]	Solving complex case adaptation problems.
E- Tourism [55]	Uses CBR techniques in the management and organizing tasks of the tourism processes and activities.

TABLE 5. SUPPORT VECTOR MACHINES PARADIGM IN E- TECHNOLOGIES

<i>Applications</i>	<i>Knowledge Types and tasks</i>
E-Business [45]	Support Vector Machines are trained on large-scale datasets and SVM classifying accuracy
E-Government information system [46]	A classification algorithm for E-government documents based on support vector machines is proposed.
E-learning [47]	The SVM classifier is approved to detect unknown samples of malware with the probability of 74 - 83 percent
E-Business [48]	The individual bio-components then can be classified as bio- or non-bio-aerosol by our SVM classifier.
E-Commerce [49]	To develop innovative services to cope with customers' evolving demands and to create customers' value
E-Business [50]	For the general workflow of E-government document circulation in the current information system, a scheme of the auto-classification and archiving based on service is presented so that it is realized in a heterogeneous information system
E-Business [51]	The experiment results show that the new parallel SVM training algorithm is efficient and the SVM classifying
E - Bank [52]	This study applies Support Vector Machines (SVM) and Artificial Neural Network (ANN) as methods for determining the visual condition of roads on an inventory and traffic data set
E-Commerce [53]	An experiment based on the SVM model shows that the proposed method can detect malware with strong resilience and high accuracy

TABLE 6. SWARM INTELLIGENCE PARADIGM IN E- TECHNOLOGIES

<i>Applications</i>	<i>Knowledge Types and Tasks</i>
E-Business [54]	focus on hierarchical task network
E-Mail policy [56]	Their initiatives for implementing business-grade electronic mail (e-mail) services for federal, state, local, and tribal agencies.
E- Learning [57]	To provide a quick sorting speed and reliable evaluation with a long lifespan, a radio frequency identification (RFID) system is often employed in practice.
E-Business [58]	Regression Test suite optimization is an effective technique to reduce the time and cost of testing.
E-Business [59]	This research with traditional AI approaches and focus on hierarchical task network (HTN) descriptions of constraints
E-Commerce [60]	In this paper, we look forward to identifying suspicious behaviors in e-government procurement systems, through the use of business intelligence techniques

From the above tables, it can be seen the following observations;

The artificial neural networks paradigm has been used for a wide range of applications of Machine learning Techniques (MLT), including classification, clustering, medical diagnosis, control and power systems, and others.

Support vector machines and swarm intelligence paradigms have been applied to many e- economic activities, producing state-of-the-art results.

Fuzzy-logic and genetic algorithms paradigms have been applied successfully to e-business, e-market, and e-commerce applications.

CBR paradigm has been used successfully in e-business and e-commerce for the following tasks, knowledge retrieval, and software realization of hybrid expert systems, induction learning technique, management requirements, and solving complex case adaptation problems.

Ontologies were developed in E-Systems, Machine learning Techniques (MLT) to facilitate knowledge sharing, refine, search and reuse [55].

Machine learning Techniques (MLT) assist in identifying patterns of criminal or terrorist behavior, identifying emerging criminal or terror threats, and predicting future criminal or terrorist actions [61].

### III. CONCLUSION

Machine learning continues to affirm its important position in the context of Information and Communication Technologies and the development of society. Data mining is considered as the main step in the knowledge discovery process that is concerned with the algorithms used to extract potentially valuable patterns, associations, trends, The value of data mining applications is often estimated to be very high. Many businesses have stored large amounts of data over years of operation, and data mining can extract very valuable knowledge from this data. The businesses are then able to leverage the extracted knowledge into more clients, more sales, and greater profits. Whereas concepts and basic foundations have been elaborated and reached maturity, further efforts are necessary for unleashing the full potential, federating systems and actors. Machine learning is a concept and a paradigm that considers pervasive presence in the environment of a variety of industries that include Converging Technologies for Smart Environments and Integrated Ecosystems, In line with this development. Although larger players in some application areas still do not recognize the potential, many of them pay high attention or even accelerate the pace by coining new terms for Machine learning and adding additional components to it. Moreover, end-users in the private and business domain have nowadays acquired a significant competence in dealing with smart devices and networked applications. As Machine learning continues to develop, the further potential is estimated by a combination with related technology approaches and concepts such as Cloud computing, Future Internet, Big Data, robotics, and semantics.

### REFERENCES

- [1] I.Kononenko (1993) Inductive and Bayesian learning in medical diagnosis, *Applied Artificial Intelligence*, 7:317-337.
- [2] I.Kononenko, I.Bratko, M.Kukar (1998) Application of machine learning to medical diagnosis. In: R.S.Michalski, I.Bratko, and M.Kubat (eds.): *Machine Learning, Data Mining, and Knowledge Discovery: Methods and Applications*, John Wiley & Sons.
- [3] I.Kononenko (1995) On Biases in Estimating Multivalued Attributes, *Proc. Int. Joint Conf. on Artificial Intelligence IJCAI-95*, Montreal, August 20-25 1995, pp. 1034-1040.
- [4] A.G.Lee (1996) *The Selected Works in the Region of Parapsychology*, Moscow, Russia: L.L.Vasilyev Fund of Parapsychology.
- [5] K.Korotkov (1998) *Aura and Consciousness: A New Stage of Scientific Understanding*, St.Petersburg, Russia: State Editing & Publishing Unit "Kultura".
- [6] Piatetsky-Shapiro and Gregory, "The Data-Mining Industry Coming of Age", in *IEEE Intelligent Systems*, vol. 14, issue 6, Nov 1999.
- [7] Hsu J., "Data Mining Trends and Developments: The Key Data Mining Technologies and Applications for the 21st Century", in the *Proceedings of the 19th Annual Conference for Information Systems Educators*.
- [8] Korosh Golnabi, Richard K. Min, Latifur Khan, and Ehab Al-Shaer, "Analysis of Firewall Policy Rules Using Data Mining Techniques", *IEEE*, 2006.
- [9] Abdul Nasir Khan, M.L. Mat Kiah, Samee U. Khan, and Sajjad A. Madani, "Towards secure mobile cloud computing: A survey", *Elsevier B.V*, 2012.
- [10] Md. Tanzim Khorshed, A.B.M. Shawkat Al, and Saleh A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing", *Elsevier B.V*, 2012. Niroshinie Fernando, Seng W. Loke, and Wenny Rahayu, "Mobile cloud computing: A survey", *Elsevier B.V*, 2013.
- [11] S. Axelsson, "Intrusion Detection Systems: A Survey and Taxonomy". Technical Report 99-15, Chalmers Univ., March <http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.1.6603>, 2000.
- [12] S.J. Stolfo, W. Lee. P. Chan, W. Fan, and E. Eskin, "Data Mining – based Intrusion Detector: An overview of the Columbia IDS Project" *ACM SIGMOD Records* vol. 30, Issue 4, 2001.
- [13] Jiawei Han and Micheline Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2nd edition 2006, 3rd edition 2011.
- [14] Bhavani, M., & Thuraisingham, Q, *Data Mining and Cyber Security* Conference: International Conference on Quality Software - QSIC, 2003.
- [15] Hong Yu, Xiaolei Huang, Xiaorong Hu, Hengwen Cai (2010) "A Comparative Study on Data Mining Algorithms for Individual Credit Risk Evaluation", International Conference on Management of e-Commerce and e-Government, 2010.
- [16] Pengyu Li. Research on the Quantitative Methods of Information Security Risk Assessment Based on AHP. The degree of Master of Management Science And Engineering in the Graduated School of Jiangxi university of finance & economics, 2012:2-4  
Bouzerdoum, Havstaad, Beghdadi, "Image Quality assessment Using a Neural Network Approach", *IEEE*, 2004, pp-330-333.
- [17] ymond Lau, Stephen Liao, K. F. Wong, and Dickson Chiu, "Web 2.0 Environmental Scanning and Adaptive Decision Support for Business Mergers and Acquisitions," analyzed company mergers and acquisitions (M&A).
- [18] Jianshe Dong, Zhanting Yuan, Qiuyu Zhang, and Yufeng Zheng, "A Novel Anti-spam Scheme for Image-based Email", First International Symposium on Data, Privacy and E-Commerce, *IEEE* 2007 <http://www.cs.jhu.edu/~mdredze/datasets/image-spam>
- [19] Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society, Series*

- B, 26, 211–246
- [20] Giorgio Fumera, Ignazio Pillai, Fabio Roli. Spam Filtering Based On The Analysis Of Text Information Embedded Into Images. *Journal of Machine Learning Research*, 7(Dec):2699--2720, 2006
- [21] Hailing Huang, Weiqiang Guo, and Yu Zhang, "A Novel Method for Image Spam Filtering", *IEEE, Ninth International Conference for Young Computer Scientists*, 2008.
- [22] G. Pahl, W. Beitz, *Engineering Design*, The Design Council, London, 1988.
- [23] S. Pugh, *Total Design*, Addison-Wesley, Wokingham, UK, 1991.
- [24] L.A. Zadeh, "Fuzzy set," *Information and Control*, vol. 8, pp. 338-353, 1965.
- [25] R.E. Bellman, L.A. Zadeh, "Decision-making in a fuzzy environment," *Management Science*, vol. 17, no. 4, pp. 141-164, 1970.
- [26] J. Davidson, W. Pedrycz, and I. Goulter, "A fuzzy decision model for the design of rural natural gas networks," *Fuzzy Sets and Systems*, vol. 53, pp. 241-252, 1993.
- [27] S.W. Hsiao, "A systematic method for color planning in product design," *Color Research and Application* vol. 20 pp. 191-205, 1995.
- [28] Y. Shimizu, T. Jindo, "A fuzzy logic analysis method for evaluating human sensitivities," *International Journal of Industrial Ergonomics*, vol. 15, pp.39-47, 1995.
- [29] C.T. Chen, "A fuzzy approach to select the location of the distribution center," *Fuzzy Sets and Systems*, vol. 118, pp. 65-73, 2001.
- [30] C. H. Cai, W. C. Fu, C. H. Cheng, and W. W. Kwong, "Mining association rules with weighted items," *The International Database Engineering and Applications Symposium*, pp. 68-77, 1998.
- [31] O. Cerdón, F. Herrera, and P. Villar, "Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the database," *IEEE Transactions on Fuzzy Systems*, Vol. 9, No. 4, 2001.
- [32] F. Herrera, M. Lozano, and J. L. Verdegay, "Fuzzy connectives based crossover operators to model genetic algorithms population diversity," *Fuzzy Sets and Systems*, Vol. 92, No. 1, pp. 21-30, 1997.
- [33] T. P. Hong, C. H. Chen, Y. L. Wu and Y. C. Lee, "Using Divide-and-Conquer GA Strategy in Fuzzy Data Mining," *The Ninth IEEE Symposium on Computers and Communications*, Vol.1, pp. 116-121, 2004.
- [34] T. P. Hong, C. H. Chen, Y. L. Wu and Y. C. Lee, "A GA-based fuzzy mining approach to achieve a trade-off between the number of rules and suitability of membership functions", *Soft Computing*, Vol. 10, No. 11, pp. 1091-1101, 2006.
- [35] M. Boddy, B. Horling, J. Phelps, R. P. Goldman, R. Vincent, A. C. Long, B. Kohout, and R. Maheswaran. *C-TAEMS Language Specification, Version 2.02*. DARPA, Arlington, VA, 2006.
- [36] R. H. Hou, T. P. Hong, S. S. Tseng, and S. Y. Kuo, "A new probabilistic induction method, " *Journal of Automatic Reasoning*, Vol. 18, pp. 5-24, 1997.
- [37] T. P. Hong, C. S. Kuo, and S. C. Chi, "Mining association rules from quantitative data", *Intelligent Data Analysis*, Vol. 3, No. 5, pp. 363-376, 1999.
- [38] T. P. Hong, C. S. Kuo, and S. C. Chi, "Trade-off between time complexity and the number of rules for fuzzy mining from quantitative data," *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, Vol. 9, No. 5, pp. 587-604, 2001.
- [39] T. P. Hong and C. Y. Lee, "Induction of fuzzy rules and membership functions from training examples," *Fuzzy Sets and Systems*, Vol. 84, pp. 33-47, 1996.
- [40] A.Aamodt, E. Plaza, *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches*. *AI Communications*, Vol.7(1994), P.39
- [41] C.H.Crawford, G.P.Bate, L. Cherbakov, K. Holley, C. Tsocanos, *Toward an On-Demand Service-Oriented Architecture*. *IBM Systems Journal*, Vol.44(2005), P.81
- [42] F. Marir, S.L. Mansar, *An Adapted Framework and Case-based Reasoning for Business Process Redesign*. (In *Proceedings of 2nd International Conference on Information Technology: Research and Education*, London, 2004)
- [43] T. Osman, D. Thakker, D. Al-Dabass, *Bridging the Gap between Workflow and Semantic-Based Web Services Composition*. (IEEE/WIC/ACM *International Conference on Web Intelligence*, France 2005)
- [44] S. Lajmi, C. Ghedira, K. Ghedira, D. Benslimane, *How to Compose Web Services via Case-Based Reason*. (Conference on e-Business Engineering, China 2006)
- [45] J.Cardoso and A. Sheth, *Adaptation and Workflow Management Systems*. (In *Proceedings of International Conference WWW/Internet*, Portugal 2005)
- [46] S.Craw, N. Wiratunga, R.C. Rowe, *Learning Adaptation Knowledge to Improve Case-Based Reasoning*. *Artificial Intelligence*, VOL.170(2006), P.1175
- [47] V.Vapnik. *The support vector method of function estimation. Nonlinear Modeling: Advanced Black-Box Techniques*. Boston: Kluwer Academic Publishers, 1998.
- [48] V. Vapnik. *Three remarks on support vector function estimation*. *Advances in Kernel Methods: Support Vector Learning*. The MIT Press, Cambridge, Massachusetts, 1999.
- [49] M. Stitson, A. Gammerman, V. Vapnik, V. Vovk, C. Watkins and J. Weston. *Support vector regression with ANOVA decomposition kernels*. *Advances in Kernel Methods: Support Vector Learning*. The MIT Press, Cambridge, Massachusetts, 1999.
- [50] J. Weston, A. Gammerman, M. Stitson, V. Vapnik, V. Vovk and C. Watkins. *Support vector density estimation*. *Advances in Kernel Methods: Support Vector Learning*. The MIT Press, Cambridge, Massachusetts, 1999.
- [51] V. Vapnik. *An overview of statistical learning theory*. *IEEE transactions on Neural Networks* 1999, 10( 5):988-1000.
- [52] V.Vapnik. *The support vector method of function estimation. Nonlinear Modeling: Advanced Black-Box Techniques*. Boston: Kluwer Academic Publishers, 1998.
- [53] Chen yisong, Wang guoping, Dong shihai. *A Progressive Transductive Inference Algorithm Based on Support Vector Machine* [J]. *Journal of Software*, 2003, 14(3): 451-46
- [54] H Drucker, D. Wu, and V. Vapnik. *Support vector machines for spam categorization*. *IEEE transactions on Neural Networks* 1999, 10( 5): 1048-1055.
- [55] O. hapelle, P. Haffner and V.Vapnik. *Support vector for histogram-based image classification*. *IEEE transactions on Neural Networks* 1999, 10( 5): 1055-1065.
- [56] *filtering of e-mail spam with content-specific misclassification costs*. *Proceedings of the TextDM'01 Workshop on Text Mining*, 2001.
- [57] M. Boddy, B. Horling, J. Phelps, R. P. Goldman, R. Vincent, A. C. Long, B. Kohout, and R. Maheswaran. *C-TAEMS Language Specification, Version 2.02*. DARPA, Arlington, VA, 2006.
- [58] S. Brueckner. *Return from the Ant: Synthetic Ecosystems for Manufacturing Control*. Dr.Rer.nat. Thesis at Humboldt University Berlin, Department of Computer Science, 2000. R
- [59] W. Chen and K. Decker. *The Analysis of Coordination in an Information System Application - Emergency Medical Services*. In P. Bresciani, P. Giorgini, B. Henderson-Sellers, and M. Winiko, Editors, *Agent-Oriented Information Systems*, vol. 3508, LNAI, pages 36-51. Springer, New York, NY, 2005.
- [60] K. Decker and J. Li. *Coordinating mutually exclusive resources using GPGP*. *Journal of Autonomous Agents and Multi-Agent Systems*, 3(2):133-158, 2000.

# Utilization Of Different Approaches For Data Security In Business Intelligence

V. Naneva\* and K. Stefanova\*\*

\* Plovdiv University "Paisii Hilendarski", Plovdiv, Bulgaria

\*\* Plovdiv University "Paisii Hilendarski", Plovdiv, Bulgaria  
vnaneva@uni-plovdiv.bg  
kstefanova@uni-plovdiv.bg

**Abstract:** Nowadays each human activity processes and produces different types of information. For the purposes of business improvement, we should use all available technologies and methodologies for this processing to be simultaneously dynamically and secure. Additionally, from the company perspective all of the enterprise related data is highly confidential so surely there should be protection rules applied. The focus of this paper is on the key approaches for data protection in each of the Business Intelligence levels, such as data transformation, storage, visualization and data shareability. We will emphasize on several appropriate tools for data protection.

## I. INTRODUCTION

Business Intelligence (BI) is a well-known sphere that is characterized by a formed approach for extracting, manipulating, storing and representing data in an appropriate way. The defined techniques for each of the BI stages help businesses to get a better understanding of their daily activities and to make predictions for future periods. In regard to enterprise data, one of the key points is concerned with the security of processing flow, due to the fact that all or part of the information may be highly confidential. This leads to the necessity of specifying rules for protection in each of the BI levels so that there will not be data leaking or unwanted access.

## II. SECURITY ON DATA PROCESSING LEVEL

Usually in the daily activities of a certain business there are various data sources that are dynamically updatable. Let us take as an example a Human Resources (HR) unit, which is ordinarily responsible for tracking of vacation and sick days, maternity period, etc. Such a unit may work with information from different sources, for instance, Sales Force platforms, HR tools and local databases. If we need to process this data and as a result to summarize all the billable days for a number of employees, we need to have a defined algorithm for combining those sources. Additionally, it is necessary also to convert the raw data into understandable information.

In this section, we will suggest an approach for data protection by considering the capabilities of Azure cloud provider from the data transformation level, dividing the security rules based on data preparation, authorization keys application and SQL server security rules manipulation.

## A. Data Preparation

The starting point in which we can adjust security rules is the data preparation level. The information firstly needs to be extracted from the multi-updatable source and then to be transformed according to the business necessities. Stream Analytics Job (SAJ) is a tool related to real time data processing. As it is mentioned in Farseev's paper, Windows Azure Stream Analytics cloud service allows using various query constructions such as "Like/Not Like patterns", "CASE statements", "Embedded procedures", etc. [1].

Azure Stream Analytics Job connects to one or more data inputs and accepts data from different kinds of event sources. It streams SQL query, written for each job, and there could be multiple tables joined. The event sources could be Event Hubs, IoT Hub, and Blob storage.

When we want to send the transformed data in an appropriate format, we need to configure the SAJ output. It can be used with a single output per job or with multiple outputs per streaming by adding multiple INTO clauses to the query. Outputs vary as types such as Azure Data Lake Storage Gen 1, Blob storage, Power BI, Azure Table storage and others.

We have several different data extractions based on time windowing, for instance, tumbling, hopping, sliding, session and snapshot. To clarify the necessities of window functions' protection, we will briefly describe the process for some of them.

If we consider the Tumbling window function, we need to take into account that it is used to segment a data stream into distinct time sections and to perform a function against them (see Figure 1).

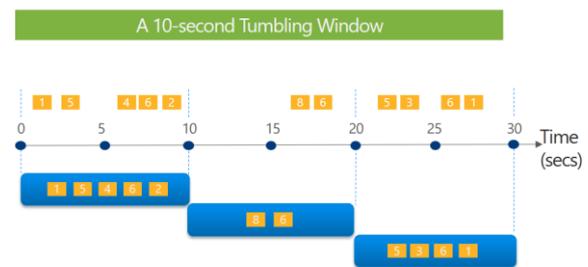


Figure 1. Tumbling window

With regard to this type of functions, we may point out that they repeat, do not overlap, and an event cannot belong to more than one tumbling window. In the following example (see Code 1) we are using **TumblingWindow()** to create a table view of the number of user queries of each employee for a specific end time. There is no restriction of the event amount for the opening period of time.

```
SELECT System.Timestamp() as WindowEndTime,
Employee, COUNT(*) AS Count
```

```
FROM HRStream TIMESTAMP BY CreatedAt
```

```
GROUP BY TimeZone, TumblingWindow(second,5)
```

Code 1. Example Function for Tumbling Window

As opposed, if the events arrive at the same time, we need to use the Session window function (see Figure 2). It is required for grouping events based on time period and if time sections with no data need to be filtered out. It has three main parameters: timeout, maximum duration, and partitioning key. The session window begins when the first event occurs. If another event is entering within the specified timeout from the last ingested event, then the window extends to include the new one.

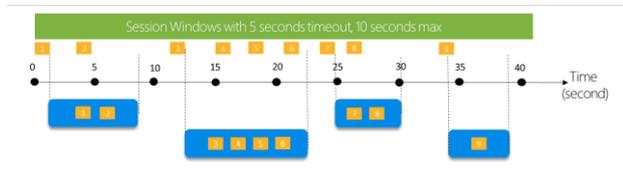


Figure 2. Session window

If no events occur within the timeout, then the window is closed and the function is waiting for another section trigger. By this kind of a function (see Code 2), we illustrate the number of users that are entering a certain software tool in a specific period of time.

```
SELECT System.Timestamp() as WindowEndTime,
Employee, COUNT(*) AS Count
```

```
FROM HRStream TIMESTAMP BY CreatedAt
```

```
GROUP BY Employee, SessionWindow(second,5,10)
```

Code 2. Example Function for Session Window

The last example that we will consider is about the Snapshot Window function (see Figure 3). This is one of the most common SAJ processes of data manipulation. Snapshot windows group events that have the same timestamp. Unlike other windowing types, which require a specific window function, we can use Snapshot to apply the process by adding **System.Timestamp()** to the **GROUP BY** clause.

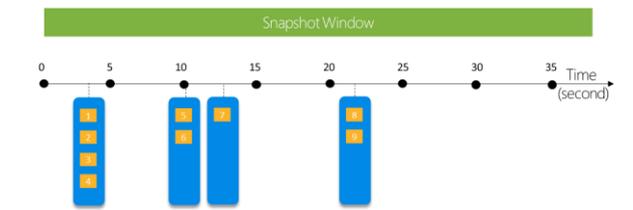


Figure 3. Snapshot Window

Commonly, snapshot windows are applied to aggregate events after they have been grouped on a single timestamp by implementing other window functions, like for example **TumblingWindow()**. This method can be considered in the **GROUP BY** clause as a key column or a snapshot window definition. The following example returns the count of records with the same employee that occur at exactly the same time:

```
SELECT Employee, COUNT(*)
```

```
FROM HRStream TIMESTAMP BY CreatedAt
```

```
GROUP BY Employee, System.Timestamp()
```

Code 3. Example Function for Snapshot Window

Aside from which window functions we are going to use, it is required security rules to be applied. If we follow the logic of Session windowing, we can implement a smaller value for maximum duration in order to provide fewer options for inappropriate data to be processed. However, other windowing methods do not support limitations in their working flow. That is why we need to follow security rules, based on the entire ASJ, such as:

- Endpoints in Stream Analytics clusters – by the help of a single tenant dedicated compute cluster, it is possible to manage private endpoints, which allows running secure outbound connection to the input and output resources.
- Identity authentication – by the configuration setting “Allow trusted services”, it could enable the secure connection with strong authentication.

#### A. Authorization keys

Another approach for data protection from the point of view of data processing is the implementation of Azure Key Vault. It is a service for securely storing and accessing secrets by which we can create access points such as API keys, passwords, certificates, or cryptographic keys. The working flow of this protection technique refers to two main users – developer and security administrator (see Figure 4). The security administrator is responsible for creating and configuring secrets or keys, and monitoring their usage. Additionally, he/she gives the developers URIs to call the key vault from their systems.

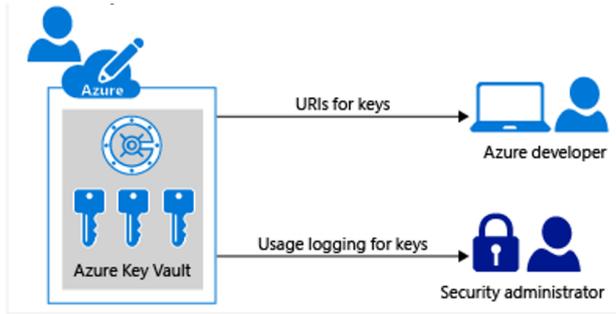


Figure 4. Azure KeyVault

The main application of Azure Key Vault in BI is concerned with the access permissions. Due to the fact that in the Azure cloud provider anyone can make a subscription and respectively to have a Tenant account, we could apply security principals by adding restrictions on read/write level of the certain Azure tenant ID. The security principal, in turn, is an object, which represents a user, service or application that requires access to Azure resources. Such kinds of applications could be SQL Server, BI tools or raw data output.

### B. Database Server security

As we already mentioned, the algorithm of BI consists of the process of extracting, transforming, storing, visualizing and analyzing data. That is why we need to take into account the possibilities to protect information with respect to data storing. In this case, we will emphasize on connection security, based on Azure SQL Server. In order to represent the actual rules application, we use **mydemoserver** (see Figure 5) as a server for a single SQL database in the Azure environment. We should add a new IP client in the rule, since we could not access the information in our database even though we are the administrator in the same subscription.

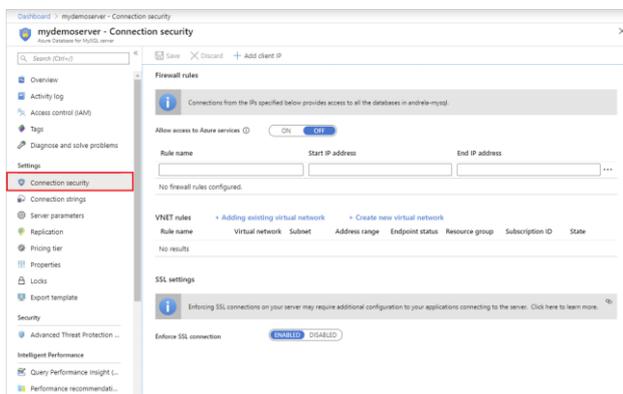


Figure 5. Connection Security in DB

For accessing certain database, we should apply Server-level or Database-level IP firewall rules, which need to be stored in the master database. To create IP rules in the Azure portal we need to enter name, Start IP address and End IP address (see Figure 6). Azure service provides the ability some of the saved clients' roles to be removed though time period.

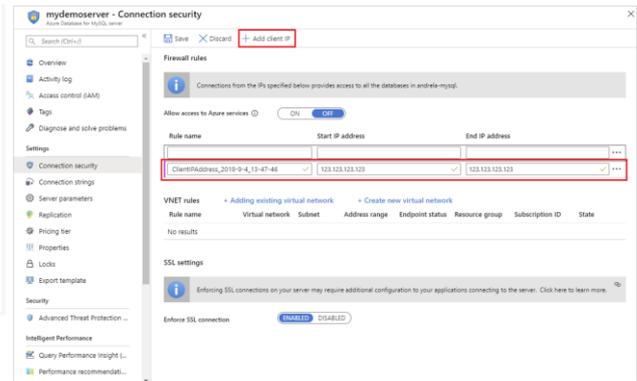


Figure 6. Creating an IP client

In order to validate the security rules, we tested the access point by entering the credentials for the database server in SQL Server Management Studio. The actual server name is blurred due to the confidential restrictions (see Figure 7). It could be noticed that with an IP address 87.97.205.43, which is different from saved one, we could not access the server or the database itself.

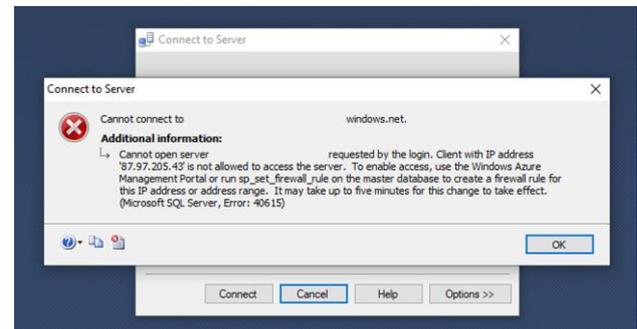


Figure 7. Connection Server Error

Working with data protection rules relies on different limitation options for inappropriate access. We need to pay attention not only to input, but also to output information. We will recall that one of the outputs could be a BI tool directly and that is why we are responsible for the provided information.

### III. SECURITY ON DATA VISUALIZATION LEVEL

The BI process ends with data visualization and analysis. There are plenty of existing software solutions which can be used for data implementation and most of them provide principles for data security on the end user level. In our paper, we will emphasize on data protection capabilities of two popular BI tools, such as Power BI and Tableau.

Power BI (PBI) is Microsoft's solution that gives business user-oriented data analysis and visualization capabilities to upgrade decision-making processes and the business visions [2]. Four features are available – Power Query, Power data, Power view and Power model, all of which provide data analysis efficiencies for enterprise users. Additionally, it maintains service in desktop, public browser and mobile environments. The end product of PBI is a report with different pages, each of which includes data visualizations.

Another popular BI tool is Tableau, which helps people to watch, observe, understand, and make decisions with a variety of data. Any type of graphs, plots, and charts can be made easily without programming [3]. As a result, Tableau builds reports by a combination of worksheets and dashboards, ordered in an appropriate way.

Both of the BI software solutions provided possibilities certain information to be hidden due to the business requirements. From the visualization point of view, for instance, restriction could be implemented on enterprise hierarchy, i.e. a regular employee will not have rights to see the salary of his/her colleagues in the company. Therefore, we will consider the role level of security in both of the tools as a data protection approach.

**A. Row Level of Security in Power BI**

Row Level of Security (RLS) in PBI is formed as a filter, which restricts data access at the row level and defines the specific representation of the data within enterprise roles. The RLS configuration for data models can be imported into Power BI Service or Power BI Desktop. For Analysis Services or Azure Analysis Services live connections, Row-Level Security can be configured in the model, not in Power BI Desktop. The security option will not show up for live connection datasets [4].

Additionally, it should be taken into account that the set RLS is applicable to the data model in each report and has no impact on the entire PBI workspace or application (see Figure 8).

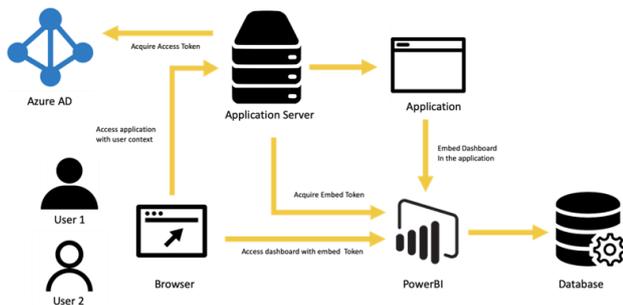


Figure 8. Role/Row Level of Security in Power BI

The process of RLS configuration in PBI starts with creating a role in Power BI Desktop and selecting a specific table for which to apply the filter. It should be a DAX rule that returns true or false. In order to define whether a user has rights to see a certain table, we have to extract the UPN with which he or she has entered the workspace [6]. By default, the row-level security filtering uses single-directional filters, regardless of whether the relationships are set to single direction or bi-directional. It is possible bi-directional cross-filter to be checked by selecting a relationship or ticking the “Apply security filter in both directions” option [7]. Even if we specify the logic for filtration, we should add a member account to the required role (see Figure 9). Surely, setting the

UPN restrictions are automatically valid in case the user table and/ or roles one are altered.

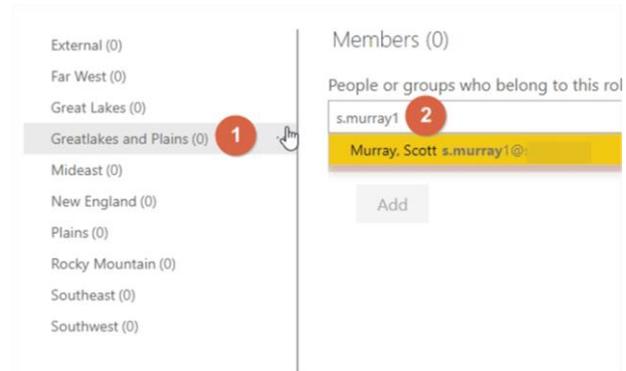


Figure 9. UPN name in a Member group

The result of implementing RLS can be observed in the security panel of Power BI Service. To test the data access of certain UPN, we can use the “View as Role” setting and simultaneously choose the required member group.

**B. Row Level of Security in Tableau**

Row Level Security (RLS) in Tableau refers to restricting the rows of data that a certain user can see in a workbook. This is different from Tableau permissions, which control the access to content and feature functionality.

Due to the fact that one of the characteristics of Tableau is server data administration, we should take additional steps to prepare it for RLS. To set up a Tableau Server content environment, it should be taken into account:

- Groups – sets of users who need the same type of access to content;
- Projects – containers for workbooks and data sources which generally represents a category of content;
- Permissions – sets of capabilities that define who can work with what content.

Moreover, in Tableau Public software the filtration can be based on the worksheet fields. As a result, similarly to the Power BI approach, the data visualization can be restricted for a specific username (see Figure 10).

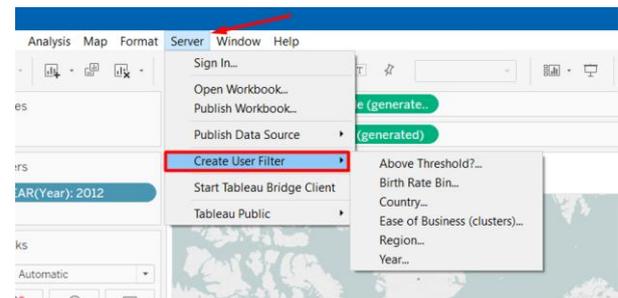


Figure 10. Role/Row Level of Security in Tableau

In Tableau, we can implement RLS by custom rule code. As an example, we will consider a deepest granularity filtration (see Code 4). By setting up an inner join between the view and the data, specific mapped rights are created. Multi-table extracts will evoke tables to match the joins, so creating the two views will simplify the resulting extraction [5]. We specify the WHERE clause in order to control the join.

```
SELECT *  
  
FROM data d INNER JOIN entitlements e ON  
  
d.attribute_a = e.attribute_a AND  
  
d.attribute_b = e.attribute_b AND ...  
  
WHERE e.username = USERNAME()
```

Code 4. Deepest granularity

The deepest granularity method can have a performance advantage when it is formed as a data hierarchy and it can be implemented as custom code in Tableau Desktop.

Role level of Security with the help of RLS allows data visualization to be protected from inappropriate sharing or representation. By this approach, the data, which has no security rules applied on the processing level, could be controlled on report level.

#### IV. CONCLUSION

There could be different approaches for security not only on the processing level, but also on a visual presentation basis. We should take into account that the information to be compromised in different ways and we should take actions before the confidential data is inappropriately shared or accessed.

#### ACKNOWLEDGMENT

This paper is partially supported by project MU21-FMI-007 of the Scientific Research Fund of Plovdiv University "Paisii Hilendarski", Bulgaria.

#### REFERENCES

- [1] Farseev, Aleksnadr & Samborskii, Ivan & Chua, Tat-Seng. bBridge: A Big Data Platform for Social Multimedia Analytics. *Proceedings of the 24th ACM international conference on Multimedia*, 2016, pp. 759-761. ISBN 978-1-4503-3603-1, DOI: 10.1145/2964284.2973836.
- [2] Bansal A., A. K. Upadhyay, Microsoft Power BI, *International Journal of Soft Computing and Engineering (IJSC)*, pp. 14-20, ISSN: 2231-2307, Volume-7 Issue-3, July 2017
- [3] <https://www.analyticssteps.com/blogs/why-is-tableau-used-for-read-the-working-and-key-features-of-tableau>, retrieved August 2021
- [4] <https://docs.microsoft.com/en-us/power-bi/admin/service-admin-rls>, retrieved August 2021
- [5] [https://help.tableau.com/current/server/en-us/rls\\_bestpractices.htm](https://help.tableau.com/current/server/en-us/rls_bestpractices.htm), retrieved August 2021
- [6] <https://www.biconnector.com/blog/power-bi-steps-to-setup-rls-row-level-security-static-and-dynamic/>, retrieved August 2021
- [7] <https://www.edureka.co/blog/row-level-security-in-power-bi/>, retrieved August 2021

# Adopting AR and Deep Learning for Gamified Fitness Mobile Apps: Yoga Trainer Case Study

M. Radenkovic\*, V. Nejkovic\*, N. Petrovic\*

\* Department of Computer Science, University of Nis, Faculty of Electronic Engineering, Nis, Serbia  
masa.radenkovic@elfak.rs, valentina.nejkovic@elfak.ni.ac.rs, nenad.petrovic@elfak.ni.ac.rs

**Abstract** – In this paper, we explore how state-of-art deep learning in synergy with augmented reality (AR) technology can be leveraged in area of fitness mobile applications adopting gamification approach. As outcome, we develop yoga instructor multiplatform mobile app relying on AR.js for marker-based mobile augmented reality in web browser and convolutional neural network (CNN) for yoga pose classification implemented in PyTorch. Moreover, some of the mobile application elements make use of AppSheet (menus and input forms) and Apps Script (trigger-based business logic), while the necessary data is stored within Google Sheets documents. Furthermore, the interaction of AR app with classification model is achieved relying on Flask. According to our results, the proposed architecture based on free-to-use cloud-based infrastructure shows acceptable performance in terms of response time. Finally, the classification model has been evaluated on publicly available dataset containing 1551 images for 5 yoga poses, reaching up to 71% correctly classified observations on test set (20% of the whole).

## I. INTRODUCTION

In recent years, due to increasing processing and graphic capabilities of smartphone devices, mobile augmented reality (AR) has become widely adopted in different areas – from education to healthcare. The main idea of augmented reality is merging realistic scenes (commonly recorded by camera) with virtual interactive 3D objects, while different outputs (texts, images, sounds, animations) can be generated as outcome of human-object interaction. This concept has been successfully applied in case of fitness mobile applications, as well [1].

On the other side, the involvement of gamification elements in AR fitness applications aiming to motivate the users to increase their physical activity [2-4] has emerged since the release of Pokemon Go [5] in 2016, reaching 500 million users back then. The goal of this game is to collect monsters and items used for battle with other players by walking and exploring the surrounding. Furthermore, recent studies show that 52.5% of currently available fitness-oriented mobile apps contain at least one element of gamification [4], approving the effectiveness of such approach.

In this paper, we explore the potential of adopting deep learning within AR fitness-oriented mobile applications. As a case study, we consider virtual yoga instructor with gamification elements. For marker-based augmented reality, we decide to use AR.js [6] web-based framework, while PyTorch [7] was used for implementation of deep learning image classification model based on convolutional neural network. User

interface and business logic of the application are implemented relying on AppSheet and Apps Script framework for rapid prototyping of multiplatform mobile apps, as described in our previous work [8].

There are several marker-based AR fitness apps in the existing literature and online stores. For example. PUMP FIT [9] is a marker-based AR gym instructor app. On the other side, Plaicise [10] is AR mobile app that makes use of body movement during exercises for control within simple games (jumping inside squares, collecting circles by jumping them and similar). However, it can be noticed that currently available fitness AR apps lack the integration with deep learning, which has enormous potential in this area as well.

## II. BACKGROUND

### A. AR.js

AR.js [6] is a lightweight web framework based on JavaScript that provides the necessary capabilities for development of augmented reality mobile applications. It covers the two main AR paradigms: 1) location-based [11]: GPS locations trigger objects to appear/disappear 2) marker-based [12]: 3D objects show up when printed, barcode-alike images are detected. Moreover, it includes the functionality of 3D primitive drawing, model loading and manipulation for various formats, such as web-friendly glTF which supports pre-defined animations. In this paper, we make use of marker-based approach, so 3D yoga instructor model appears when corresponding marker is detected.

The syntax of AR.js is intuitive, relies on special HTML tags which give the ability to define complex AR applications in just several lines of code. The highest-level tag for AR.js app represents virtual scene `<a-scene>` which consists of camera stream with additional 3D objects that appear over it. Furthermore, inside the scene, we define the markers which trigger the objects popping up, using tag `<a-marker>`. Various types of marker presets are supported such as standard *hiro*, *kanji* and numeric barcode markers. Finally, the objects that appear placed inside the marker tag using and denoted as `<a-entity>`, which can be either 3D primitive (such as `<a-box>`) or model loaded from external file. In the later case, it is necessary to define the value of `gltf-model` attribute containing the path to the 3D model file. In Fig. 1, an example of minimal marker-based AR.js application is shown. In this example, we have interactive 3D objects, each is assigned to marker and specific event handler is activated when we click on them.

```

<script>
// Event handlers go here
...
AFRAME.registerComponent('clickhandlers', {
  tick: function () { ... }
...
});
</script>
<body style='margin : 0px; overflow: hidden;'>
  <a-scene embedded arjs='sourceType: webcam;'>
    ...
    <a-marker id='markerN' preset='kanji' cursor='rayOrigin: mouse'>
      <a-entity gltf-model='./assets/models/yogali/scene.gltf'
        rotation='0 30 0' scale='1 1 1' clickhandlerN />
    </a-marker>
    ...
    <a-entity camera>
    </a-entity>
  </a-scene>
</body>
</html>
    
```

Figure 1. Example of minimal marker-based AR.js app

### B. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) [13] for image classification relies on convolutional layers which enable extracting feature map starting from raw image pixels. This kind of layer extracts spatial relations making use of small windows processing the input data, by applying filters, the so-called convolutional kernels. Convolutional layers aim to learn the filters as suitable as possible for recognizing specific objects and patterns. On the other side, the classification process is performed by fully connected layers. Fig. 2 illustrates the operation performed by convolutional layer over the given input image.

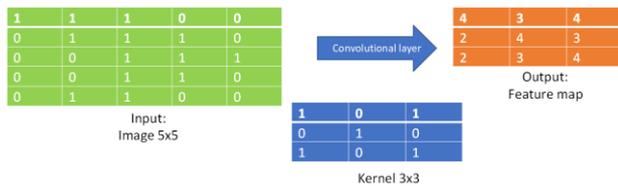


Figure 2. Illustration of convolution operation

The main parameters which define the feature map size are the following: 1) *depth* – number of filters used for convolution 2) *stride* – number of pixels between two centers of convolution filter operation in the input image 3) *padding* – adding zeros around the input corner 4) *kernel size* – dimensions of the applied convolutional filter. Formula which gives the feature map size starting from input size and previously described parameters is given as:

$$\frac{(input - kernel + 2 \cdot padding)}{stride} + 1 = input^* \quad (1)$$

Additionally, pooling layers performing the *downsampling* operation are placed between the convolutional in order to reduce the number of dimensions in the feature map, while preserving the most relevant information. This way, the processing time becomes shorter, especially for larger images. The most common types of pooling in CNNs are: 1) max pooling – returns the maximum value from the kernel area 2) average pooling – returns the average value of the pixels covered by kernel. For example, if we apply pooling layer with 2x2 kernel size over 4x4 input feature map, the output would be 2x2. In this paper, we rely on PyTorch [6] an object-oriented framework released by Facebook in 2016 for deep learning in Python, when it comes to CNN implementation.

### C. Flask

Flask [14] is Python framework aiming the development of web applications. It provides all the necessary functionalities for implementation of server-side in Python programming language, such as HTTP request handling and template rendering. Its main advantages are lightweight core with huge set of features and simple philosophy of writing code. Moreover, Flask enables the implementation of quite complex functionalities in less lines of code compared to other similar solutions, such as Django, requiring just several lines of boilerplate code. It is quite popular and LinkedIn is an example of web-based platform relying on Flask.

In this paper, Flask is used to achieve the interaction of AR.js application with PyTorch model deployed within Google Colab [15] execution environment. However, as Google Colab relies on isolated virtual machines, it is necessary to make the implemented services visible outside its environment. For this purpose, flask-ngrok tool is used [16-18], which makes the implemented services accessible via internet using REST API. In Fig. 3, an excerpt of Python code showing how Flask is used to receive the user-provided image forward it to classification model and return the prediction results. However, before that, the prediction model class inheriting Module has to be defined, corresponding object instantiated and model trained.

```

class YogaClassifier(torch.nn.module)
...
class YogaDataset(Dataset):
...
predictor=YogaClassifier()
#Training of image classification model
...
app = Flask(__name__)
run_with_ngrok(app)

@app.route('/yoga_classify', methods = ['POST'])
def analyze_image():
    if request.method == 'POST':
        uploaded_file = request.files['attachment1']

        if uploaded_file.filename != '':
            uploaded_file.save(uploaded_file.filename)

            img_name = uploaded_file.filename
            image = io.imread(img_name)
            image = dataset.transform(image)
            image=image.to(device)

            inputs = image.unsqueeze(0)
            outputs = net(inputs)
            pose_class=torch.argmax(outputs)

            return str(pose_class.item())

if __name__ == "__main__":
    app.run()
    
```

Figure 3. Excerpt of server-side Python code showing how the received image is forwarded to PyTorch image classification model

## III. IMPLEMENTATION

### A. Architecture overview

User accesses the AR application via web browser. After that, it is necessary to grant permissions for using camera by the app and video stream becomes visible on

the screen. After that, user is able to point camera to paper markers, which are detectable by the app. When marker is detected, the 3D model of fitness instructor specific yoga pose appears above it. By clicking on the model, user is redirected to AppSheet page providing an user interface for image submission. User has to take a picture of her/him repeating the given yoga pose, while the image is stored on Google Drive. Once the image is submitted, Apps Script trigger sending HTTP request with user's image as an attachment to Flask server run in Google Colab environment. Moreover, Flask server extract the image from the received request, transforms it to suitable form and forwards it to pre-trained yoga pose classification neural network. As outcome, neural network returns the predicted label of the user-provided image and forwards it as HTTP response. Finally, Apps Script triggers receives the response and checks the predicted label with the expected value stored in Google Sheets document. In case that the predicted and expected values match, the user's score kept in Google Sheets document will be updated by the trigger. Additionally, if user has enough points collected for performing the given exercises, it is possible to claim discount coupons through AppSheet interface which are sent to the provided mail address using Apps Script trigger, relying on Gmail service. The previously described working principle is depicted in Fig. 4.

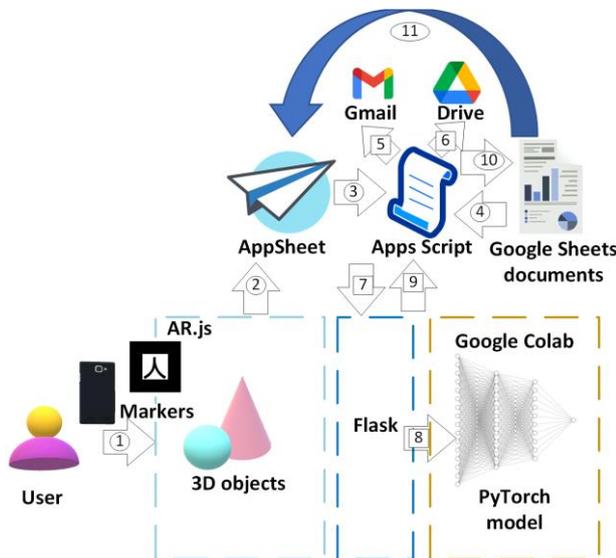


Figure 4. Architecture and generalized working principle overview: 1) HTML page of AR.js app 2) Click/touch on 3D object 3) Apps Script trigger activation after entering data into AppSheet form 4) Retrieval of necessary data from Google Sheets document while executing business logic triggers 5) Sending coupon e-mail 6) Image retrieval from Google Drive disk 7) HTTP request to the prediction model for given input image 8) Request routing and parameter forwarding to the corresponding Flask request handler 9) Prediction results returned as HTTP response 10) Storing the results of business logic outcome 11) Refreshing the view on data

### B. Image classification model

The image classification model is based on deep learning convolutional neural network. The first part consists of 5 hidden convolutional layers for feature map extraction, while each of them is followed by batch normalization, ReLU activation and max pooling.

Additionally, except the first layer, the other convolutional layers also have dropout layer before normalization. The role of dropout is to promote the independence of feature maps, by randomly setting some of the input channels to 0 with pre-defined probability  $p$ . After the convolutional layers, the obtained featured maps are forwarded to the fully connected part (or multilayer perceptron – MLP). The second part of the network actually performs the classification task and consists of 4 hidden layers. Before classification, the images are prepared by reducing their resolution to 256x256 in order to avoid long processing delays, especially in training mode. Table I summarizes the architecture of the CNN used for yoga pose classification in this paper. The model is able to distinguish between 5 different yoga exercise poses.

TABLE I. SUMMARY OF CNN ARCHITECTURE FOR YOGA POSE CLASSIFICATION

Layer type	Kernel	Stride	Padding	Input	Output	Channels	Activation	
Conv. 1	5x5	1	0	256x256	252x252	6	ReLU	
Max pooling 1	2x2	2	0	252x252	126x126	6		
Conv. 2	3x3	1	0	126x126	124x124	16		
Dropout, p=0.5								
Max pooling 2	2x2	2	0	124x124	62x62	16		
Conv. 3	3x3	1	0	62x62	60x60	32		
Dropout, p=0.3								
Max pooling 3	2x2	2	0	60x60	30x30	32		
Conv. 4	3x3	1	0	30x30	28x28	64		
Dropout, p=0.4								
Max pooling 4	2x2	2	0	28x28	14x14	64		
Konvolucio ni 5	3x3	1	0	14x14	12x12	128		
Dropout, p=0.5								
Max pooling 5	2x2	2	0	12x12	6x6	128		
Fully connected 1	-	-	-	6x6x128	256	-		
Fully connected 2	-	-	-	256	128	-		
Fully connected 3	-	-	-	128	64	-		
Fully connected 4	-	-	-	64	32	-		
Output	-	-	-	32	5	-	Softmax	

### C. AR mobile app

It is assumed that recognizable markers in printed form are put near the significant city locations, such as monuments, historical buildings and archeological sites. When user turns camera towards the showing some specific exercise pose pops up. The object itself is interactive and can be rotated by touching, while long click on it redirects to exercise submission form. Here, it is expected that users will provide camera-taken images showing how they successfully repeat the previously seen exercise. Screenshots of the previously described views in the app are given in Fig. 5.



Figure 5. Yoga instructor mobile app leveraging marker-based augmented reality

Once submitted, it is checked whether the pose on user-provided image matches with the correct answer for given yoga instructor marker. In case of correct answer, user's score is increased for a pre-defined value, depending on the pose difficulty. The points collected this way can be later used for prizes digital discounts received via email. The idea is to gamify the learning of new yoga exercises and keep the user motivated for learning new poses this way. In Fig. 6, the AppSheet views responsible for prizes-related functionalities are displayed.

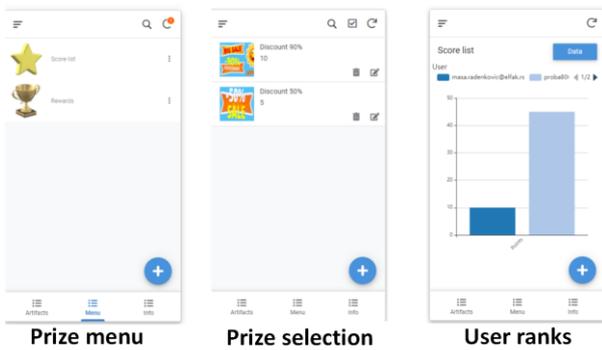


Figure 6. Prize claim menu in AppSheet

#### IV. EXPERIMENTS AND EVALUATION

In this section, we evaluate the implemented application considering two relevant aspects: response time and classification performance. When it comes to response delay, the crucial execution paths are covered: checking whether the user-provided pose on image is correct and reward claim. On the other side, the CNN model for pose image classification is evaluated on publicly available online dataset [19]. Based on this dataset, there are five possible poses which can be detected, as depicted in Fig. 7.



Figure 7. Five possible poses from yoga exercise dataset

In Table II, the experiments and their results are summarized. For each of the experiments, the following details are provided: 1) software components covered by the experiment 2) metric used for evaluation 3)

experiment conditions and parameter values 4) value of the selected evaluation metric for experiment.

TABLE II. EXPERIMENT DESCRIPTION AND EVALUATION RESULTS

Experiment	Component	Metric	Conditions	Value
1	Prize claim Apps Script trigger	Execution time [s]	User has enough points to claim the prize, so the coupon will be generated and sent using Gmail.	1.28 s
2	1. Pose check Apps Script trigger 2. Image classification	Execution time [s]	Correct pose image provided, so it also covers score update. Image label prediction on GPU.	4.73 s
3	Image classification model	Correctly classified images [%]	1240 training samples 311 test samples $\alpha=0.003$ , batch_size=128, num_epochs=14 Adam optimizer	71 %
		GPU training execution time [s]		227 s
		CPU training execution time [s]		368 s

Furthermore, to illustrate the convergence of classification model, we provide the graph showing how the number of correctly classified images changes through training epochs, as it can be seen in Fig. 8.

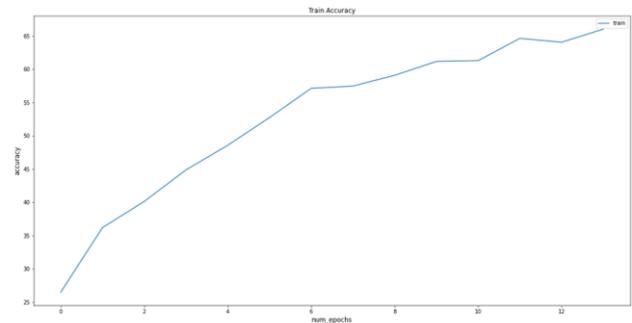


Figure 8. Percentage of correctly classified images against the number of training epochs

For evaluation purposes, we used Android smartphone Xiaomi Redmi Note 8 (4GB RAM, 2GHz octa-core processor). Regarding the classification model execution environment it was trained and deployed in cloud relying on Google Colab, using the randomly allocated high-performance GPU (Nvidia K80, T4, P4 or P100). When it comes to AppSheet and Apps Script components, they were deployed on Google's cloud infrastructure. The AR.js web app was delivered online using GitHub pages.

#### V. CONCLUSION AND FUTURE WORK

In this paper, we have shown how AR fitness-oriented mobile app can be integrated with services providing predictions relying on deep learning in case of gamified virtual yoga instructor. Furthermore, the adopted image classification model using CNN was evaluated on realistic dataset for 5 yoga poses, reaching up to 71% of correct classifications in test phase. The proposed set of technologies enables rapid prototyping relying on free cloud resources (however, with limited number of users).

In future, it we plan to explore the potential of integrating such approach within tourist sightseeing AR mobile apps in context of post COVID-19 tourism [20].

#### ACKNOWLEDGMENT

This work has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

#### REFERENCES

- [1] Augmented Reality in the Workout Industry [online], available on: <https://riseapps.co/fitness-augmented-reality/>, last accessed: 03/07/2021.
- [2] How AR VR Apps Are Gamifying Exercise And Workout? [online], available on: <https://appinventiv.com/blog/ar-vr-gamifying-fitness-workout-industry/>, last accessed: 03/07/2021.
- [3] H. R. Prial, "Gamification in Healthcare and Fitness application", preprint, pp. 1-4, 2020. <https://doi.org/10.13140/RG.2.2.29502.10566/>
- [4] Pokemon Go [online], available on: <https://pokemongolive.com/en/>, last accessed: 05/10/2021.
- [5] K. Shokurova, Fitness App Gamification In 2021: A Trend You Cannot Miss [online], available on: <https://medium.com/@kateshokurova/fitness-app-gamification-in-2021-a-trend-you-cannot-miss-778dc1e0b7e7>, last accessed: 15/08/2021.
- [6] AR.js - Augmented Reality on the Web [online], available on: <https://ar-js-org.github.io/AR.js-Docs/>, last accessed: 05/10/2021.
- [7] E. Stevens, L. Antiga, T. Viehmann, *Deep Learning with PyTorch*, Manning Publications, 2020.
- [8] N. Petrović, V. Roblek, M. Radenković, V. Nejković, "Approach to Rapid Development of Data-Driven Applications for Smart Cities using AppSheet and Apps Script", AIIT 2020 International conference on Applied Internet and Information Technologies, pp. 77-81, 2020.
- [9] A. Binoy, Anmol Srivastava, "PUMP FIT: An Augmented Reality Application Which Helps People with Their Workout More Efficiently", ISS '20: Interactive Surfaces and Spaces, pp. 91-93.
- [10] Plaicise: AR Fitness Games [online], available on: <https://apps.apple.com/by/app/plaicise-ar-fitness-games/id1535050335>, last accessed: 14/07/2021.
- [11] AR.js: Location Based [online], available on: <https://ar-js-org.github.io/AR.js-Docs/location-based/>, last accessed: 03/07/2021.
- [12] AR.js: Location Based [online], available on: <https://ar-js-org.github.io/AR.js-Docs/marker-based/>, last accessed: 05/10/2021.
- [13] S. Khan et al., *A Guide to Convolutional Neural Networks for Computer Vision*, Morgan & Claypool, 2018. <https://doi.org/10.2200/S00822ED1V01Y201712COV015>
- [14] Flask [online], available on: <https://flask.palletsprojects.com/en/2.0.x/>, last accessed 03/07/2021.
- [15] Google Colaboratory [online], available on: <https://colab.research.google.com>, last accessed: 05/10/2021.
- [16] How to run Flask app in Google Colab [online], available on: <https://www.geeksforgeeks.org/how-to-run-flask-app-on-google-colab/>, last accessed: 03/07/2021.
- [17] flask-ngrok [online], available on: <https://pypi.org/project/flask-ngrok/>, last accessed: 03/07/2021.
- [18] Deploying PyTorch in Python via Rest API with Flask [online], available on: [https://pytorch.org/tutorials/intermediate/flask\\_rest\\_api\\_tutorial.html](https://pytorch.org/tutorials/intermediate/flask_rest_api_tutorial.html), last accessed: 03/07/2021.
- [19] N. Pandit, Yoga Poses Dataset [online], available on: <https://www.kaggle.com/niharika41298/yoga-poses-dataset>, last accessed: 15/08/2021.
- [20] N. Petrović, V. Roblek, M. Khokhobaia, I. Gagnidze, "AR-Enabled Mobile Apps to Support Post COVID-19 Tourism", TELSIS 2021, pp. 1-4, 2021.

# Lean Production and Industry 4.0

Sanja Stanisavljev\*, Zlatko Košut\*, Saša Zec \*\*, Branko Markoski\*, Željko Stojanović\*

\* University of Novi Sad, Technical faculty "Mihajlo Pupin" in Zrenjanin, Republic of Serbia

\*\* Rosewood Hotel Georgia, Vancouver, British Columbia, Canada

sanja.stanisavljev@tfzr.rs, zlatko.kosut@tfzr.rs, sashazec@hotmail.com, branko.markoski@tfzr.rs,  
zeljkoelemir@gmail.com

**Abstract - The aim of this paper is to analyze and consider the advantages of modern concepts, available methodologies, and the benefits they provide, in order to point out the necessity of their application to gain survival and the advantages in a modern, dynamic, turbulent market. Modern, competitive production is namely the lean production that enables a timely response to market demands, one that achieves fast and efficient production by using modern information systems, digitization, robotics and the like. Companies need to be flexible, dynamic and innovative in every way possible.**

**Key words: Flexibility production, Lean concept, Industry 4.0.**

## I. INTRODUCTION

Organizations, if they want to survive, must adapt to changes and the needs of the market and customers. Traditional business and production methods must be replaced by new, modern and flexible ones.

Potential improvements or modernization of business activities, as well as the ways in which they can be implemented, are the subject of numerous discussions and represent an important challenge to which a society must pay appropriate attention if they want to ensure a prosperous future. One of the current concepts of improving and modernizing a wide range of business activities along the entire value chain are the concepts of Lean and Industry 4.0, which do not represent the immediate future, but the present that needs to be adequately addressed.

Efficient and flexible production is a modern production with more difficult companies aware of the necessity of timely response and responds to market demands dictated by the customer who expects a quality, functional product. The buyer does not want to pay for production errors and should be minimized. A stock accumulation should be avoided, which creates costs and produces the "right" time.

With rapid development in the fields of information technology and hardware, the world is about to witness a fourth industrial revolution [1].

Present traditional production processes are centralized, and Industry 4.0 will transform them into decentralized ones. Decentralization will enable network elements to make their own decisions, and to initiate certain activities based on those decisions, which will greatly change the production as we know today.

The concept of Industry 4.0 is going to change the way India manufactures, designs and refurbishes the products [2]. Driven by the power of big data, high computing capacity, artificial intelligence and analytics, Industry 4.0 aims to completely digitize the manufacturing sector.

In the literature, industry 4.0, which is also called "the industry of the future" or "smart manufacturing", does not seem to refer to a single, common definition. Nevertheless, experts agree that digital technology will play an essential role: industry 4.0 production processes will incorporate autonomous technology and tools that will communicate with one another throughout the value chain, thereby disrupting the role of operators and managers [3].

## II. LEAN PRODUCTION

Flexible production involves continuous and interrupted production flows, that is, the benefit and advantage of these flows, compensating for their shortcomings, with the aim of continuing the flows of materials within the production. This means that the system can respond to the requests of each individual customer in the shortest possible time, without stalling in the production, accumulation of unfinished products, and the like [4].

Pull production is the opposite of Push. This means that the products are made only when the buyer requests it or pulls it out, and not earlier. Accordingly, there is no production without the demands of customers. After the introduction of tactile production, single current flow of materials and reliable machines, the enterprise can introduce pull production, providing products at the moment when the customer needs them.

The lean transformation leader has the role of change agent, helping the organization and its leaders embed the principles of continuous improvement into organizational values to make sure everybody in the organization starts improving his or her processes [5].

In order for the entire Lean concept to be implemented, there are a number of methods and techniques that must be included within the company. The two basic principles – continuous flow and extraction system – must be fully met, while the rest are acquired through nurturing the Lean philosophy in the company, and everything is focused on continuous improvement.

If all the principles of the Lean concept are effectively implemented in the company and implemented on a daily basis, such a company will have numerous benefits, such as:

- reduced stocks;
- shortened production cycle time;
- increased operational readiness;
- increased product quality;
- increased efficiency of employees;
- reduced number of machine failures;
- increased utilization of machines and space;
- smaller warehouses.

One of the modern concepts available to manufacturing companies that want to achieve a competitive edge is certainly a lean concept. This concept seeks to eliminate all losses in production, avoids scattering of materials, resources available, with the aim of achieving high quality, flexibility and price that the buyer is ready to pay. It strives for continuity, optimization and integration of machines, people and other resources in the production process.

5 S is one of the tools used at the first level of lean maturity that helps setting the standards on how a workplace is organized. By identifying and demonstrating how a workstation should be organized to work efficiently, and how the work in process is managed, deviations from this way of working will be shown [5]. When implemented right, a deviation from the standards should lead to action before it leads to delays. A visual management tool that can prevent waste generation is the 5S tool.

Traditional methods such as Six Sigma  $6\sigma$  focus more on quality than it does at speed.

All major six sigma problems are solved using the DMAIC methodology (Define - Defining or Determining, Measure - Measurement, Analyze - Analyzing, Improve - Improving or Improvement and Control - control or management). This technique can be used to teach techniques: how to collect data, how much to use, and how often they count and how to be flexible.

The Six Sigma concept is a system that combines tools for continuous improvements that are on to make it possible to draw attention on processes, analyze them and compare each other, and objectively determine resources for those processes that require that they pay more attention to them.

### III. INDUSTRY 4.0

The beginnings of the Industry 4.0 concept are linked to the high-tech strategic plan of the German government, which was presented in 2006. However, the beginning of Industry 4.0 is marked in 2011, when a new, fourth industrial revolution was announced at the Hanover Fair, based on the further development of global industry through the production automation using modern (cyber-physical) means of production [6].

Contrary to the way the first three industrial revolutions took place, Industry 4.0 relies on breakthroughs in the ways the available information is being used. Since no specific event or epochal

technological breakthrough is associated with it, there are controversies whether it is a matter of revolution or evolution, such as a trend of improvement and modernization of existing processes. In general, Industry 4.0 encompasses the development and integration of innovative information and communication technologies in industry, with the aim of encouraging intelligent networking of processes (production and business) and products along the entire value chain [7].

A simple definition, proposed by Bidet-Mayer [8], describes "connected factories that are made flexible and smart through networks of machines, products and individuals". Schumacher et al. [9] provide a more detailed definition of Industry 4.0 by referring to recent technological developments where the internet and support technologies provide a backbone for incorporating physical objects, human beings, smart machines, product lines and processes across organizational boundaries in order to form a new sort of smart, networked, agile value chain. Industry 4.0 therefore implies constant communication with all of the parts of the value chain that are outside the company, especially suppliers and customers. Real-time sharing of information among all the different stakeholders makes the supply chain more responsive [10]. In addition to this horizontal integration of the value chain, there is also vertical integration: all the operations carried out in the company may now be connected and optimized in an integrated network [3].

Examples of Industry 4.0 are technologically innovative solutions that, in accordance with marketing and strategic activities, affect the change in the perception of modern business. Smart systems that contribute to the betterment are based on technology and the quality connections with the economy. The key business changes brought by Industry 4.0 are focused on reaction and communication speed, agility and flexibility, digital automation and the growth of trust. The elements that have been improved consequently affect the growth of productivity and the quality of the realization of operations [11].

Companies around the world, as well as governments, are working hard to implement advanced technologies into the industry's manufacturing processes and thus leverage new information and communication technologies (ICTs) in order to produce more productively and flexibly. The German government has marked the strategy of digitalization of production processes under the name "Industry 4.0", where it wants to maintain a leading position in the production and development of technologies and standards, so that they will be the first to export ready-made solutions. The answer to "Industry 4.0" is given by the USA under the name "Advanced Manufacturing Partnership 2.0", which aims to create high quality products, to create a renaissance of production, that is, to connect the industry with the Internet. The Government of Japan is responding to "Industry 4.0" with a strategy called "Revitalization and Robots Strategy" in order to increase productivity in industry by promoting the development and implementation of the robotics industry, as well as revitalizing the digital society and industrial sector. The

Chinese government is adopting a strategy called "Made in China 2025" which wants such technological development to make China a leading technological power by 2025, by applying innovations in order to improve global competitiveness, research and apply new jobs through production adjustment [12].

Regarding the fourth industrial revolution, i.e. "Industry 4.0", China cooperates with all industrialized countries. Thus, they are working with Germany on standardization, a German-Chinese alliance for vocational training and education has been created, as well as a common framework for the action plan named "Design innovation together!" [12] China also cooperates with other countries, such as Switzerland, France, etc. The Chinese government is taking all steps to make Chinese industry and its development play an important role in the development of the world market, in other words, to make China one of the most technologically advanced countries in the world by 2025.

It is estimated that in the next five years, digitalization will increase the annual income of the European industrial sector by more than 110 billion Euros. Gunther Oettinger, Commissioner for the Digital Economy and Society, says: "In order for Europe to maintain its leading position, it must successfully and rapidly digitize its entire industry." In the European Union, investments of 1.350 billion Euros in Industry 4.0 are planned by 2030, which means that it is committed to the digitalization and development of Industry 4.0. [13].

#### IV. CONCLUSIONS

We are witnessing a fast, dynamic development of technologies, especially information technologies that are increasingly being used in the production and changing the concept of the industry.

Industry 4.0 targets the implementation of interconnecting, smart, and self-controlled structures of processes and systems [5].

Changes are inevitable and companies must follow them, go one step ahead of them for survival and competitiveness. Only in that way and with a quick response to the requests and wishes of customers will the market.

It is critical to understand the role of Industry 4.0 in today's fast-changing and competitive business environment; where companies are facing challenges in dealing with big data and rapid decision making for the improved productivity [10].

On the changes and demands of the market and customers, in the sea of competitors, companies can respond with the improvement and efficiency of their production and process, which can be enabled by the application of six sigma, lean concepts and modern methods and techniques. In the world, the concept is increasingly being applied, with the tendency to achieve efficient and so-called fit production.

The concept of continuous improvement, adaptation, monitoring of changes, reduction and elimination of losses is the only correct and the right way to achieve

efficient production. By applying the modern, mentioned philosophies of the company, they can respond in a timely manner to the demands of customers who will be loyal and satisfied with the quality and price of the product.

Serbia is not the best example when it comes to the application of modern, flexible concepts, tools, Lean, and especially Industry 4.0. There is little interest and very modest results in the development of Industry 4.0. It is the orientation and aspiration to Industry 4.0 that could provide a chance for its revitalization. Shift and contribution to the development of Industry 4.0 may occur in some scientific institutions and IT companies where there are experts who are well versed in digital technologies, developing parts of industrial software.

If the industry wants to prosper, whether in production or in services, it is necessary to invest in equipment, to apply available methodologies, to strive for flexible production, which is the industry 4.0, while in some countries that can serve as an example and model, there is already a lot of talk about the industry 5.0.

#### REFERENCES

- [1] Reiner, A. (2014) Industrie 4.0 - Advanced Engineering of Smart Products and Smart Production. Int. Semin. High Technol. 1–14.
- [2] Bowman E. and Hurry D. (1993) Strategy through the Option Lens: An Integrated View of Resource Investments and the Incremental Choice Process, *Academy of Management Review*; October, 18 (4), 760-782.
- [3] FEL, Jennifer CAYLA, Valentina CARBONE, Can Industry 4.0 foster the backshoring of production? Insights from a Delphi study in France, 25<sup>th</sup> International EurOMA Conference, 2018, Budapest, Hungary.
- [4] Zelenović D., Projektovanje proizvodnih sistema, Fakultet tehničkih nauka, 272-286.
- [5] Thijs Panneman (2017), Lean Transformations, When and how to use lean tools and climb the four steps of lean maturity, pp 8. This Is What the Fourth Industrial Revolution Looks Like, April 2015, available at: [https://www.vice.com/en\\_us/article/4x3p43/life-after-the-fourth-industrial-revolution](https://www.vice.com/en_us/article/4x3p43/life-after-the-fourth-industrial-revolution)
- [6] Barreto, L. Amaral, A. Pereira, T. 2017. *Industry 4.0. implications in Logistics, Manufacturing Engineering*. Society International Conference 2017, Science Direct.
- [7] Bidet-Mayer, T. (2016), L'industrie du futur : unecompétitionmondiale, Paris, Presses des Mines
- [8] Schumacher A., Erol S., and Sih W. (2016), A Maturity Model for Assessing Industry 4.0.
- [9] Thoben, K.-D., Wiesner, S., and Wuest, T., (2017), "Industrie 4.0" and Smart Manufacturing – A Review of Research Issues and Application Examples", *Int. J. of Automation Technology*, Vol.11 No.1, p.4-16.
- [10] Stajić, N. 2020. *Industrija 4.0 u kontekstu upravljanja odnosima u lancu snabdevanja*. *Ekonomске ideje i praksa*. Broj 39. Decembar 2020.
- [11] Karabegović, I. i dr. 2018. *Implementacija „Industrije 4.0“ primjenom robotaidigitalnetehnologije u proizvodnim procesima u Kini*, *Tehnika – Mašinstvo* 67 (2018) Univerzitet u Bihaću, Tehničkifakultet, Bihać.
- [12] Nikola Marković, *Industrija 4.0 i budućnost*, PC Press 09/2016, str 8

# Use of CNNs on mobile devices to protect data from malware and unauthorized attacks

S. Mihajlović\*, D. Ivetić \*\*, I. Berković\*

\* Technical Faculty “Mihajlo Pupin”, Zrenjanin, Republic of Serbia

\*\* Faculty of Technical Sciences, Novi Sad, Republic of Serbia

mihajlovicsinisa25@gmail.com, ivetic@uns.ac.rs, ivana.berkovic62@gmail.com

**Abstract – Convolutional neural networks (CNNs) have a wide application, so they are also used to protect data from malware and unauthorized attack on mobile devices. Mobile devices are used and developed daily and have a wide range of software that can be used. Therefore, mobile devices are an easy target for malware. There are various methods in detecting malware, and they are all based on CNNs. The methods are very successful and effective in preventing and detecting malware. Mobile devices use different lock methods. Based on the trained model, CNNs improve data protection on mobile devices, thus reducing the possibility of unauthorized to data.**

## I. INTRODUCTION

With the daily development of mobile devices and deep learning, smart mobile applications that use deep learning technologies have emerged. In recent years, deep learning has achieved tremendous success in image processing, natural language processing, language analysis, and other fields of research. The resources required to run these models have increased significantly, and therefore pose a challenge for the application of such applications on resource-limited mobile devices. The development of smartphones, laptops and other new mobile devices has further promoted the development of AI (Artificial Intelligence) applications on mobile devices [1].

Mobile computing devices are very popular and powerful, and people want to embrace the benefits of CNN (Convolutional Neural Network) on their mobile devices [5].

## II. RELATED WORKS

A deep model has been defined that has been trained and applied to a specific service and is designed to run on mobile devices as a mobile deep learning application (MDLA). Its training can be cloud-based, or it can be based on edge devices that use federal learning technology. Cameras, microphones, and sensors can receive different types of information such as video, audio, and real-world acceleration. All this data is provided to MDLAs. MDLAs have evolved rapidly because of their benefits to users. MDLAs benefit users by performing malware detection, application recommendation, user verification, visual tasks for mobile devices, optimization of mobile internet browsing, human activity monitoring, medical health monitoring, and other smart fields. For network operators and third-party service providers, MDLA can support distributed machine

learning, federated learning, multiple smart IoT (Internet of Things) applications, and more. Intelligence in mobile applications is changing the way people live, work, and communicate with the world. While these apps are great for smartphones, they require a lot of storage space, calculations and high-power consumption and network bandwidth. These requirements are a constraint on mobile devices and prevent the development of MDLAs. CNN (Convolutional Neural Network) is the main method of mobile vision of tasks and is known for its great temporal and spatial complexity. The main idea of MDLA is to reduce the need for resources to perform deep learning tasks, or to design a deep learning framework suitable for mobile devices to optimize computing. To reduce resource requirements, the number of calculations and storage space required by the model itself is reduced. For different inputs, some calculation results are reused to reduce the demand for computer resources. Multiple deep learning tasks are performed on one mobile device [1].

Primary authentication on smartphones is based on PIN (Personal Identification Number), GUI (Graphical User Interface) passwords, fingerprint sensor and face unlock mechanisms. Users are either satisfied or do not use these mechanisms to protect private data and prevent unauthorized use of the device. One suggestion is a handwritten signature, a pseudo signature drawn on a touch screen smartphone for authentication. The advantage of using a signature is that the signature is not forgotten, and no one can imitate the signature of other people perfectly. If a signature is copied, that user is blocked from accessing the smartphone. The device verifies the signature with the already trained CNN model and approves the approach based on the classifier results [2].

Smart homes provide comfort, efficiency, convenience and secure connection of various appliances and devices. The appliances and devices can communicate with each other and can be remotely controlled by smartphones via the Internet. Smartphones carry sensitive and personal data and for this reason it is important to protect the security and privacy of smartphones and stored data. The Smartphones Sensor User Identification (SSUI) framework facilitates user identification based on the relationship between different types of sensor data and smartphone users. In SSUI, time and frequency characteristics are extracted and learned separately using CNN. CNN outputs are processed using recurrent neural networks (RNNs). The SSUI framework has an accuracy of 91.45% in different scenarios, on the Heterogeneity Dataset for Human Activity Recognition - HHAR [3].

With the daily use of smart devices, the number of applications is increasing. Almost all functions performed with the help of computers are performed more on mobile devices. Personal data is misused through malicious applications in the devices and these applications make the devices unusable. A model based on permissions to detect malicious applications on mobile devices running the Android operating system has been implemented. The AndroTracker dataset is used to identify whether Android-based apps are malicious or not. CNN is used as a learning algorithm. CNN achieved a success rate of 96.71% [4].

CNN can learn without first extracting features and fits well in the fast iteration of Android malware. The traditional solution for detecting Android malware requires constant learning through pre-selected features to maintain high-performance malware identification. A color-based Android malware detection system (R2-D2) has been developed. The system can convert bytecode from the Android archive to RGB (Red Green Blue) color code and save it as a fixed size color image. The color image is fed into the convolutional neural network for automatic feature extraction and training. Approximately 2 million benign and malicious Android applications for experiments have been collected [5].

Android is a mobile operating system that is increasingly the target of malware. Malicious applications are designed to turn mobile devices into bots as part of a larger botnet, so they have become quite common and pose a serious threat. The deep learning approach to discovering Android botnets is based on CNN. It is implemented as a CNN-based model that is trained in 342 static application functions to distinguish botnet applications from normal ones [6].

### III. MALWARE PROTECTION ON MOBILE DEVICES

Existing solutions for protecting privacy and security on mobile devices remain ineffective in many respects, and many analysts warn that malware families and their variants for Android are growing rapidly [14].

There are many techniques to avoid detecting most antimalware. For example, repackaging, an attacker decompiles a trusted application to obtain the source code, then adds malicious payloads and recompiles the payload application to make it available on various market alternatives, and sometimes on the official market [15].

Detection of signature-based malware, the most widespread technique adopted by free and commercial mobile antimalware, is often unintelligent. This is expensive because the process of obtaining and classifying malware signatures is laborious and time consuming [16].

Android malware detection technique based on dynamic analysis that considers the sequences of system calls that can occur in malware and legitimate applications. Behavioral similarities between malicious applications can be used to detect new malware [17].

Detection of malicious behavior for mobile applications can be considered as a special case of more general problems that reveal common phenomena and automatic learning of models from training examples is a very attractive way to solve this. The aim is to classify

such concurrent phenomena as true or false carriers of malicious behavior for applications that contain it. The CNN training input consists of pre-marked system call sequences, each of which represents a trace of the application and does not require special pre-processing. The network learns a set of reliability scores for each input sequence [18].

CNN can classify malicious mobile applications. For the classification process, the input units are rows composed of words, i.e., words consist of strings of special symbols, each representing a particular operation. A series of operations coded as specific symbols make up a sentence. Each row represents the behavior of the application via the Android system. In the training series, each row is marked with a "genre", assuming that one of the values is "pos" and "neg". Positively marked lines correspond to reliable applications, applications with negative markings indicate malicious ones. Each word in a row represents a system call extracted by the execution of the application. Among the features that characterize the detection of malware, there is a typology and sequence of system calls that are called during the application lifecycle. First, river system calls are built, so that each textual designation of the system call is represented by a numeric index, in this way each input line, which corresponds to a series of system calls made by the application can be represented as a numeric vector. Data vectorization is achieved on the TensorFlow platform [19].

The growth of Android malware is leading to improved access to malware detection. DeepClassifyDroid, is a new Android malware detection system based on deep learning. DeepClassifyDroid has a three-step approach to feature extraction, feature embedding, and deep learning-based discovery. The first and second steps perform a broad static analysis and generate five different sets of characteristics. The final step to detect malware based on convolutional neural networks. DeepClassifyDroid detects 97.4% of malware with several false alarms [22].

### IV. UNAUTHORIZED ATTACKS ON MOBILE DEVICES

The problem of detecting an attack consist in answering whether the captured biometric sample is original or not. Image data is transformed before the camera collects it on the user's device. The real samples are collected as one shot by directly photographing authentic user, in case of an attack the biometric sensor recaptures the previously captured image of the user, which is displayed on the screen. The camera captures and pre-processes the image of the target user's face, and when that image is displayed on the attack screen, it is further modulated using its own reproduction, geometric and reflective characteristics of the media. Face recognition is not trivial because it affects even the slightest variation in the user's facial features, and lighting conditions. Most methods are based on common assumptions and characteristics [7].

Based on liveness or motion detection. The archetypal method is eye blink detection, which can be effective if the attacker uses a photo but is easily circumvented by

repeated video playback attacks or even by cutting holes in a printed face image and simulating blinking with your own eyes [8].

The second class of methods attempts to detect the subtle movements of a living human face, using an estimate of the optical flow, an increase in motion, or a temporal extension of a low-level texture descriptor. Some methods use correlations of movement between foreground and background and other stage clues. If the attacker does not use fixed support when performing a print or display attack but would otherwise fail. These methods have the disadvantage of requiring a series of shots for a single prediction, and most of them can be bypassed by false eye blinking and careful handling of attack devices [9].

Based on physics or geometry. The face is usually a false representation of the user on a flat surface, which has different reflection properties compared to a living face. Some methods seek to detect this “flatness” or abnormal reflection by physical or geometric motivations. One method attempts to capture in-depth information using the Structure-from-Motion technique. It is proposed to detect differences in movement between areas of the face by optical flow estimation or explicit modeling of 3D projective invariants. Another option is to model local curvatures using multiple images. Methods require user collaboration to be successful [10].

Lambert’s model of reflection, it is necessary to model the interaction between illuminators and the reflecting surface to separate albedo and normal maps, which are then used to distinguish the right approach from attack patterns. Real-world lighting is mixed and uncontrolled, so assumptions are not valid in practice. It is possible to model diffuse and specular components [11].

Based on texture, noise analysis, or image quality. These methods seek to detect artifacts left by the re-capture process or to assess the degradation of the overall image quality. Texture characterization is usually motivated as a means of distinguishing the intrinsic textural properties of attack instruments and living faces, but it also encompasses other types of high-frequency information. Frequency-specific data can be captured by filtering Difference of Gaussians or Fourier analysis [12].

Methods based on low-level descriptor texture or high-frequency information can be effective in detecting paper texture and pattern sums, but efficiency depends on the exact shooting conditions and the camera's ability to solve fine details [13].

A method based on CNN training to distinguish between the original approach and attack images. The architecture considers detection of representation attacks, formulating the problem as a Class 2 classification and using only aligned face-to-face images for network training. The use of variable-resolution face patches during training reduces over-adaptation to user-specific characteristics and promotes the learning of more robust non-single-scale representations and a loss function that more closely models the goal [7].

A well-known problem is the smudge attack, when the user draws an unlock pattern to unlock the lock pattern on

the smartphone screen, the extraction of the pattern sometimes becomes difficult due to the presence of fatty residues around. Residues blur the phone screen, and significantly reduce the pattern of guessing the pattern lock. To address this, there is a CNN-based attack method. CNN shows high accuracy in image classification but using only CNN for an attack is not enough, as there are 389.112 possible patterns, and training CNN for all cases is difficult. There are two solutions, screen segmentation where the crane is divided into four segments to reduce the number of possible templates to 1470 in each segment. Another is the use of pruning rules, which reduces the number of total template cases by combining a pattern in each segment. Applying Android pattern lock restrictions reduces the number of possible patterns. Ideally, when the actual locking pattern is entered, without oil residues, the proposed scheme supports excellent performance [20].

The Deep Convolutional Neural Network (DCNN) is an architecture that aims to continuously authenticate on mobile devices. The complexity of the network is reduced by learning indirect traits such as gender and hair color instead identity. DCNN is based on multiple attribute detection tasks. The DCNN architecture is efficient in terms of speed, power consumption and accuracy on mobile devices [21].

Malware detection is basically a software classification problem based on information gathered from program analysis. The classification of Android applications is based on the system API (Application Programming Interface) call sequences and Deep Neural Networks (DNNs). The ability of DNNs to learn complex and flexible functions leads to the effective detection of malware [23].

## V. DISCUSSION

There are a lot of solutions to protect data from malware, as well as unauthorized attack and access to mobile devices. They are very efficient but require high performance. CNNs have excelled in classifying malicious malware as well as preventing unauthorized access. To prevent the spread of malware, antimalware needs to be constantly updated. The difference in malware affects performance and this applies to all classification methods. The more complex the application, the greater the diversity and thus the more difficult the classification. CNNs are used as a learning algorithm. Different approaches give different results and solutions.

Mobile phones are evolving every day and have better and better performance. Mobile phones are used daily, and people store and use their personal data on them, which is of great importance. There are various security methods for unlocking mobile devices and unauthorized access. The most used protection is the PIN because it is the easiest to remember, as is the template. The suggestion to use the user's signature is a great solution because the signature is unique and difficult to copy. The device verifies the signature based on the trained CNN model.

One of the common methods of protecting mobile devices is face recognition. There are several face recognition solutions. One solution can be tricked by using a printed image or a video that simulates blinking,

while the other solution requires a series of frames. Face recognition is affected by the slightest variation in the facial features of the user, and this is one of the biggest obstacles. It is necessary to find a solution that will recognize the face of the user in different situations and variations. With the help of CNNs and trained models, user attacks are prevented. To differentiate between 2D and 3D face recognition, and thus reduce the possibility of fraud and unauthorized access. The protection used daily is drawing a pattern to unlock mobile devices. The disadvantage is that greasy stains from the fingers remain, and it significantly reduces the accuracy of unlocking the phone. CNNs help in classification and thus increase the accuracy of unlocking.

The role of CNNs in protecting data from malware and unauthorized attacks is in the classification of data. CNN recognizes whether malware is benign or malicious based on trained models. CNN checks to see if the user unlocking the phone is accessing the phone authorized or unauthorized and trying to commit fraud.

It is recommended to continue in this direction because the results achieved so far are excellent. What needs to be addressed in further work is to minimize the demand for resources of mobile devices. In addition, applications should be as simple as possible for the user and not have too many requirements to run on the devices themselves. In terms of preventing unauthorized attack and access, it is necessary to reduce the amount of data required for the user to enter, and to keep the process as short as possible. It should go in the direction of using something unique and difficult to falsify for authorization, and thus prevent fraud. It is extremely difficult to maintain high accuracy, precision, speed, performance, and to keep resource requirements to a minimum.

## VI. CONCLUSION

Convolutional neural networks are evolving daily and have different applications. It has also found its application on mobile devices. The paper describes applications on mobile devices in the prevention of malware, as well as the prevention of unauthorized access to mobile devices. Both applications have one thing in common, and that is to protect the data of mobile device users.

Mobile devices are used every day, and where people store their personal data. There are malicious programs that can damage data on mobile devices. In addition to malware, people are struggling to prevent unauthorized access to personal data on mobile devices. There are many solutions that succeed in this, but there is always room for improvement. Just as antimalware is being promoted, so are malware, and the same is true for unauthorized access methods. By classifying the data and training the CNNs model, positive results are obtained, and great success and accuracy is achieved. The need for resources needs to be reduced because mobile devices are compact, and small, and do not have as much power as computers.

## ACKNOWLEDGMENT

Ministry of Education, Science and Technological Development, Republic of Serbia financially supported

this research, under the project number TR32044: "The development of software tools for business process analysis and improvement", 2011-2021.

## REFERENCES

- [1] Y. Wang, J. Wang, W. Zhang, Y. Zhan, S. Guo, Q. Zheng, X. Wang, A survey on deploying mobile deep learning applications: A systemic and technical perspective, *Digital Communications and Networks*, 2021.
- [2] C. J. Jose, M.S. Rajasree, User Authentication for Mobile Phones by Pseudo Signature Using CNN, Ranganathan G., Chen J., Rocha A. (eds) *Inventive Communication and Computational Technologies. Lecture Notes in Networks and Systems*, vol 89. Springer, Singapore, 2020.
- [3] Z. Qin et al., Learning-Aided User Identification Using Smartphone Sensors for Smart Homes, *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7760-7772, Oct. 2019.
- [4] A. I. Karabey, Using convolutional neural network for Android malware detection, *The 17th INTERNATIONAL SCIENTIFIC CONFERENCE INFORMATION TECHNOLOGIES AND MANAGEMENT*, Vol. 23, pp. 29-35, 2019.
- [5] T. H. Huang and H. Kao, R2-D2: Color-inspired Convolutional Neural Network (CNN)-based Android Malware Detections, *IEEE International Conference on Big Data (Big Data)*, pp. 2633-2642, 2018.
- [6] S. Y. Yerima and M. K. Alzaylaee, Mobile Botnet Detection: A Deep Learning Approach Using Convolutional Neural Networks, *International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, pp. 1-8, 2020.
- [7] Almeida, Waldir & Andaló, Fernanda & Padilha, Rafael & Bertocco, Gabriel & Dias, William & Torres, Ricardo & Wainer, Jacques & Rocha, Anderson, Detecting face presentation attacks in mobile devices with a patch-based CNN and a sensor-aware loss function, *PLoS ONE*. 15. e0238058, 2020.
- [8] P. Gang, L. Sun, Z. Wu, W. Yueming, Monocular camera-based face liveness detection by combining eyeblink and scene context, *Telecommunication Systems*, Vol. 47, pp. 215-225, 2011.
- [9] W. Bao, H. Li, N. Li and W. Jiang, A liveness detection method for face recognition based on optical flow field, *International Conference on Image Analysis and Signal Processing*, pp. 233-236, 2009.
- [10] K. Kollreider, H. Fronthaler, J. Bigun, Evaluating liveness by face images and the structure tensor, *IEEE Workshop Automat. Identification Advanced Technol*, pp. 75-80., 2005.
- [11] M. De Marsico, M. Nappi, D. Riccio and J. Dugelay, Moving face spoofing detection via 3D projective invariants, *5th IAPR International Conference on Biometrics (ICB)*, pp. 73-78, 2012.
- [12] J. Li, Y. Wang, T. Tan, A.K. Jain, Live face detection based on the analysis of fourier spectra, *Biometric Technol. Human Identification*, Vol. 5404, pp. 296-304, 2004.
- [13] G. Kim, S. Eum, J.K. Suhre, D.I. Kim, K.R. Park, J. Kim, Face liveness detection based on texture and frequency analyses, *IAPR Int. Conf. Biometrics*, pp. 67-72, 2012.
- [14] C. Marforio, F. Aurelien, C. Srdjan, Application collusion attack on the permission-based security model and its implications for modern smartphone systems, *CTIT technical reports series 724*, 2010.
- [15] R. Fedler, J. Schütte, On the Effectiveness of Malware Protection on Android An evaluation of Android antivirus, 2013.
- [16] X. Jiang and Y. Zhou, *Android Malware*, SpringerBriefs in Computer Science, 2013.
- [17] G. Canfora, F. Mercaldo, C.A. Visaggio, P. Di Notte, *Metamorphic Malware Detection Using Code Metrics*, *Information Security Journal A Global Perspective*, Vol. 23, pp. 57-67, 2014.
- [18] F. Mercaldo, V. Nardone, A. Santone, C. A. Visaggio, Hey malware, i can find you!, *IEEE 25th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pp. 261-262, 2016.

- [19] F. Martinelli, F. Marulli, F. Mercaldo, Evaluating Convolutional Neural Network for Effective Mobile Malware Detection, *Procedia Computer Science*, Vol. 112, pp. 2372-2381, 2017.
- [20] H. Shin, S. Sim, H. Kwon, et al., A new smart smudge attack using CNN, *Int. J. Inf. Secur*, 2021.
- [21] P. Samangouei, C. Rama, Convolutional neural networks for attribute-based active authentication on mobile devices, pp. 1-8, 2016.
- [22] Y. Zhang, Y. Yang, and X. Wang, A Novel Android Malware Detection Approach Based on Convolutional Neural Network, *ICCSP 2018: Proceedings of the 2nd International Conference on Cryptography, Security and Privacy*, pp. 144-149, 2018.
- [23] R. Nix and J. Zhang, Classification of Android apps and malware using deep neural networks, *International Joint Conference on Neural Networks (IJCNN)*, pp. 1871-1878, 2017.

# Creating a mobile application using the Kotlin programming language

B. Babić\*, E. Brtko\*, I. Vecštejn\*

\* University of Novi Sad, Technical faculty "Mihajlo Pupin", Đure Đakovića bb, 23000, Zrenjanin, Republic of Serbia  
bojanbabic023@gmail.com; eleonorabrtka@gmail.com; igor.vecstejn@gmail.com

**Abstract - Mobile applications are a rapidly growing segment of the global mobile market. In this paper, Android mobile platform for the mobile application development, using the Kotlin programming language is discussed. If you are familiar with Java it won't be hard to understand Kotlin.**

## I. INTRODUCTION

Android Studio is the official Integrated Development Environment (IDE) for Android app development, based on IntelliJ IDEA. On top of IntelliJ's powerful code editor and developer tools, Android Studio offers even more features that enhance your productivity when building Android apps. Since the first release of Google's mobile operating system, developers have been developing applications mostly using Java and, in some specific scenarios, using C++. Developers use the official IDE, named Android Studio, and a SDK (Software Development Kit) provided by Google to build Android applications. The SDK compiles the Java code into Dalvik bytecode, which is then packaged on an apk. Then, developers submit those apks to application stores, such as the official, named Google Play Store, which currently has more than 3 million applications in its official store [1]. Android users can install applications into their Android devices by downloading those apks directly from the applications stores. Despite the Android operating system's success, different development approaches and frameworks have emerged to ease mobile applications' development during the last years [2, 3]. For instance, approaches such as PhoneGap/Cordova [4] from Adobe, Xamarin from Microsoft [5], React-Native from Facebook [6], and more recently Flutter from Google [7], aim at facilitating the development of multi-platform mobile applications by allowing developers to write applications using a non-native programming language and then to obtain a version of a native version for each platform (Android and iOS). Meanwhile, Google and Apple continue evolving their development toolkits to build native applications to avoid mobile developers migrating to such third-party development frameworks [8].

In an attempt to ease mobile programming practice, in June 2014, Apple released Swift, a modern, multi-paradigm language that combines imperative, object-oriented, and functional programming for developing iOS applications [9]. Similarly, in 2017 Google announced the Kotlin programming language as an officially supported language for Android development [10]. Kotlin is a pragmatic programming language that runs on the Java

Virtual Machine (JVM) and it is fully interoperable with Java because both Java and Kotlin code is compiled to JVM bytecode. Consequently, it is possible to mix Kotlin and Java code in the same application, call Kotlin code from Java code, and the opposite. Kotlin provides a different approach to write applications because it combines object-oriented and functional features, some of them not present in Java or not available for Android development [11].

The paper describes an android application that was made in the Android Studio and which is connected to a database that is located in the cloud, Firebase. The programming language used for the application is Kotlin.

## II. BACKGROUND

An Android app is a software application running on the Android platform. Because the Android platform is built for mobile devices, a typical Android app is designed for a smartphone or a tablet PC running on the Android OS. Developers may download the Android software development kit (SDK) from the Android website. The SDK includes tools, sample code and relevant documents for creating Android apps [12].

The main concepts that make up an Android application are:

- **Activities:** An activity is a basic user task, such as adding a contact to an agenda or taking a photo.
- **Views and View Groups:** the view is a widget for the interface, such as a button or text entry; views are grouped into groups of views, which represent hierarchical organizations of appearance and content.
- **Intent:** intent is a specification of a request for action.
- **Events and event listeners:** an event is a phenomenon within an activity, which can be managed directly by a business method connected to the display element that produced the event [13].

### A. Android Studio

Android Studio is the official integrated development environment (IDE) for Google's Android operating system, built on JetBrains' IntelliJ IDEA software and designed specifically for Android development. It is

available for download on Windows, macOS and Linux based operating systems. It is a replacement for the Eclipse Android Development Tools (ADT) as primary IDE for native Android application development [14].

### B. Kotlin

Kotlin is one of many Java Virtual Machine (JVM) languages and was developed by JetBrains, the maker of IntelliJ IDEA. Previously, developers struggled to update Java quickly as it was almost impossible to make bigger changes in syntax due to backward compatibility.

While the developers of JetBrains began Kotlin in 2010, it took over five years to release Kotlin version 1.0. One year later, in 2017, at the Google I/O conference, Kotlin was announced as an official language for Android development.

Nowadays, Kotlin is the second most used JVM language and is increasing in popularity due to the fact that it is developed under an open-source license and is easy to use. Developers also like Kotlin for it's:

- Expressiveness
- Interoperability
- Language features
- Structured concurrency [15]

### C. Firebase

Firebase is a toolset to “build, improve, and grow your app”, and the tools it gives you cover a large portion of the services that developers would normally have to build themselves, but don't really want to build, because they'd rather be focusing on the app experience itself. This includes things like analytics, authentication, databases, configuration, file storage, push messaging, and the list goes on. The services are hosted in the cloud, and scale with little to no effort on the part of the developer [16].

Meaning “hosted in the cloud” is that the products have backend components that are fully maintained and operated by Google. Client SDKs provided by Firebase interact with these backend services directly, with no need to establish any middleware between your app and the service. So, if you're using one of the Firebase database options, you typically write code to query the database in your client app. This is different than traditional app development, represented in fig. 1. which typically involves writing both frontend and backend software. The frontend code just invokes API endpoints exposed by the backend, and the backend code actually does the work. However, with Firebase products, the traditional backend is bypassed, putting the work into the client. Administrative access to each of these products is provided by the Firebase console.

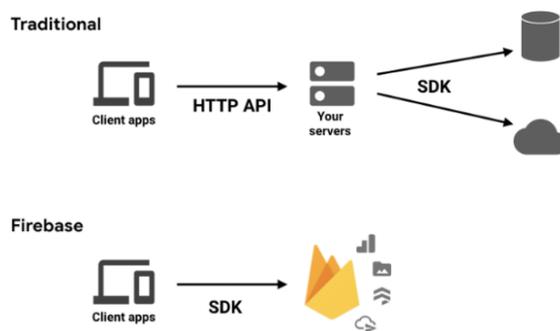


Figure 1. The difference between a traditional database and a Firebase database

## III. INSTAGRAM CLONE

### A. Starting the project

Before you start anything you need to have a created project with Bottom navigation activity. You need to change the color of your application because we don't want it to be by default. All icons and pictures that we used in the project must be in the project folder called “drawable”. After that all dependencies and plugins need to be set at the exact place in the project and with the wright version of that plugin or dependence. After all this you can connect to Firebase real time data base and Firebase storage for storing your posts and profile pictures.

### B. Fragments

A fragment represents a modular portion of the user interface within an activity. A fragment has its own lifecycle, receives its own input events, and you can add or remove fragments while the containing activity is running [17].

- HomeFragment - This fragment shows us posts that the people you follow are posting. The logic is such that it checks whether the person who posted the post is on the list of followers, and if so, his post is displayed on the home page.
- SearchFragment - This fragment is for searching app users who have an account. The search is programmed so that as soon as we start typing someone's name or username it will show us all users with those letters even though we have not finished typing. When someone's name is typed, the program turns everything into lowercase letters and checks it that way.
- ProfileFragment - This fragment shows the current user. If so, then we have options for setting up an account as well as for logout. You can also see the numbers for how many people we follow, how many people follow us and the number of posts. All this is stored in a real time database from where the program takes all the necessary information to make it easier for the user.

- NotificationFragment - This fragment is for getting messages when someone comments on us, likes the post.

### C. Layout

A layout defines the structure for a user interface in your app, such as in an activity. All elements in the layout are built using a hierarchy of View and ViewGroup objects. A View usually draws something the user can see and interact with. Whereas a ViewGroup is an invisible container that defines the layout structure for View and other ViewGroup objects [18]. In fig. 2. we can see an example of layout on profile page.

The unit of measurement used for all measurements in the layouts is dp<sup>1</sup>.

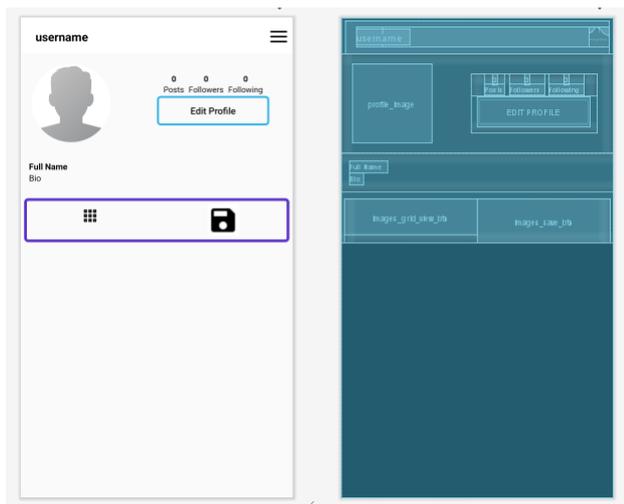


Figure 2. Example of layout, profile page (fragment)

### D. Activities

An activity is a single, focused thing that the user can do. Almost all activities interact with the user, so the Activity class takes care of creating a window for you in which you can place your UI with #setContentView. While activities are often presented to the user as full-screen windows, they can also be used in other ways: as floating windows, Multi-Window mode or embedded into other windows. There are two methods almost all subclasses of Activity will implement:

- #onCreate is where you initialize your activity. Most importantly, here you will usually call setContentView(int) with a layout resource defining your UI, and using findViewById to retrieve the widgets in that UI that you need to interact with programmatically.
- onPause is where you deal with the user pausing active interaction with the activity. Any changes made by the user should at this point be committed. In this state the activity is still visible on screen [19].

<sup>1</sup> dp-Density-independent pixels. An abstract unit of measurement based on the physical display density of the device in relation to the baseline of the screen of 160 dpi.

- SignUpActivity - We examine what the user has compressed and whether he has entered all the data correctly. In case the user already has an account, he can click “Already have an account? Sign in ”and the app will take you to the login page. If we do not have an account, it is necessary to enter the appropriate data. The program takes the data entered by the user as the basis of his account, checks it and stores it in a database. The user in the database has his own ID that this application generates itself, full username, username, biography (something about himself) and profile picture.
- SignInActiviy - Logging into the application is done after a successfully created user account in SignUpActivity. When logging in, the program checks the entered e-mail values and passwords with the database. If successful, the program will log you into your account in a few seconds.
- MainActivity - This activity is the main activity. This means that other activities of this program are called from MainActivity. Part of the code of this activity are Bottom Navigation and there are consists of:
  - If the home icon is clicked, it will display the HomeFragment with all new companion posts.
  - If the search icon is clicked, the program will switch to the SearchFragment to search for the users we want to track.
  - If the add icon is clicked, the program will switch to the AddPostActivity, activity where images can be added and new user posts posted.
  - If the notifications icon is pressed, the program will switch to the NotificationFragment fragment.
  - If the profile icon is clicked, the program will switch to the ProfileFragment fragment, where we set the appearance and description of the user who uses it.
- AddPostActivity - Adding a post is coded so we have to enter an image, the comment is not so important. When an image is published, that image with attributes from the class model appears in the database for publication.
- AccountSettingActivity - With user profile options, you can change the image, profile description (biography) and logout from the profile. When changing the profile picture, the option to crop the picture as desired and to write a biography has been added. In case

both items change, we have the `uploadImageAndUpdateInfo` function that changes and saves the new biography and the new image to the database. In case only the biography is changed, there is a function that changes only the biography of the `updateUserInfoOnly` user. The default activity itself has a `userInfo` function that displays a description and profile picture from the database. This means that every time the profile is changed everything will be up to date. When logging out, it switches to the page where the user logs in to their user account (`SignInActivity`).

### E. Models

A class model is a class where we describe the attributes of that class in order to connect them to the appropriate adapter and perform activities on that class through it. In our case, these are the `User` and `Post` classes.

- `User` - This model represents the user who has his default constructor and the constructor that receives the attributes of the class, as well as the get and set methods for those attributes. Attributes that describe the user are `username`, `fullname`, `biography`, `image`, and `unique user number (uid)`. Each of these attributes is also found in an online database. We also have an email but it is not among the attributes because we only need it to log in to the account. In fig. 3. we can see the appearance of user data from the real time database.

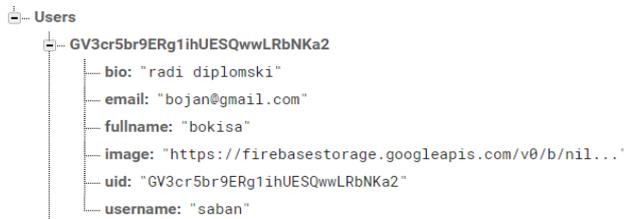


Figure 3. Appearance of user data from the real time database

- `Post` - The post model is connected to an adapter that performs activities on the publishing data. The default constructor, the constructor with passed publish attributes, and the get and set methods are included. The attributes that describe the post are the unique post number, the user who posted, the description, and the postimage. All of these attributes also describe the `RealTime` database on the Internet. In fig. 4. we can see the appearance of publication data from `RealTime` database.



Figure 4. Appearance of publication data from RealTime database

### F. Adapters

An Adapter object acts as a bridge between an `AdapterView` and the underlying data for that view. The Adapter provides access to the data items. The Adapter is also responsible for making an `android.view.View` for each item in the data set [20].

- `UserAdapter` - The user adapter is used to regulate what happens in the database. Used wherever user iteration is enabled. Its function is to check and monitor what the user is doing, e.g. if a user wants to track someone, that person must be added to their watchlist from the database. Conversely, if we no longer want to track it, the adapter must remove it from the watchlist in the database.
- `PostAdapter` - The attributes of one post are described with this adapter. It displays the posts stored in the database under the corresponding variables and displays them e.g. used for `HomeFragment`. When posting on `HomeFragment`, the posts described in `PostAdapter` will be displayed, that is: who posted that post, his profile picture, their username and full username.

### G. Preview of finished application

When launching the application and when successfully logging in to the user account, all the posts of the people we follow will be posted on the home page, we can see that in fig. 5. As it is explained how user search works, a picture is shown which obviously shows that as soon as we start typing a word, it throws out users with those letters of the username.

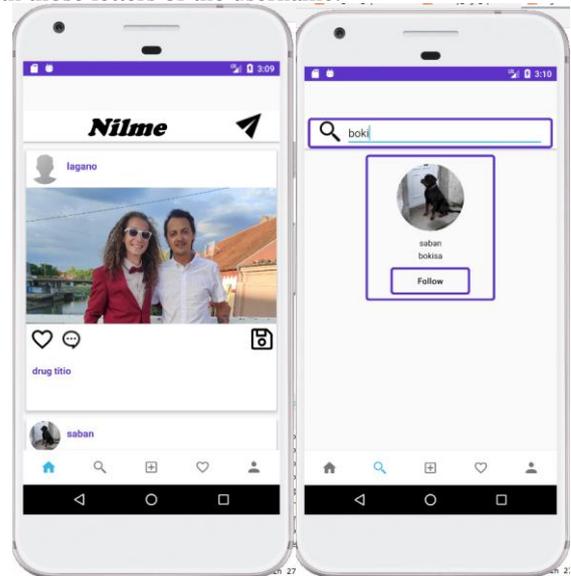


Figure 5. Preview of posts and search function

## IV. CONCLUSION

The key advantages of Kotlin include excellent code, best performance, less runtime and useful libraries that made it the first choice for programs compared to Java.

All its rich functions and advantages have made it the future program of Android applications.

With less coding and debugging, it speeds up the development process as well as the use of existing codes, and makes it easier for developers to learn and implement into existing or new projects. This first-class official programming language aims to provide more opportunities for developers by setting up a non-profit foundation in partnership with Google to preserve the language.

We will not know for sure whether Kotlin is the future of mobile development or not. But for now, you should definitely take it as a language to learn. It is extremely flexible, powerful and easy to learn. It also shows no signs of slowing down any time soon, so 2021 is the ideal time to learn Kotlin.

#### ACKNOWLEDGMENT

This research is financially supported by Ministry of Education and Science of the Republic of Serbia under the project number TR32044 “The development of software tools for business process analysis and improvement”, 2011-2021.

#### REFERENCES

- [1] AppBrain. Android and Google Play statistics. <https://www.appbrain.com/stats>. (Accessed on 1/10/2021).
- [2] Meiyappan Nagappan and Emad Shihab. “Future Trends in Software Engineering Research for Mobile Apps”. In: 2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER) (2016), pp. 21–32. issn: 9781509018550. doi: 10.1109/SANER.2016.88. url: <http://ieeexplore.ieee.org/document/7476770/>.
- [3] Matias Martinez and Sylvain Lecomte. “Towards the Quality Improvement of Crossplatform Mobile Applications”. In: Proceedings of the 4th International Conference on Mobile Software Engineering and Systems. MOBILESoft '17. Buenos Aires, Argentina: IEEE Press, 2017, pp. 184–188. isbn: 978-1-5386-2669-6. doi: 10.1109/MOBILESoft.2017.30.
- [4] Apache. Apache Cordova Official Website. <https://cordova.apache.org/>. (Accessed on 1/10/2021).
- [5] Microsoft. Xamarin | Open-source mobile app platform for .NET. <https://dotnet.microsoft.com/apps/xamarin>. (Accessed on 1/10/2021).
- [6] Facebook. React Native • A framework for building native apps using React. <https://reactnative.dev/>. (Accessed on 1/10/2021).
- [7] Google. Flutter - Beautiful native apps in record time. <https://flutter.dev/>. (Accessed on 1/10/2021).
- [8] Bruno Gois Mateus. Towards high-quality Android applications development with Kotlin. Mobile Computing. Université Polytechnique Hauts-de-France, 2021.
- [9] M. Rebouças, G. Pinto, F. Ebert, W. Torres, A. Serebrenik, and F. Castor. “An Empirical Study on the Usage of the Swift Programming Language”. In: 2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER). Vol. 1. Mar. 2016, pp. 634–638. doi: 10.1109/SANER.2016.66.
- [10] Mike Cleron. Android Developers Blog: Android Announces Support for Kotlin. <https://android-developers.googleblog.com/2017/05/android-announces-support-for-kotlin.html>. (Accessed on 1/10/2021). May 2017.
- [11] Comparison to Java Programming Language. Online; accessed 01-October-2021. June. 2021. url: <https://kotlinlang.org/docs/reference/comparison-to-java.html>.
- [12] What Does Android App Mean? Online; accessed 01-September-2021. April 2012. url: <https://www.techopedia.com/definition/25099/android-app>
- [13] Marco Brambilla, Piero Fraternali, Interaction Flow Modeling Language 1st Edition, Morgan Kaufmann, 2014, (url: <https://www.elsevier.com/books/T/A/9780128001080>)
- [14] Android Studio. [https://android.fandom.com/wiki/Android\\_Studio](https://android.fandom.com/wiki/Android_Studio) (Accessed on 1/9/2021).
- [15] Java Basics: What Is Kotlin? Online; accessed 02-September-2021. January 2021. url: <https://www.jrebel.com/blog/what-is-kotlin>
- [16] What is Firebase? The complete story, abridged. Online; accessed 02-September-2021. September 2018. url: <https://medium.com/firebase-developers/what-is-firebase-the-complete-story-abridged-bcc730c5f2c0>
- [17] Create a fragment. <https://developer.android.com/guide/fragments/create> (Accessed on 3/9/2021).
- [18] Layouts. <https://developer.android.com/guide/topics/ui/declaring-layout> (Accessed on 3/9/2021).
- [19] Activity. <https://developer.android.com/reference/kotlin/android/app/Activity> (Accessed on 3/9/2021).
- [20] Adapter. <https://developer.android.com/reference/kotlin/android/widget/Adapter> (Accessed on 3/9/2021).

# Review of challenges in identifying microservices from software artifacts

A. Stojkov and Ž. Stojanov

University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Republic of Serbia  
aleksandra.stojkov@tfzr.rs, zeljko.stojanov@uns.ac.rs

**Abstract - Microservices are a well-known and widely used architecture pattern nowadays. They can be used to build a completely new system, but also for transitioning old legacy systems for achieving better operation and maintenance performances. In this paper, the analysis of literature sources is performed in order to identify on which parts of existing software systems microservices can be extracted. Challenges related to identification of microservices from variety software artifacts are identified, systematized, and discussed.**

## I. INTRODUCTION

Many software systems are in use today. Some of those systems that are used for long time are named as a legacy software [1]. Legacy software systems are mostly built with monolithic architecture. These monolithic software systems are easier for development, but their modifications are quite hard to implement, leading to very high complexity and high costs of maintenance activities. In many cases, modifications of such systems introduce new errors that may cause problems in their use and future maintenance. The main disadvantage of these systems is their inability to adapt to new requirements or new technologies [2].

Upgrades of legacy systems with new features become difficult over time due to its rigor structure. Modifications of old software systems can be performed in one of the following ways [3]: (1) *removing the system* when its functionalities should be completely changed, (2) *continuing the system use* when small-scale changes are required, (3) *system reengineering for achieving better maintenance* when system quality and performances decrease, and (4) *replacing parts or the whole system* when it cannot support required functionalities.

Reengineering of the system is performed when the old system cannot fulfil changed requirements. In the process of reengineering, the system gets a new form, but keeps the old functionality [4]. One of the most popular evolution type of software system nowadays is reengineering existing software towards microservices architecture. One way to do this is to identify the microservice from the existing software components – software artifacts.

The objective of this paper is a review of challenges in identifying microservices from software artifacts in existing software systems. Challenges are extracted through analyses of scientific literature. The next section present review of literature review studies related to adoption and migration to microservice architectures. The

method and the results of the research are explained in the third and fourth sections respectively. The last section contains conclusions and future research directions.

In the continuation of this section, explanations of the terms *Software artifact* and *Migration to microservices* are given.

### A. Software artifacts

Artifacts are the most important parts of the software systems because they represent the foundation for building, operating, and maintaining software systems. The software elements can have multiple artifacts [5]. The main goal of the artifacts is to define program parts and determine their functionality. They are used in the process of development and maintenance of software components [6]. Efficient software development and maintenance require organization of software artifacts in a meaningful manner, which means that artifacts must contain their own rules for creating, controlling, and interacting with other artifacts. All artifacts should be made to be upgradeable and free for access [7].

There are a lot of artifacts in software life cycle, which can be classified into three main categories [5]:

- **Code Artifacts** (Test Suites, Setup Scripts, Compiled Code, Test Outputs)
- **Documentation** (Diagrams, End-User Agreements, Internal Documents)
- **Project Management Artifacts** (User Cases/Stories, Criteria)

### B. Migration to microservices

Since the monolith software applications are not the best choice, their replacement should be considered and implemented. The best alternative is based cloud architecture [8]. Cloud computing use the virtual servers to store computing resources. Microservices are part of cloud architecture which decomposes entire application into multiple very small and fine-grained services. The advantage of services is that each service is separated and independent from the others. Each microservice is independent relative to development, deployment, and implementation process [9]. In addition, microservices specification should be independent from the selection of programming language, operating system, and hardware platform. Microservice independence means that the changes and errors in one service will not affect other services and the entire system, which makes system maintenance easier [10].

The modernization and migration of legacy system is structured as a process with multiple steps. The process is very complex and depends on decomposition goals, types of relationships, system structures, etc. Microservice identification process can be done manually, or automatically. In the literature can be found several proposed migration processes [11].

## II. LITERATURE REWIEWS ON MICROSERVICES

Migration of existing software systems to microservices architectures has gained significant attention by research community and practitioners from industry, which resulted with the increasing number of research publications. This has led to a great number of scientific publications that should be summarized and synthesized to provide a basis for further research and implementation in the practice. A common approach for literature reviews in software engineering are: Systematic Literature Review (SLR) [12][13], Systematic Mapping Study (SMS) [14][15], and informal literature review [16]. This section provides a short overview of literature review studies related to adoption and migration to microservice architectures.

Several literature reviews have been published recently, dealing with different aspects of microservice architectures and systems. Kalske et al. [17] presented an informal literature survey of challenges in architectural transitions from monolithic architecture to microservices and identified organizational and architectural challenges as two groups of challenges. Taibi et al. [18] presented a SMS on architectural patterns of microservices. Hassan et al. [19] presented a SMS on various views, approaches, and activities for transition to microservices. Aksakalli et al. [9] presented a SLR on deployment and communication patterns in microservice architectures. Velepucha and Flores [2] presented a SMS on migration problems and challenges from monolithic architecture to microservices. Li et al. [20] presented a SLR on quality attributes of microservices architectures. Neves [21] conducted a SMS on technical challenges of microservices migration within his master thesis. Ghani et al. [22] presented a SLR on microservices testing approaches. Razzaq [23] presented a SLR on adoption of microservices architectures for IoT systems.

Insight into published literature reviews revealed that there is a great interest for studying issues related to microservice architectures, but also that there are a lot of spaces for additional research. This paper aims to contribute to the field of microservice literature reviews

by providing review of challenges in identifying microservices from different software artifacts.

## III. REVIEW METHOD

The literature review method consists of several phases. The first phase relates to selecting keywords and forming strings for literature searching on digital libraries with scientific publications. Only papers related to the research topic are included in the analysis, and these studies are called primary studies (ps).

In this research, the following keywords were included: reengineering, transition, migration, identification, decomposition, candidate, and microservices. From combination of two keywords the following search strings were formed:

“reengineering” AND “microservices”  
“transition” AND “microservices”  
“migration” AND “microservices”  
“identification” AND “microservices”  
“decomposing” AND “microservices”  
“candidate” AND “microservices”

A literature search was performed on Google Scholar and the following digital libraries: ScienceDirect, IEEE Xplore, and Springer.

## IV. RESULTS

Using the specified search method, 110 research papers were found for the proposed time range from 2016 to 2021. Based on text analysis of the collected research papers, it was concluded that 16 of them do not correspond to the topic. Other 94 papers are closely related to the research topic **Identification of Microservices from existing software** and were included in more detailed analysis.

The challenges were identified through analysis of 94 selected papers.. All challenges are collected and grouped based on the specified subject. These groups are named to reflect challenge in migrating to microservice architecture. One of the created groups is **Microservice identification from artifacts**, with 10 papers relating to this topic. These 10 papers were selected as primary studies (ps) and presented in Table 1. From the primary studies 10 challenges were identified (CH) and presented in Table 2.

Challenges were formed based on the type of artifacts. A presentation of the identified challenges was also made in the form of a diagram presented at Figure 1.

TABLE I. PRIMARY STUDIES

No.	Reference
ps1	G. Kecskemeti, A. C. Marosi and A. Kertesz (2016) " <i>The ENTICE approach to decompose monolithic services into microservices</i> ", 2016 International Conference on High Performance Computing & Simulation (HPCS), 2016, pp. 591-596, doi: <a href="https://doi.org/10.1109/HPCSim.2016.7568389">https://doi.org/10.1109/HPCSim.2016.7568389</a> .
ps2	R. Chen, S. Li and Z. Li (2017) " <i>From Monolith to Microservices: A Dataflow-Driven Approach</i> ", 2017 24th Asia-Pacific Software Engineering Conference (APSEC), 2017, pp. 466-475, doi: <a href="https://doi.org/10.1109/APSEC.2017.53">https://doi.org/10.1109/APSEC.2017.53</a> .
ps3	Baresi L., Garriga M., De Renzis A. (2017) " <i>Microservices Identification Through Interface Analysis</i> ", In: De Paoli F., Schulte S., Broch Johnsen E. (eds) Service-Oriented and Cloud Computing. ESOC 2017. Lecture Notes in Computer Science, vol 10465. Springer, Cham. doi: <a href="https://doi.org/10.1007/978-3-319-67262-5_2">https://doi.org/10.1007/978-3-319-67262-5_2</a>
ps4	Safa Habibullah, Xiaodong Liu, and Zhiyuan Tan (2018) " <i>An Approach to Evolving Legacy Enterprise System to Microservice-Based Architecture through Feature-Driven Evolution Rules</i> ", International Journal of Computer Theory and Engineering vol. 10, no. 5, pp. 164-169, 2018, doi: <a href="https://doi.org/10.7763/IJCTE.2018.V10.1219">https://doi.org/10.7763/IJCTE.2018.V10.1219</a>
ps5	Sinan Eski, Feza Buzluca, (2018) " <i>An automatic extraction approach: transition to microservices architecture from monolithic application</i> ", XP '18: Proceedings of the 19th International Conference on Agile Software Development: Companion, May 2018, Article No.: 25, Pages 1–6, doi: <a href="https://doi.org/10.1145/3234152.3234195">https://doi.org/10.1145/3234152.3234195</a>
ps6	De Alwis A.A.C., Barros A., Polyvyanyy A., Fidge C. (2018) " <i>Function-Splitting Heuristics for Discovery of Microservices in Enterprise Systems</i> ", In: Pahl C., Vukovic M., Yin J., Yu Q. (eds) Service-Oriented Computing. ICSSOC 2018. Lecture Notes in Computer Science, vol 11236. Springer, Cham. <a href="https://doi.org/10.1007/978-3-030-03596-9_3">https://doi.org/10.1007/978-3-030-03596-9_3</a>
ps7	Garriga M. (2018) " <i>Towards a Taxonomy of Microservices Architectures</i> ", In: Cerone A., Roveri M. (eds) Software Engineering and Formal Methods. SEFM 2017. Lecture Notes in Computer Science, vol 10729. Springer, Cham. doi: <a href="https://doi.org/10.1007/978-3-319-74781-1_15">https://doi.org/10.1007/978-3-319-74781-1_15</a>
ps8	Shanshan Li, He Zhang, Zijia Jia, Zheng Li, Cheng Zhang, Jiaqi Li, Qiuya Gao, Jidong Ge, Zhihao Shan, A (2019) " <i>Dataflow-driven approach to identifying microservices from monolithic applications</i> ", Journal of Systems and Software, Volume 157, 2019, 110380, ISSN 0164-1212, doi: <a href="https://doi.org/10.1016/j.jss.2019.07.008">https://doi.org/10.1016/j.jss.2019.07.008</a>
ps9	Saidani I., Ouni A., Mkaouer M.W., Saied A. (2019) " <i>Towards Automated Microservices Extraction Using Multi-objective Evolutionary Search</i> ", In: Yangui S., Bouassida Rodriguez I., Drira K., Tari Z. (eds) Service-Oriented Computing. ICSSOC 2019. Lecture Notes in Computer Science, vol 11895. Springer, Cham. doi: <a href="https://doi.org/10.1007/978-3-030-33702-5_5">https://doi.org/10.1007/978-3-030-33702-5_5</a>
ps10	M. H. Gomes Barbosa and P. H. M. Maia (2020) " <i>Towards Identifying Microservice Candidates from Business Rules Implemented in Stored Procedures</i> ", 2020 IEEE International Conference on Software Architecture Companion (ICSA-C), 2020, pp. 41-48, doi: <a href="https://doi.org/10.1109/ICSA-C50368.2020.00015">https://doi.org/10.1109/ICSA-C50368.2020.00015</a> .

TABLE II. CHALLENGES IN IDENTIFYING FROM SOFTWARE ARTIFACTS

CH	Challenge	Artifact	Explanation
CH1	Microservice delivery based on building with common parts. [ps1]	common parts	Refers to the creation of patterns based on common parts from already defined microservice images.
CH2	Generalization monolithic services fragmentation into microservices. [ps1]	services	Prohibit decomposition unless it introduces alternative protocols to communicate between fragmentations.
CH3	Improvement of the initially generated microservice candidates by including post-decomposition criteria. [ps2]	post-decomposition criteria	After the decomposition process, extracted candidate for microservices need expert judgement based on pre-established criteria.
CH4	Microservice decomposition with non-functional aspects. [ps3, ps9]	non-functional criteria	Extending approach by adding response time, resource allocation, cost, scalability and availability of the system.
CH5	Transformation of large enterprise system by applying enhanced evolution rules. [ps4]	evolution rules	Settings of new rules for migration process.
CH6	Extraction microservices with graphs. [ps5]	graphs	Using graph representation of the system, migration process and microservices extraction could be more accurate.
CH7	Microservice discovery process using database artifacts. [ps6] [ps10]	database	Database artifacts are collection of tables (Business Objects), triggers, functions, views...
CH8	Latent use of serverless architectures to deploy and manage microservices. [ps7]	serverless architecture	"Serverless functions are single-purpose, programmatic functions that are hosted on managed infrastructure." [26]
CH9	Searching for factors to determine a process of migration. [ps8]	different factors	Detection of factors which can improve the migration process. Examples of factors are: frequency of communication among processes, the execution time of processes, runtime data monitoring and collection...
CH10	Expansion of the microservice identification process using by code artifacts [ps10]	code	Example of code artifacts are listed in introduction section.

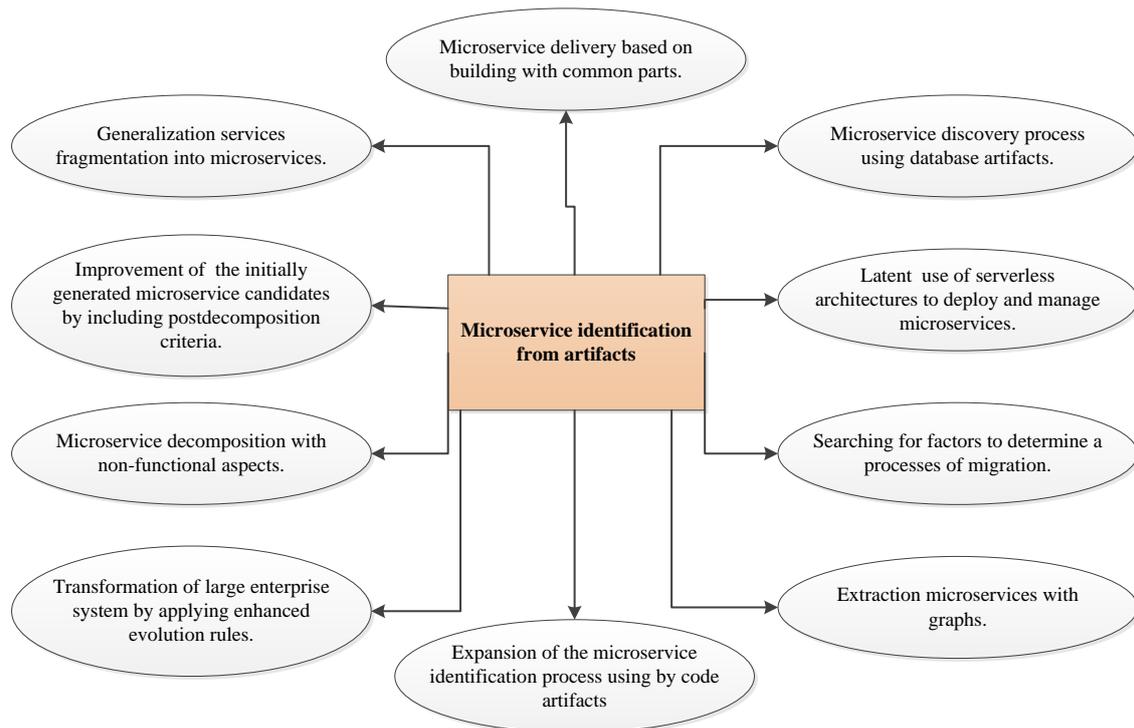


Figure 1. Identified challenges

## V. CONCLUSION

The theme of this paper are challenges in identifying software artifacts for migration of existing software systems to microservices. This is important subject because there are many software systems in use that are developed in monolithic architecture. This architecture is difficult and expensive to maintain. One way to overcome the imperfections of these system is to migrate them to service oriented architectures. Microservices architecture provides easy maintenance and upgrades that extend the life of applications and allows to follow the trends of modern technologies. The main contribution of this work is summary view and analysis of challenges in identifying microservices from software artifacts. This analysis can help other researchers that plan to research migration to microservices, and engineers that can find references on dealing with specific types of artifacts in migration process. The limitation of this work is the small number of references and possibility that some specific software artifacts are missed in the analysis. Since microservices appear in the last ten years, inclusion of papers published before 2016 can help to get insight into their development and early implementations.

For future work is planned to increase the number of references and the inclusion of older works to get more comprehensive insight into this topic. Detailed review of

other migration challenges is also future research direction. Examination of the area of software artifacts in more details is future research direction that is important for proposing a method and a tool for migrating specific monolithic software systems to microservices architecture.

## REFERENCES

- [1] D. Wolfart, W. K. G. Assunção, Ivonei F. da Silva, D. C. P. Domingos, E. Schmeing, G. L. Donin Villaca, and Diogo do N. Paza. (2021), "Modernizing Legacy Systems with Microservices: A Roadmap". In *Evaluation and Assessment in Software Engineering (EASE 2021)*, June 21–23, 2021, Trondheim, Norway. ACM, New York, NY, USA, 11 pages, doi: 10.1145/3463274.3463334
- [2] V. Velepucha and P. Flores (2021), "Monoliths to microservices - Migration Problems and Challenges: A SMS," 2021 Second International Conference on Information Systems and Software Technologies (ICI2ST), 2021, pp. 135-142, doi: 10.1109/ICI2ST51859.2021.00027.
- [3] I. Sommerville, (2010), "Software Engineering". Harlow, England: Addison-Wesley. ISBN: 978-0-13-703515-1
- [4] P. Bourque, R. E. Fairley (2014). "SWEBOK: Guide to the Software Engineering Body of Knowledge". Los Alamitos, CA: IEEE Computer Society. ISBN: 978-0-7695-5166-1
- [5] <https://artifacts.ai/what-is-an-artifact/>
- [6] N. Nazar, Y. Hu, H. Jiang (2016) "Summarizing software artifacts: A literature review". *Journal of computer science and technology* 31(5): 883–909 Sept. 2016, doi: 10.1007/s11390-016-1671-1
- [7] M. Aksit, F. Marcelloni, B. Tekinerdogan, K. van den Berg, P. van den Broek (1998) "Active Software Artifacts". *Lecture Notes in Computer Science*, 307–310, doi: 10.1007/3-540-69687-3\_62
- [8] O. Al-Debagy, P. Martinek (2021) "Dependencies-based microservices decomposition method". *International Journal of Computers and Applications*, doi: 10.1080/1206212X.2021.1915444

- [9] I. Karabey Aksakalli, T. Çelik, A. Burak Can, B. Tekinerdoğan (2021) "Deployment and communication patterns in microservice architectures: A systematic literature review". *Journal of Systems and Software*, Volume 180, 2021, 111014, ISSN 0164-1212, doi: 10.1016/j.jss.2021.111014.
- [10] F. Auer, V. Lenarduzzi, M. Felderer, D. Taibi (2021) "From monolithic systems to Microservices: An assessment framework". *Information and Software Technology*, Volume 137, 2021, 106600, ISSN 0950-5849, doi: 10.1016/j.infsof.2021.106600.
- [11] L. Kirby, E. Boerstra, Z. Anderson and J. Rubin (2021) "Weighing the Evidence: On Relationship Types in Microservice Extraction," in 2021 IEEE/ACM 29th International Conference on Program Comprehension (ICPC), 2021 pp. 358-368. doi: 10.1109/ICPC52881.2021.00041
- [12] B.A. Kitchenham, (2004) "Procedures for Undertaking Systematic Reviews, Joint Technical Report", Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd. ( 0400011T.1). Keele, UK, 2004.
- [13] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, S. Linkman (2008) "Systematic literature reviews in software engineering - A systematic literature review". *Information and Software Technology*, vol. 51, no. 1, pp. 7-15. doi: 10.1016/j.infsof.2008.09.009.
- [14] K. Petersen, R. Feldt, S. Mujtaba, M. Mattsson (2008) "Systematic mapping studies in software engineering". In *Proceedings of the 12th international conference on Evaluation and Assessment in Software Engineering (EASE'08)*, pp. 68-77. Swinton, UK2008.
- [15] K. Petersen, S. Vakkalanka, L. Kuzniarz (2015) "Guidelines for conducting systematic mapping studies in software engineering: An update". *Information and Software Technology*, vol. 64, pp. 1-18. 2015. doi: 10.1016/j.infsof.2015.03.007.
- [16] M. Niazi (2015) "Do Systematic Literature Reviews Outperform Informal Literature Reviews in the Software Engineering Domain? An Initial Case Study". *Arabian Journal for Science & Engineering*, vol. 40, no. 3, pp. 845-855. 2015. doi: 10.1007/s13369-015-1586-0.
- [17] M. Kalske, N. Mäkitalo, T. Mikkonen (2017) "Challenges When Moving from Monolith to Microservice Architecture". In: Garrigós I, Wimmer M. (eds) *Current Trends in Web Engineering (ICWE 2017)*, pp. 32-47. *Lecture Notes in Computer Science*, vol 10544. Springer, Cham. doi: 10.1007/978-3-319-74433-9\_3.
- [18] D. Taibi, V. Lenarduzzi, C. Pahl (2018) "Architectural Patterns for Microservices: A Systematic Mapping Study". In *Proceedings of the 8th International Conference on Cloud Computing and Services Science - CLOSER*, pp. 221-232. 2018. doi: 10.5220/0006798302210232.
- [19] S. Hassan, R. Bahsoon, R. Kazman (2020) "Microservice transition and its granularity problem: A systematic mapping study". *Software: Practice and Experience*, vol. 50, no 9, pp. 1651 - 1681. 2020. doi: 10.1002/spe.2869.
- [20] S. Li, H. Zhang, Z. Jia, C. Zhong, C. Zhang, Z. Shan, J. Shen, M. A. Babar (2021) "Understanding and addressing quality attributes of microservices architecture: A Systematic literature review". *Information and Software Technology*, vol. 131, pp. 106449, 2021. doi: 10.1016/j.infsof.2020.106449.
- [21] J. C. R. D. Neves (2019) "Technical Challenges of Microservices Migration". Master Thesis. ISEP - Instituto Superior de Engenharia do Porto. Politécnico do Porto. Porto, 2019.
- [22] I. Ghani, W. M. Wan-Kadir, A. Mustafa, M. Imran Babar (2019) "Microservice Testing Approaches: A Systematic Literature Review". *International Journal of Integrated Engineering*, vol. 11, no. 8, pp. 65-80. 2019.
- [23] A. Razzaq (2020) "A Systematic Review on Software Architectures for IoT Systems and Future Direction to the Adoption of Microservices Architecture". *SN Computer Science*, vol. 1, 350. 2020. doi:10.1007/s42979-020-00359-w.
- [24] <https://www.pubnub.com/blog/what-is-a-serverless-function/>

# Green Cloud Computing in the Purpose of Energy Efficiency

V. Mikić\*, M. Ilić\*, A. Zakić\* and D. Zlatković\*\*

\* Alfa University/Faculty of Information Technology, Belgrade, Serbia

\*\* Alfa University/Faculty of Mathematics and Computer science, Belgrade, Serbia

vladimir.mikic@alfa.edu.rs, milos.ilic@alfa.edu.rs, aleksandar.zakic@alfa.edu.rs, dragan.zlatkovic@alfa.edu.rs

**Abstract** - Cloud computing represents the delivery of computing services via the internet. The number of people and companies using some of the cloud services is rapidly growing. The increasing demand for cloud infrastructure has resulted in a significant increase in data center energy consumption. Addressing this critical issue is necessary in order to reduce energy consumption and thus reduce carbon dioxide emissions. Green cloud computing is recognized as a possible solution to this problem, as its primary goal is to maximize energy efficiency. In this paper, in addition to defining cloud computing with the main cloud service models, deployment models and computing platforms, a presentation of green cloud computing solutions for energy savings is given. These solutions were then classified, and their results in reducing energy consumption were compared and discussed.

## I. INTRODUCTION

The term cloud computing already represents one of the key phrases in the IT sector. It encourages users to upload their data to the remote cloud, and then access it in the most simple way. Cloud technology has a wide range of applications, offering scalability, reliability and security, and good performance at comparatively affordable rates [1]. The evolution and development of these technologies has modified modern networking, and provided promising environmental protection perspectives, and also economic advantages. These technologies hence tend to improve energy efficiency and system reliability. Minimizing energy consumption is an important element in this approach. In addition, by enhancing resource usage, energy consumption could be considerably reduced. The use of energy in computer systems such as clouds also leads to numerous other important problems such as carbon emissions and system stability. The current support for green or sustainable computing has received considerable attention in recent years. Green cloud computing cannot be constrained to main computing components, but can develop into a broader spectrum of computing resources including equipment and the physical area these resources require. The use of energy and resources in clouds is highly interconnected. In particular, when used in adequate measure, resources with a low utilization rate still consume inappropriate amounts of energy in comparison to their energy consumption when they are optimized.

In this paper an overview of green cloud energy-saving solutions and their energy efficiency is presented and discussed.

## II. CLOUD COMPUTING

Cloud computing represents the use of computer system resources on-demand, for the purpose of storing data, developing, networking, etc [2]. Cloud computing infrastructure refers both to applications that are provided as Internet services and to hardware and system software in data centers. All the information that users need to access is located in the cloud or virtual space. Users do not have to be in a specific place to access it, because they can access all the data via the Internet. There are many benefits of cloud computing, such as cost savings, flexibility, scalability, speed, efficiency, and data security.

### A. Cloud service models

Services proposed by cloud providers can be mainly categorized into software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS).

SaaS presents a distribution mechanism for software where applications are hosted and made available over the network to clients by a service provider [3]. SaaS applications must also communicate with other data and other applications in a broad range of contexts and platforms. Examples of implementation of this service such as Salesforce.com, Google apps, Amazon apps, have developed their SaaS architectures to host exterior applications as well as their own online services.

PaaS provides an environment where developers can develop and design apps and do not need to worry about how much memory and other resources their work will demand [4]. The PaaS approach offers benefits to developers for designing applications and their maintenance in terms of completing the whole software life cycle. It also offers a full range of technology to construct and operate SaaS applications. PaaS's The distinction between PaaS and SaaS lies in the fact that SaaS only represents completed cloud applications, whereas PaaS provides a cloud application development platform.

IaaS represents a combination of hosting, hardware and basic cloud services [5]. The basic purpose of the IaaS paradigm is to supply high-performance and user-friendly computing power. IaaS design seeks to increase

the hardware utilization of the entire platform, optimize the deployment of virtual machines and raise demand for service.

### B. Cloud deployment models

According to [6] four main cloud deployment models are categorized: private cloud, public cloud, community cloud, and hybrid cloud.

The private cloud is also called the internal cloud, and is administered and managed by a single company or group [7]. A consistent degree of security, privacy and governance control is the benefit of an infrastructure that is managed and controlled by just one organization.

Public cloud, also known as external cloud represents an open cloud environment used by general public [8]. It may be operated and managed by the government, or an organization or some sort of combination between them.

Community cloud implies a specially designed cloud environment that is shared and operated in a common domain by a number of associated organizations [9]. This approach shares resources in a community that has shared interests with various groups.

Hybrid cloud consists of two or more independent cloud infrastructures but is connected through conventional data portability technology [10]. It offers benefits for numerous deployment models and allows the companies to manage the continuous workload on private clouds.

### C. Cloud computing platforms

The most known cloud computing platforms are: Amazon Web Service, Google Cloud, Microsoft Azure and IBM Cloud.

Amazon Web Service (AWS), a daughter company of Amazon, is a leading cloud provider in the world. It offers hundreds of different services, including IaaS and PaaS solutions. It provides services and products in all major spheres, including machine learning, databases, computing, IoT, migration and transfer, data storage, etc. Facebook, Twitter, Adobe and Netflix are just some of the world-renowned companies that use AWS services.

Google Cloud, offered by Google, provides solutions for data storage, data analytics, machine learning, networking and many more. Like AWS, Google Cloud also offers a wide set of IaaS and PaaS services. When it comes to pricing, Google claims to be the most cost-effective compared to other competitors.

Microsoft Azure, or simply Azure, was created by Microsoft Corporation. Azure offers a wide array of services within various categories, such as analytics, AI/machine learning, databases, data storage, blockchain, etc. Unlike the previous two cloud services, in addition to IaaS and PaaS services, it also offers SaaS.

IBM Cloud, developed by IBM, is a cloud platform for business. It includes SaaS, IaaS and PaaS offered through public, private and hybrid cloud deployment models. IBM Cloud offers a large number of products and services that can be used by programmers to develop,

manage and run all kinds of applications. This platform covers AI/machine learning, analytics, automation, blockchain, IoT, containers, data storage, etc.

## III. GREEN CLOUD COMPUTING

Green cloud computing, or simply Green cloud, represents the sustainable way to use the cloud [11]. With the increasing growth of cloud computing usage, energy consumption is also increasing, and because of that, green cloud computing has become an interesting topic for researchers, especially in finding energy-efficient solutions. Green cloud computing entails designing, manufacturing, and utilizing digital spaces in a way that minimizes their environmental impact [12]. Green cloud computing aims to reduce energy usage as well as the amount of waste disposed to the environment. That is why energy-efficient solutions are needed to limit the environmental impact of cloud computing. The benefits of green cloud computing include energy savings, reduction of carbon dioxide emissions and the use of hazardous materials, less e-waste, recycling of e-waste, decreasing the amount of paper used, etc.

As stated before, one of the challenging research issues in the area of cloud computing is the problem of energy efficiency, and that is why many authors propose new solutions for reducing computational energy consumption. Jeba et al. [13] present an efficient energy optimization framework based on dynamic resource scheduling for virtual machine migration in cloud data centers. Dashti et al. [14] propose dynamic virtual machines placement for energy efficiency by particle swarm optimization in cloud computing. Jena [15] uses a clonal selection algorithm for task scheduling (TSCSA) in order to optimize energy and processing time. Sigwele et al. [16] proposes an energy reduction technique based on baseband workload consolidation using virtualized general-purpose processors (GPPs) in the cloud. Usman et al. [11] propose an energy-efficient flower pollination algorithm (EE-FPA) for optimal resource allocation of data center virtual machines and also for resource under-utilization. Yang et al. [18] propose an energy-efficient cloud system with two novel methods, the dynamic resource allocation method and the energy-saving method. Aslam et al. [19] present a virtual machine consolidation approach based on an artificial neural network to enhance energy efficiency in the green cloud. Wen et al. [20] propose an energy-efficient virtual resource dynamic integration method to minimize the energy consumption of a data center. Gholipour et al. [21] propose a novel energy-aware resource management technique using joint virtual machines and container consolidation approach for green computing in cloud data centers. Xu et al. [22] present a virtual machine scheduling algorithm based on the gravitational effect to optimize the energy efficiency of cloud computing systems. Ajmal et al. [23] propose a green cloud computing algorithm named cost-based energy-efficient scheduling technique for dynamic voltage frequency scaling systems (CEEST) which reduces energy consumption without compromising the quality of service. Hussain et al. [24] propose energy and performance-efficient task scheduling algorithm (EPETS)

in a heterogeneous virtualized cloud to resolve the issue of energy consumption.

#### IV. DISCUSSION AND RESULTS

As previously stated, green cloud computing represents a paradigm of computing models that are utilized to make the resources more energy-efficient in terms of expense and power. Through the usage of resources there is a number of factors that should be addressed. In present day society, a society that remotely controls the data centers and servers under cloud computing models, green cloud computing is necessary to make them more energy-efficient and economical. The main, and most complex aim of green cloud computing is maximizing that energy efficiency. As data increases exponentially, green cloud computing has problems connected to computation infrastructures that not only decrease energy consumption but also increase the performance and economic efficiency of cloud services. After researching numerous applications of green cloud computing, this section provides the main concept categorization which is identified by reviewed publications. Also, the results of their energy efficiency are portrayed.

The trend of research, and a number of publications per year is presented in Figure 1.

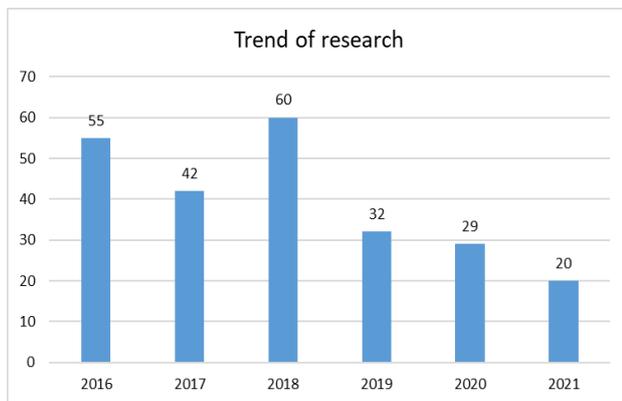


Figure 1. Number of publications per year

The graph is founded on data extracted from Google Scholar database, hence it is the biggest database of scholarly work. The extraction process was conducted through Harzing’s Publish and Perish [25] software. The process of filtering was carried out with the use of parameters (keywords) in the title field, and keyword field. A search query for title field was based on the string “Green cloud computing”, and for keyword field “Energy efficiency”. The publications time scope parameter was set to filter related works from the period of 2016 to 2021. The trend of research shows a slight decrease in a number of publications per year, in a period of the last 3 years which is normal as it relates to various areas of the field being extensively explored in past periods.

Numerous authors published research work on the topic of green cloud computing and various concepts for obtaining better energy efficiency. Table I represents the

conceptual matrix of green cloud computing areas and their interconnection in terms of environment and energy optimization.

TABLE I. CONCEPT CLASSIFICATION OF LISTED SOLUTIONS

Authors	Virtual machine	Network infrastructure	Task scheduling	Resource allocation
Jeba et al. [13]	x			
Dashti et al. [14]	x			
Jena [15]			x	
Sigwele et al. [16]		x		
Usman et al. [17]				x
Yang et al. [18]				x
Aslam et al. [19]	x	x		
Wen et al. [20]	x			
Gholipour et al. [21]	x			
Xu et al. [22]	x			
Ajmal et al. [23]			x	
Hussain et al. [24]			x	

Table II shows the energy saving results, which are extracted from 12 relevant publications included in the green cloud computing section. The results indicate a percentage improvement in energy savings after the application of the proposed solutions.

TABLE II. ENERGY SAVINGS OF LISTED SOLUTIONS

Authors	Energy savings for proposed solutions
Jeba et al. [13]	30%
Dashti et al. [14]	14%
Jena [15]	30%
Sigwele et al. [16]	38%
Usman et al. [17]	35%
Yang et al. [18]	39,89%
Aslam et al. [19]	30%
Wen et al. [20]	45%
Gholipour et al. [21]	9,9%
Xu et al. [22]	20%
Ajmal et al. [23]	30%
Hussain et al. [24]	20%

A comparison of results from Table II is presented in Figure 2. It can be seen from this chart, that the proposed solution in [20] minimizes the energy consumption by 45%, which represents the largest energy savings compared to others. The lowest energy savings in a given sample were in [21] with only 9,9%. On average, proposed solutions achieved energy savings of 28,48%. The standard deviation of 10,63 indicates that the percentages of the proposed solutions do not deviate much from the average energy savings. The coefficient of variation is 37,32%, so it suggests that the arithmetic

mean is a satisfactory measure of the central tendency, respectively, that the data is satisfactorily grouped around the arithmetic mean.

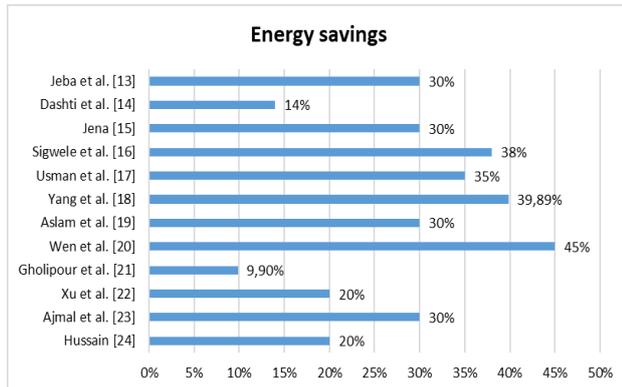


Figure 2. Comparison of energy savings from listed solutions

## V. CONCLUSION

Green cloud computing is recognized to be a wide and promising field for study. To benefit from different assets, it has produced optimal strategies for using resources, improving networking infrastructures, scheduling tasks, and allocating resources, and all in favor of making them more energy efficient. A number of research initiatives and technologies that aim to improve the energy efficiency of different elements of green cloud computing were categorized, and their energy efficiency results were compared. From this review, it is concluded that energy efficiency and power management solutions are important factors in the transition from traditional cloud computing technologies to more green and sustainable ones.

## REFERENCES

- Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., & Ghalsasi, A. (2011). Cloud computing—The business perspective. *Decision support systems*, 51(1), 176-189.
- Wang, L., Von Laszewski, G., Younge, A., He, X., Kunze, M., Tao, J., & Fu, C. (2010). Cloud computing: a perspective study. *New generation computing*, 28(2), 137-146.
- Cusumano, M. (2010). Cloud computing and SaaS as new computing platforms. *Communications of the ACM*, 53(4), 27-29.
- Beimborn, D., Miletzki, T., & Wenzel, S. (2011). Platform as a service (PaaS). *Business & Information Systems Engineering*, 3(6), 381-384.
- Manvi, S. S., & Shyam, G. K. (2014). Resource management for Infrastructure as a Service (IaaS) in cloud computing: A survey. *Journal of network and computer applications*, 41, 424-440.
- Diaby, T., & Rad, B. B. (2017). Cloud computing: a review of the concepts and deployment models. *International Journal of Information Technology and Computer Science*, 9(6), 50-58.
- Basmadjian, R., De Meer, H., Lent, R., & Giuliani, G. (2012). Cloud computing and its interest in saving energy: the use case of a private cloud. *Journal of Cloud Computing: Advances, Systems and Applications*, 1(1), 1-25.
- Qian, L., Luo, Z., Du, Y., & Guo, L. (2009, December). Cloud computing: An overview. In *IEEE International Conference on Cloud Computing* (pp. 626-631). Springer, Berlin, Heidelberg.
- Marinos, A., & Briscoe, G. (2009, December). Community cloud computing. In *IEEE international conference on cloud computing* (pp. 472-484). Springer, Berlin, Heidelberg.
- Weinman, J. (2016). Hybrid cloud economics. *IEEE Cloud Computing*, 3(1), 18-22.
- Radu, L. D. (2017). Green cloud computing: A literature survey. *Symmetry*, 9(12), 295.
- Jing, S. Y., Ali, S., She, K., & Zhong, Y. (2013). State-of-the-art research study for green cloud computing. *The Journal of Supercomputing*, 65(1), 445-468.
- Jeba, J. A., Roy, S., Rashid, M. O., Atik, S. T., & Whaiduzzaman, M. (2021). Towards green cloud computing an algorithmic approach for energy minimization in cloud data centers. In *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing* (pp. 846-872). IGI Global.
- Dashti, S. E., & Rahmani, A. M. (2016). Dynamic VMs placement for energy efficiency by PSO in cloud computing. *Journal of Experimental & Theoretical Artificial Intelligence*, 28(1-2), 97-112.
- Jena, R. K. (2017). Energy efficient task scheduling in cloud environment. *Energy Procedia*, 141, 222-227.
- Sigwele, T., Alam, A. S., Pillai, P., & Hu, Y. F. (2017). Energy-efficient cloud radio access networks by cloud based workload consolidation for 5G. *Journal of Network and Computer Applications*, 78, 1-8.
- Usman, M. J., Ismail, A. S., Gital, A. Y., Aliyu, A., & Abubakar, T. (2018). Energy-Efficient Resource Allocation Technique Using Flower Pollination Algorithm for Cloud Datacenters. In *International Conference of Reliable Information and Communication Technology* (pp. 15-29). Springer, Cham.
- Yang, C. T., Chen, S. T., Liu, J. C., Chan, Y. W., Chen, C. C., & Verma, V. K. (2019). An energy-efficient cloud system with novel dynamic resource allocation methods. *The Journal of Supercomputing*, 75(8), 4408-4429.
- Aslam, A. M., & Kalra, M. (2019). Using Artificial Neural Network for VM Consolidation Approach to Enhance Energy Efficiency in Green Cloud. In *Advances in Data and Information Sciences* (pp. 139-154). Springer, Singapore.
- Wen, Y., Li, Z., Jin, S., Lin, C., & Liu, Z. (2017). Energy-efficient virtual resource dynamic integration method in cloud computing. *IEEE Access*, 5, 12214-12223.
- Gholipour, N., Ariyanan, E., & Buyya, R. (2020). A novel energy-aware resource management technique using joint VM and container consolidation approach for green computing in cloud data centers. *Simulation Modelling Practice and Theory*, 104, 102127.
- Xu, X., Zhang, Q., Maneas, S., Sotiriadis, S., Gavan, C., & Bessis, N. (2019). VMSAGE: A virtual machine scheduling algorithm based on the gravitational effect for green cloud computing. *Simulation Modelling Practice and Theory*, 93, 87-103.
- Ajmal, M. S., Iqbal, Z., Khan, F. Z., Bilal, M., & Mehmood, R. M. (2021). Cost-based Energy Efficient Scheduling Technique for Dynamic Voltage and Frequency Scaling System in cloud computing. *Sustainable Energy Technologies and Assessments*, 101210.
- Hussain, M., Wei, L. F., Lakhani, A., Wali, S., Ali, S., & Hussain, A. (2021). Energy and performance-efficient task scheduling in heterogeneous virtualized cloud computing. *Sustainable Computing: Informatics and Systems*, 30, 100517.
- Harzing, Anne-Wil. (2010) *The publish or perish book*. Melbourne, Australia: Tarma Software Research Pty Limited.

# Designing the prototype of a scalable smart gardening system for testing and evaluation

S. Felbab\*, D. Dobrilović\*, Z. Čović\*\*, J. Simonc\*\*\*

\* University of Novi Sad, Technical Faculty "Mihajlo Pupin, Zrenjanin, Serbia

\*\* Subotica Tech, Subotica, Serbia

\*\*\* University of Szeged, Faculty of Engineering, Szeged, Hungary

srdjan.felbab@tfzr.rs, dalibor.dobrilovic@tfzr.rs

**Abstract** – The wide involvement of Internet of Things (IoT) in variety of modern systems is evident and grows continuously. The smart agriculture as very important segment of our lives is equally affected with these advances in technology. Considering that the smart agriculture is very complex system, the numerous applications can be developed for this improvement. This paper is targeting one small section of agriculture, plant watering systems. Particularly, this paper investigated the design of portable and scalable smart plant watering system that can be applied in a broad range of scenarios, ranging from one plant watering system in the apartment, to the larger area home and larger area gardens with multiple watering targets. In this paper the evolution of the system designed from local centralized system to the IoT layered architecture system is presented. Together with the system evolution and development, the system architecture and the prototype, called the SmartGrowth system, are presented as well.

## I. INTRODUCTION

The rapid development of the technologies strongly reflects to the world changing our ways of living. The application of the technology has or should have the purpose of improving our lives. The concept of appliance of emerging technologies such as Internet of Things (IoT) is deeply involved in variety of segments of everyday life. The approaches in IoT applications and deployments develop persistently. The smart agriculture as vital portion of our lives is not bypassed with this phenomenon [1]. Taking into account that the smart agriculture is very complex and that has variety of subgroups, we should be aware that various applications can be designed within this field. All these applications should have one thing in common, the optimization of food and plant growth.

This paper is focusing on one segment of agriculture, plant watering systems. To be more precise, this paper investigates the approach in design of scalable, portable plant watering system, an adaptable system that can be applied in a wide scope of situations, stretching from single plant watering, to covering the bigger regions with multiple watering locations.

This paper presents the approach in the evolvement of the centralized local system to the IoT layered architecture plant watering system. The IoT layered architecture should provide the portability and scalability of the system. With these features system can be used as the

platform for development, testing and evaluation of smart home plant watering systems and home gardens.

This paper is structured as follows. After the introduction, the smart home and smart gardening systems are briefly explained. Next, the system architecture and its development and evolution are explained, followed with the description of system prototype. At the end the concluding remarks are given.

## II. SMART HOME AND SMART GARDENING

The smart home projects and systems are numerous. Very large portion of these projects is built upon the open-source hardware and software, with extensive usage of Arduino, NodeMCU/ESP8266 microcontroller boards and Raspberry Pi single board computers. One of these projects is described in [2]. The goal of this project is the design of the system of voice and remote control of home appliances via World Wide Web. In the system presented in [3] the authors integrated Olivia into smart door lock for 'The Smart Home Surveillance System' implemented on Raspberry Pi. The system is designed to identify and differentiate between the owner and stranger using face recognition. Olivia can interact with the stranger at the door in case the owner is not present at home and will notify the owner with Email and SMS.

The integration of similar open-source hardware and software platforms is applicable in smart watering and irrigation systems. The paper [4] presents a farm monitoring and automatic irrigation system that has three modules: unified sensor pole (USP) as intelligent IoT based module, irrigation unit (IU), and sensor information unit (SIU). The IU is responsible for parsing and writing incoming data from USP. SIU stores the sensor data by sending data stream to Mosquitto broker and to log files. The data are used by the HTTP server to allow remote sensor data monitoring.

Paper [5] has in focus the implementation of intelligent technologies like the internet of things (IoT), computer vision etc. for facilitating farming activities and providing flexible farm operations. This paper emphasizes farm management and appropriate monitoring of farm parameters as an indispensable for productive farming in smart cities or rural areas. The paper presents an IoT based Smart AgroTech system for urban farming that considers humidity, temperature, and soil moisture as

necessary farming parameters deciding whether the irrigation action should begin or stop.

The goal of the paper [6] is design of watering system for monitoring status of the plant health. This approach in design involves analyzing the moisture content, temperature of surrounding area where plant grows as parameters for watering schedule. Adafruit cloud page is used as an IOT platform and it is further used to analyze the plant health and send an email alerts. The presented Plant watering system that is innovative, time-saving, and user-friendly and it consists of: Arduino UNO board, sensors, soil moisture sensor (YL38), pH sensor, flame sensor, DHT sensor (DHT11) and Wi-Fi module (ESP8266). The system uses the Machine Learning techniques as well.

The paper [7] deals with automation of gardening systems using IoT considering the facts that water resources should be carefully managed leading to imperative that we use water extremely judiciously. The system tracks the environmental conditions such as humidity of the soil, and temperature and runs algorithms which help keep the garden in good condition and at the same time using resources optimally. The Arduino board is used as a hardware, and systems contains module for fertilizer suggestions for different plants etc.

The application of internet of things communication technologies in smart agriculture and challenges is very interesting area to investigate. The [8] gives literature on this topic. The literature search is made in three different databases: ScienceDirect, IEEE Xplore, and Scopus and total of 94 research articles were reviewed among the total of 886 titles found.

The research presented in [9] shows that NodeMCU can be used to monitor the low wetness and warm temperature and to show the results using Blynk App. After detecting, the message is passed between NodeMCU and Blynk App and it automatically starts the motor in home gardening, farm, etc.

The paper [10] investigates the critical parameters for automating sustainable vertical gardening (VG) systems by using the IoT concept in smart cities. This paper analyzed 30 peer-reviewed publications published between 2004 and 2018, and eight international patents on VGs. The research resulted with identification of research challenges and four potential future investigation ideas on how to approach these challenges in principle.

### III. SYSTEM ARCHITECTURE AND ITS DEVELOPMENT

In this section the evolution of this scalable home gardening system is presented. It started with centralized system focused on managing single plant watering system intended for home indoor usage.

The system is centered around NodeMCU and with directly connected DHT11 and soil moisture sensors, relay and watering pump. The system was wirelessly connected with Mobile phone for displaying data with Blynk application. The system is presented in Fig. 1 with the specification of wiring the components of the system. The system is tested in the scenario of watering one plant.

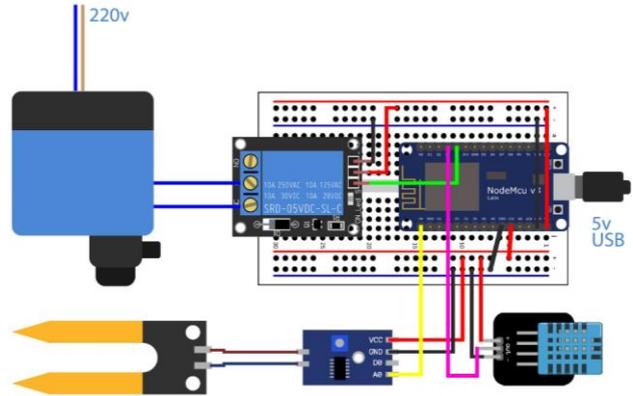


Figure 1. Initial stand-alone prototype system

The model of the same system is presented in Figure 2.

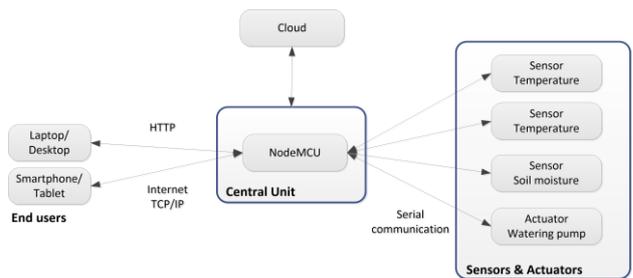


Figure 2. Architecture of centralized local system

With the goal of creating scalable and portable platform testing home gardening/plan watering system, the layered IoT architecture is proposed. This architecture is presented in Fig. 3 and in accordance with the IoT architectures presented in [11]. The system is partially inspired on the project presented in [12].

The architecture scalability requirement was tailored to the need to make this system easily adoptable to the different environments. This architecture should provide the deployment of the system on the small-scale level, such as one watering plant in the small apartment, then to the larger apartments and houses with several watering targets, to the home gardens in outdoor environments.

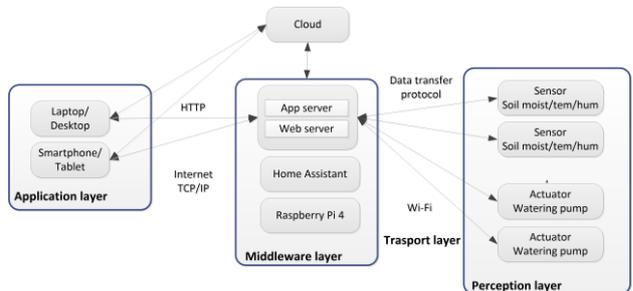


Figure 3. Proposed IoT layered architecture of smart gardening system

The model of the system is used as a starting point for designing the prototype. The prototype, called the SmartGrowth, is explained in the next section. One of the most important components of the prototype are the software, Home Assistant, an open-source software installed on Raspberry Pi. The second important feature of

the prototype is its scalable design allowing this system to be used equally in indoor single plant watering scenarios, and in wider scale outdoor home gardening and small smart farming watering scenarios.

#### IV. SYSTEM PROTOTYPE

The goal of the smart garden system prototype is to optimize water used for watering the plants thus allowing the optimal amounts of water needed for plant growth. It also saves labor time required to manipulate irrigation system turning it on and off. This prototype is a system for

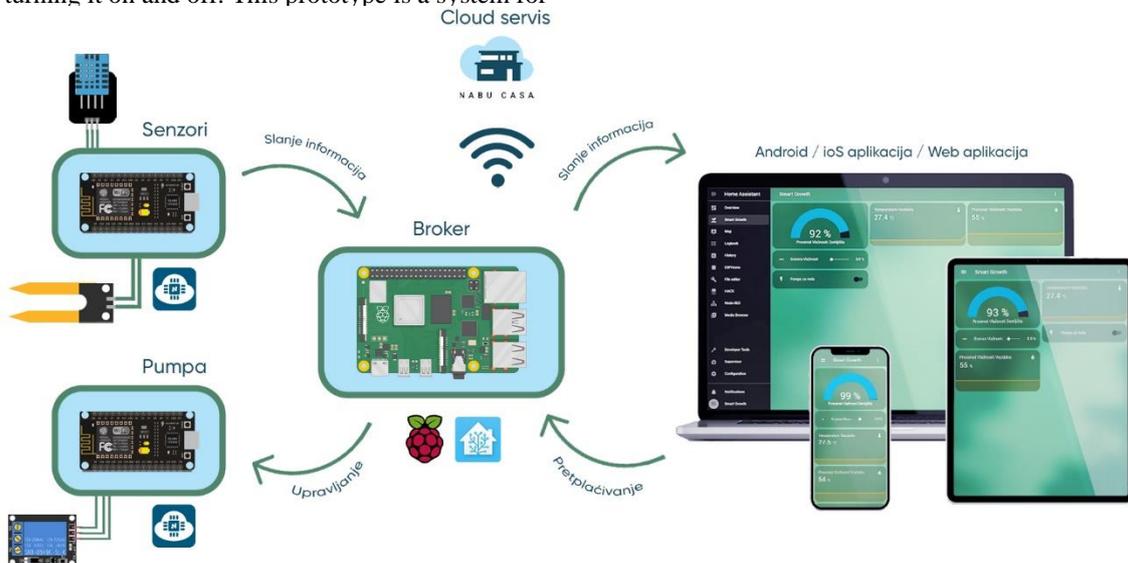


Figure 4. Smart gardening irrigation prototype system SmartGrowth

The development of a prototype smart garden system is based on the usage of open-source hardware and software. The communication between the components is based on Wi-Fi technology, e.g. between devices connected via NodeMCU development boards that use the ESP8266 microchip and the Raspberry Pi 4 as the main computer. The control is done through an open-source system called Home Assistant, which receives feedback from the devices via the TCP/IP protocol. Home Assistant is a system installed on the Raspberry Pi and it plays a major role in controlling and managing the smart garden system. The system defines all the necessary parameters for communication between devices, as well as the executing commands.

The advantage of this approach in building the system is its modularity. Sensors as well as other devices in the system can be added or changed in a very simple way. In addition to its modularity, a great advantage of the system is its cloud-based architecture, enabling control and data monitoring from a remote location.

The data exchange between the data components is achieved with Wi-Fi technology. The central component of the system, Raspberry Pi 4, has integrated Wi-Fi interface, while the sensor nodes are connected to the NodeMCU development boards. Because the NodeMCU board has integrated ESP8266 chip, the transfer of data collected from the sensor, as well as the transfer of command to the actuators is made with Wi-Fi technology and TCP/IP protocol.

digital irrigation, which combines data processing, sensor used for data collection, display of results using smartphone (Android, iOS) and\*or Web applications for monitoring. The system is shown in Fig. 4.

The prototype system measures soil moisture, based on the measuring parameters defined by the system user. When the soil moisture goes below defined levels, the irrigation is turned on, and vice versa. The user adjusts the soil moisture limit via the web-based application for each zone.

The soil moisture sensor has the task of examining the soil moisture and sending information to the NodeMCU development board. The collected information is sent to a central computer which further processes it. Based on the data collected by the sensor and the information set by the user via the application, the system forwards the data further to another NodeMCU development board. Depending on the soil moisture measures and the minimum value set by the user, the second development board switches the water pump on or off. In addition to the soil moisture sensor, a DHT11 soil temperature and humidity sensor is also connected. The DHT11 sensor has the task of testing the temperature and humidity of the air.

##### A. Components of the system

List of components used for system prototype.

- NodeMCU ESP 8266,
- FC-28 soil moisture sensor,
- DHT11 temperature and humidity sensor,
- relay,
- USB Cable,
- prototype board,
- watering pump, and
- Raspberry Pi 4

The Raspberry Pi 4 hosts Home Assistant, the central software component of the system. It is an open-source platform that began its development as a hobby project, dealing with the control of Philips Hue. Currently, it has

community of more than 70,000 IoT fans. This software is used to control smart home devices. Home Assistant supports over 1,000 services and pieces of hardware according to Schoutsen, and this includes Nest, Philips Hue, Sonos, WeMo, Ikea, Arlo, Ecobee, Ring, Dyson, Xiaomi and August devices. Like Samsung SmartThings, it acts as a hub for home devices.

Connecting NodeMCU development boards and communication with the Home Assistant system is done using the ESPHome add-on. This add-on should be additionally installed on the Home Assistant system. After installation, the two NodeMCU boards are configured, one of them has the role of collecting information using a sensor, while the other has the task of turning the relay on and off, i.e., water pumps. For communication between Home Assistant installed and two NodeMCU clients is ESPHome native API. The ESPHome native API is based on a custom TCP protocol using protocol buffers. For the future development, the implementation of more standardized MQTT will be considered.

### B. Device configuration

Home Assistant uses YAML syntax for configuration of devices such as sensor nodes, actuators, etc. YAML is a data serialization language that is often used to create configuration files and works simultaneously with any programming language. The YAML language is designed for human interaction, and it allows creation of complex configurations. It is a strict JSON superset which is another language for data serialization. YAML one of the main differences is that newlines and indents actually have a meaning in YAML, as opposed to JSON which uses parentheses and headers.

The yaml code used for sensor node is shown in Listing 1.

```
esphome:
  name: senzori
  platform: ESP8266
  board: esp01_1m

# Enable logging
logger:

# Enable Home Assistant API
api:

ota:
  password: "OTAPassword"

wifi:
  ssid: "MySSID"
  password: "Mypassword"
  # Enable fallback hotspot (captive portal) in case
  # wifi connection fails
  ap:
    ssid: "Senzori Fallback Hotspot"
    password: "HotspotPassword"

captive_portal:

sensor:
  - platform: dht
    pin: 0
    temperature:
      name: "Temperatura-Vazduha"
    humidity:
      name: "Vlaznost-Vazduha"
    update_interval: 5s

  - platform: adc
    unit_of_measurement: "%"
    icon: "mdi:flower-outline"
    accuracy_decimals: 0
```

```
pin: A0
filters:
  - lambda: |-
      if (x > 0.7) {
        return 0;
      } else if (x < 0.38) {
        return 100;
      } else {
        return (0.7-x) / (0.7-0.38) * 100.0;
      }
name: "Procentat-Vlaznosti-Zemljista"
update_interval: 5s
```

Listing 1. Code for sensor nodes

## V. RESULTS

The system is tested in home environment during the several days with one watering plant in several scenarios (Fig. 5).



Figure 5. The components of the prototype SmartGrowth system in operation

Here is explained the example scenario of prototype operations. The user set a minimum soil moisture percentage of 50%, while the sensor read a soil moisture value of 37% (Fig. 6). As the user setpoint is higher than the value read by the sensor, the water pump will switch on and run until the soil percentage is above 50% (Fig. 7). In this way, where the user sets the minimum value of soil moisture, depending on the plant, we can accurately determine the optimal amount of water it needs. During the scenario the prototype shows its effectiveness, and reliability.

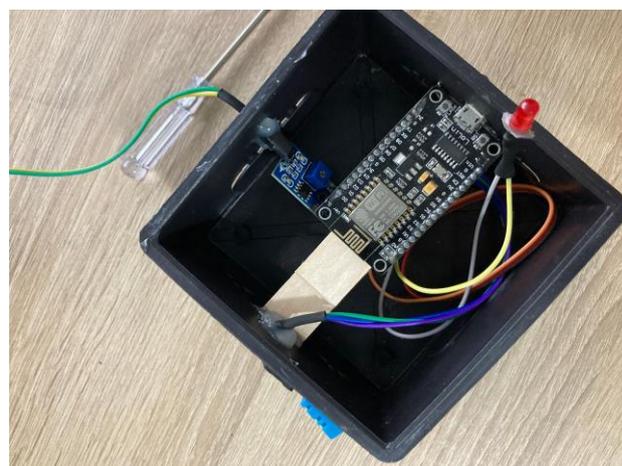


Figure 6. NodeMCU sensor node with temperature and soil moisture sensors

Additionally, the modularity of the system allows easy addition of multiple sensor and watering plant systems, giving the opportunity for easy system expansion. On the other hand, the usage of ESP8266 as portable sensor nodes, with the support of Wi-Fi, gives the opportunity for the system to be deployed in large outdoor the areas with the range offered by the Wi-Fi technology.



Figure 7. The watering pump component of the system



Figure 8. Smart Growth system

## VI. CONCLUSION

The successful testing of the presented prototype justified the choice of Home Assistant as a prototyping platform. The tested prototype will be used as a base for further research. The first phase of further research will include implementation of MQTT protocol in the system. Further, the addition of multiple portable sensor nodes will be tested with the system, together it the analyses of the system behavior with increased numbers of devices. Third, direct of the research will be pointed to the experiments with the range of NodeMCU centered inopen

outdoor and closed indoor areas for assessing the system deployment in gardening or building plant watering.

## REFERENCES

- [1] A. Tzounis, N. Katsoulas, T. Bartzanas, and C. Kittas, "Internet of Things in agriculture, recent advances and future challenges," *Biosystems Engineering*, vol. 164, pp. 31-48, 2017, <https://doi.org/10.1016/j.biosystemseng.2017.09.007>.
- [2] M. Al-Amin and S. Z. Aman, "Design of an Intelligent Home Assistant," 2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), 2016, pp. 69-71, doi: 10.1109/ISMS.2016.34.
- [3] S. Khattar, A. Sachdeva, R. Kumar and R. Gupta, "Smart Home With Virtual Assistant Using Raspberry Pi," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2019, pp. 576-579, doi: 10.1109/CONFLUENCE.2019.8776918.
- [4] N. K. Nawandar, and V. R. Satpute, "IoT based low cost and intelligent module for smart irrigation system," *Computers and Electronics in Agriculture*, vol. 162, pp. 979-990, 2019, <https://doi.org/10.1016/j.compag.2019.05.027>.
- [5] A. K. Podder, A. Al Bukhari, S. Islam, S. Mia, M. A. Mohammed, N. M. Kumar, K. Cengiz, and K. H. Abdulkareem, "IoT based smart agrotech system for verification of Urban farming parameters," *Microprocessors and Microsystems*, vol. 82, 2021, 104025, <https://doi.org/10.1016/j.micpro.2021.104025>.
- [6] A. Reghukumar, and V. Vijayakumar, "Smart Plant Watering System with Cloud Analysis and Plant Health Prediction," *Procedia Computer Science*, vol. 165, pp. 126-135, 2019, <https://doi.org/10.1016/j.procs.2020.01.088>.
- [7] N. Sukhdev, N. Nahata, S. Sridhara and G. Swamy, "IoT Enabled Smart Gardening," In *Proceedings of 14<sup>th</sup> International Conference on Information Processing (ICINPRO)*, pp. 1-2, 2018, doi: 10.1109/ICINPRO43533.2018.9096672.
- [8] W. Tao, L. Zhao, G. Wang, and R. Liang, "Review of the internet of things communication technologies in smart agriculture and challenges," *Computers and Electronics in Agriculture*, vol. 189, 2021, <https://doi.org/10.1016/j.compag.2021.106352>.
- [9] M. Sheth, and P. Rupani, "Smart Gardening Automation using IoT With BLYNK App," In *Proceedings of 3<sup>rd</sup> International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 266-270, 2019, doi: 10.1109/ICOEI.2019.8862591.
- [10] M. N. Halgamuge, A. Bojovschi, P. M.J. Fisher, T. C. Le, S. Adeloju, and S. Murphy, "Internet of Things and autonomous control for vertical cultivation walls towards smart food growing: A review," *Urban Forestry & Urban Greening*, vol. 61, 127094, 2021, <https://doi.org/10.1016/j.ufug.2021.127094>.
- [11] F. Flammini, D. Dobrilović, A. Gaglione, and D. Tokody, "LoRaWAN Technology Mapping to Layered IoT Architecture", In *Proceedings of 10<sup>th</sup> International Conference on Applied Information and Internet Technologies (AIIT2020)*, pp. 131-137, October 16th, Zrenjanin, Serbia, 2020.
- [12] T. Stanciu, R. Bogdan, and M. Marcu, "Developing a low-cost Smart Terrarium in the Context of Home Automation Applications," In *Proceedings of 10<sup>th</sup> International Conference on Applied Information and Internet Technologies (AIIT2020)*, pp. 131-137, October 16th, Zrenjanin, Serbia, 2020.

# A Model for Integration of Internet of Things Systems in a Smart City

H. Dimova Popovska\*, T. Dimovski\*, and I. Hristoski\*\*

\* "St Kliment Ohridski" University – Bitola / Faculty of ICTs – Bitola, N. Macedonia

\*\* "St Kliment Ohridski" University – Bitola / Faculty of Economics – Prilep, N. Macedonia

hristinadimova95@gmail.com, {tome.dimovski, ilija.hristoski}@uklo.edu.mk

**Abstract** - Nowadays, the Internet of Things (IoT) has received a great deal of attention from researchers as it becomes an important technology that promises the life of a smart human being, allowing communication between objects, machines, and all things together with humans. The Internet of Things is a system consisting of real-world objects and sensors attached to or combined with these things, connected to the Internet through a wired and wireless network structure. Things made available on the Internet of Things will share information about the state of things and the surrounding environment with people, software systems, and other machines. With the technology of the Internet of Things, the world will become smarter in all respects, as the Internet of Things provides the means for smart cities, smart healthcare, smart homes, and buildings, in addition to many important applications such as smart power, grid, transport, management and monitoring of waste. In this paper, a model for the integration of IoT systems is presented, including its main components.

## I. INTRODUCTION

The challenges posed by the rapid development of technologies are closely related to information technology. As a new technology, the Internet of Things (IoT) is attracting a lot of attention and expectations to make a major contribution in many areas. The concept of IoT was conceived by Kevin Ashton, a member of the radio frequency identification (RFID) development community (RFID) in 1999 [1], and recently became more relevant to the practical world mostly due to the growth of mobile devices, embedded and ubiquitous communication, cloud computing and data analytics. The world of IoT is populated by billions of facilities that can sense, communicate and share information, all interconnected via public or private Internet Protocol networks. These interconnected objects have data regularly collected, analyzed, and used to initiate action, providing tremendous intelligence for planning, managing, and decision-making. The common definition states that the IoT is a network of physical objects. The Internet is not just a network of computers, but it has evolved into a network of devices of all kinds and sizes: vehicles, smartphones, home appliances, toys, cameras, medical instruments and industrial systems, animals, people, buildings, all interconnected to communicate and share information based on certain protocols to achieve smart reorganizations, positioning, tracing, security and control, and even personal real-time Internet monitoring, Internet upgrades, process control, and administration. With the rapid development of the IoT, large manufacturers are committed to exploring the combination of multiple

sensors, in the sense of obtaining more environmental data with a single device. The deployment of wireless sensor networks in Smart City infrastructures has led to very large amounts of data being generated every day in a variety of domains, with applications including environmental monitoring, health insurance, and transportation monitoring. To take advantage of increased data volumes, new methods and techniques for efficient data management and analysis need to be used to generate information that can help manage resource usage both intelligently and dynamically.

In this paper, we propose a model for the integration of IoT systems in a Smart City. The model is presented as a multi-level architecture. The UML data flow diagram of the IoT systems integration model is also presented, depicting the integration and communication between the individual entities.

The rest of the paper is structured as follows. Section II focuses on some of the most popular data integration tools. Section III discusses the proposed model for integration of the IoT systems in a Smart City as a multi-level architecture. The last section highlights the conclusions.

## II. INTERNET OF THINGS, SMART CITY, AND INTEGRATION TOOLS

As a new technology, the Internet of Things (IoT) is attracting a great deal of attention and expectation to make a major contribution in many areas. The Internet of Things refers to the unique identification of values and their virtual display in the Internet structure. The term Internet of Things was first used by Kevin Ashton in 1999 and has become synonymous with linking things online. RFID [2], sensors, actuators, and cell phones are often seen as prerequisites for the advent of the Internet of Things. The last few years have seen an explosive growth of information and communication technologies to improve the design of hardware and software. The use of ICT in cities in different forms for different urban activities has led to increased efficiency of city operations and these cities are marked by the use of many terms such as "cyber Ville", "digital city", "electronic city", "flexi city", "information city", "telicity", "wired city" and "smart city". An intelligent city is the biggest abstraction in the labels used because it encompasses other labels used for cities [3]. Data integration is the process of combining data from different sources into a single, unified view. The most popular integration tools are: SQL Server Integration Services (SSIS) is a platform that is part of the SQL

Server Management Studio, developed by Microsoft. The platform is used for data integration and workflow applications, for building high-performance data integration solutions. SSIS includes graphical building tools for debugging, tasks for performing workflow functions, such as FTP operations, executing SQL statements, sending emails, source and destination of download and write data, their transformations, calculations, etc. [4]. A widely used data integration software product, Oracle Data Integrator [5] provides a new declarative design approach to defining data transformation and integration processes, resulting in faster and simpler development and maintenance. Based on the unique ELT architecture, Oracle Data Integrator not only guarantees the highest possible level of performance for performing data transformation and validation processes but is also the most efficient solution available today. Oracle Data Integrator provides a unified infrastructure for streamlining data and application integration projects. Data integration is at the heart of the entire Talend Data Fabric platform. Convenient self-service tools make it easy to enter data from almost any source, and the built-in functionality ensures that your data is usable from day one. From fast data downloads to cloud data warehousing to the most sophisticated multi-cloud projects, Talend Data Fabric can meet your needs [6, 7].

### III. A MODEL FOR INTEGRATION OF IOT SYSTEMS IN A SMART CITY

In this section, the proposed model for integration of the IoT systems in a Smart City is presented as a multi-level intelligent city architecture. It contains five levels, as follows (Fig. 1):

**Level 0:** Read data from the integration database. This level reads the IoT systems configuration data stored in the integration database, which is important for the next level.

**Level 1:** Communication and connection parameters required. The communication medium plays an important role in achieving the concept of connecting IoT systems in an intelligent city. The main goal is to establish a connection with sensors from different IoT systems.

**Level 2:** Download data. At this level, the data/information collected by the sensors is stored for further processing. The collected formats are further processed using semantic web technologies to convert them into a common format.

**Level 3:** Data preparation / formatting. The information collected is summarized before transmission, using semantic web technologies analysis and fusion. The main goal at this level is to convert the collected heterogeneous information into a common format. RDF11 is the most common way to exchange information over the Internet and facilitates heterogeneous data sharing and integration across different domains of the Smart City.

**Level 4:** Data storage. At this level, data is stored that is processed and ready to be shared via sensors throughout the architecture of the intelligent city.

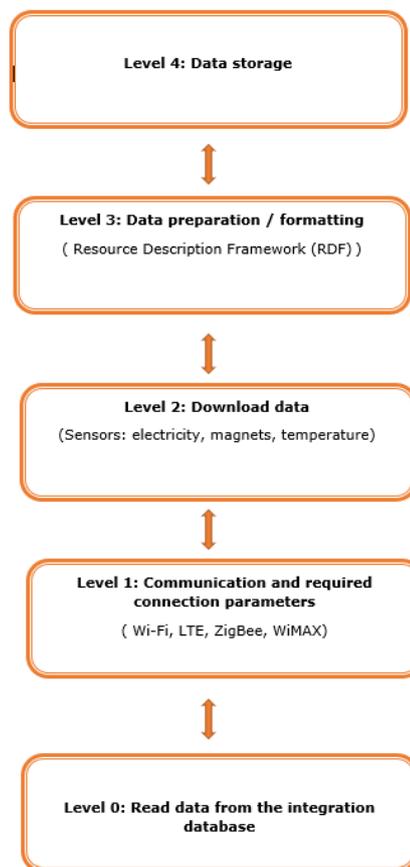


Figure 1. The five levels of the multi-level architecture

The holistic view of the proposed city architecture is depicted in Fig. 2. The approach focuses on city management as a subsystem system. Each autonomous subsystem is connected to the proposed multi-level architecture, which is fully integrated and interconnected with each subsystem. Each system shares its data with the proposed multi-level architecture that can provide cross-domain services to citizens. The proposed multi-level architecture acts as an integration point for information coming from the subsystems. The proposed multi-level architecture can use the information and data available to them to make better decisions in real-time.

The UML data flow diagram of the IoT Systems Integration Model, presented in Fig. 3, shows integration and communication between the individual IoT entities. The model consists of several entities that are interconnected and in sync. The entities, i.e. the separate parts of the model, can be physically located at different locations, but at the same time, they function as if they are a single system or a single whole.

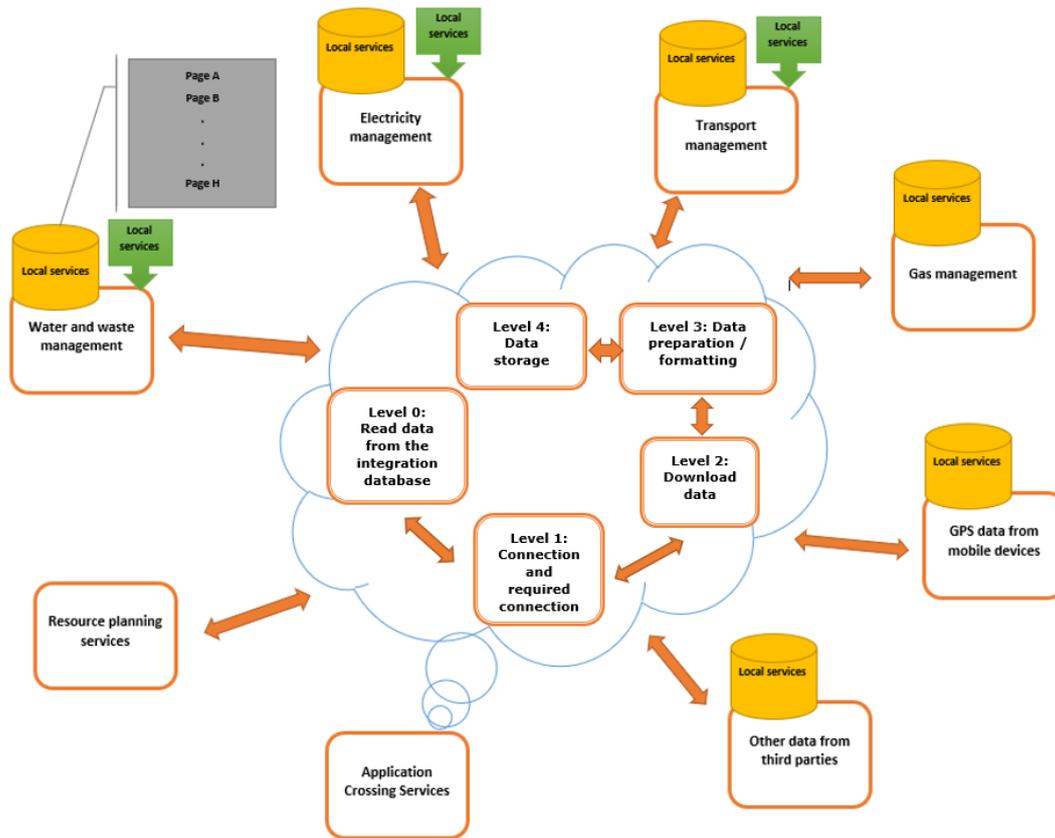


Figure 2. A holistic view of the proposed Smart City architecture

The UML diagram consists of the following entities:

- Data Warehouse
- Integration System
- Integration Database
- IoT System N ((Node N), (Sensor N), (Database for system N)).

The integration system is defined as the process of bringing entities together into a single system and ensuring that entities function together as a system. Integration system integrates IoT systems that are part of Smart City. Initially, the integration system addresses the integration database from which it reads parameters and information about IoT systems and their nodes to which they need to connect and download data, as shown in the diagram. The integration system can receive information from the nodes of IoT systems in two ways:

- Option 1 (Data pulling): The integration system sends a request to Node N for data.
- Option 2 (Push system): Node N sends data whenever there is something new.

An integration database is a database that contains data for IoT systems used in Smart City and in which data should be integrated and stored in a common data warehouse. IoT systems are configured in the database nodes i.e. the sensor nodes that make up those systems

and the information for connecting and accessing sensor node data.

The UML diagram presented in Fig. 3 shows an example of System N, which is just one IoT system of all the systems that can be part of Smart City. System N consists of Node N, Sensor N, and Database for System N. Node N sends data requests from Sensor N. Sensor N returns data, while Node N writes the data to its database.

A data warehouse is a database management system designed to enable and support business intelligence and analytics. Data warehouses are intended primarily for research and analysis, and they often contain large amounts of historical data. The data in the data warehouse is usually obtained from a wide variety of sources, such as application logs and transaction-based applications. The data warehouse centralizes and consolidates large amounts of data from multiple sources. Its analytical capabilities enable organizations to extract valuable business insights from their data to improve decision-making. Over time, it builds a historical record that can be invaluable to data scientists and business analysts. Because of these possibilities, the data warehouse can be considered the "only source of truth" of the organization. In the diagram, the data warehouse writes data from System N. When the data is going to be written, the data warehouse returns a response stating that the data is written. The integration system will try to write the data until the data warehouse returns a relevant response.

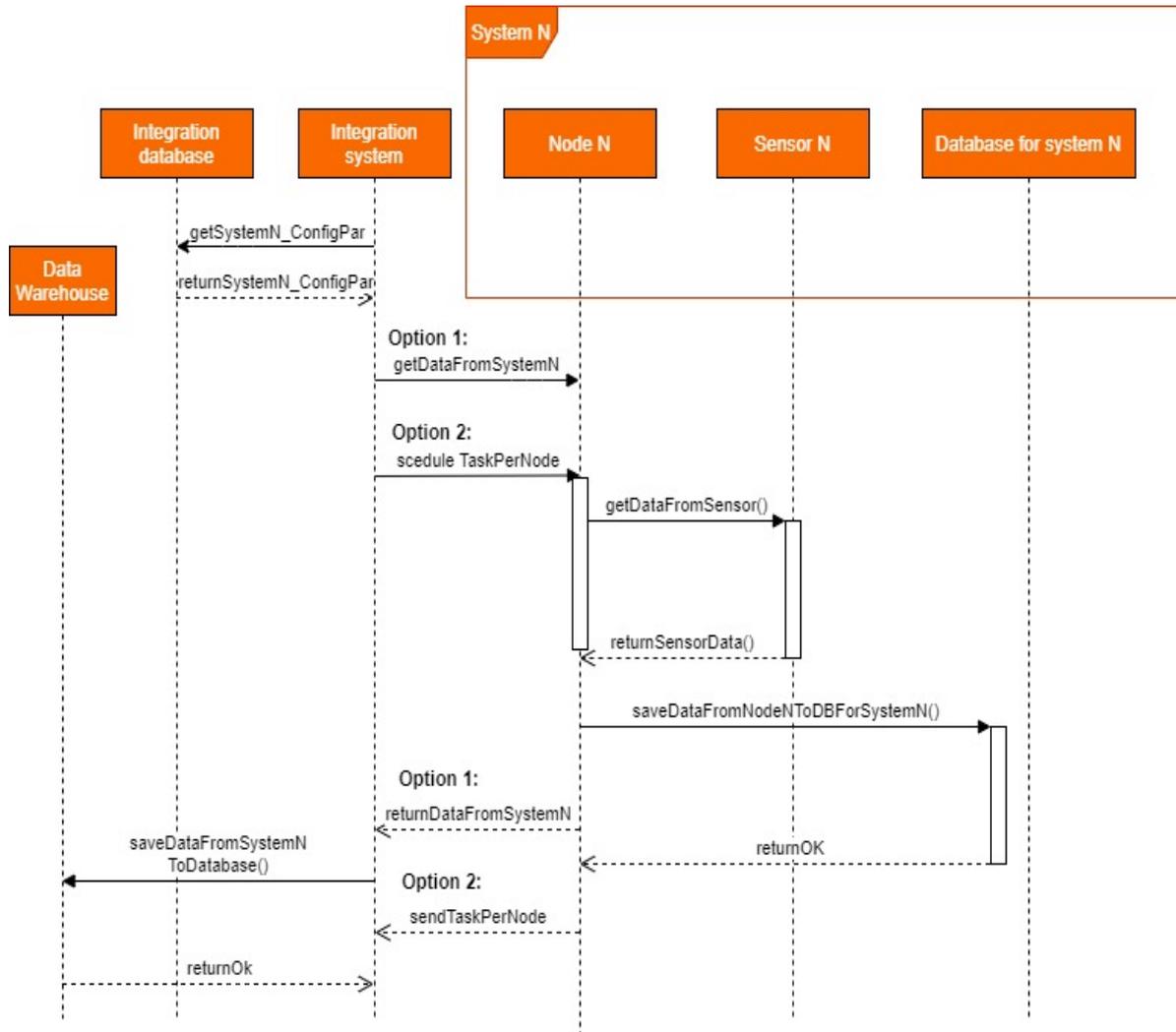


Figure 3. UML diagram of the IoT Systems Integration Model

#### IV. CONCLUSION

In this paper, we propose a model for the integration of IoT systems in a Smart City. The model is shown as a multi-level architecture. The UML data flow diagram of the IoT systems integration model is presented, which shows integration and communication between the individual IoT entities. The model consists of several entities that are interconnected and in sync. The entities, i.e. the separate parts of the model, can be physically located at different locations, but, at the same time, they function as if they are a single system or a single whole.

Our future work is going to be directed towards the detailed design of the integration database and the implementation of the integration system based on the hereby presented model.

#### REFERENCES

[1] W3C Proposed Recommendation, "Resource Description Framework (RDF) Model and Syntax Specification", 1999. [Online]. Available: <http://www.w3.org/TR/PR-rdf-syntax/>

[2] D. Washburn, U. Sindhu, "Helping CIOs Understand Smart City Initiatives", 2010. [Online]. Available: [https://s3-us-west-2.amazonaws.com/itworldcanada/archive/Themes/Hubs/Brainstorm/forrester\\_help\\_cios\\_smart\\_city.pdf](https://s3-us-west-2.amazonaws.com/itworldcanada/archive/Themes/Hubs/Brainstorm/forrester_help_cios_smart_city.pdf)

[3] S. Luis, J. R. Santana, L. Munoz, et al., "SmartSantander: IoT experimentation over a smart city testbed." *Computer Networks*, vol. 61, pp. 217-238, 2013.

[4] K. Anbarasan, "SQL Server Integration Services (SSIS) – Step by step Tutorial", eBook, 2011.

[5] Oracle, "Oracle Data Integrator", [Online]. Available: [https://docs.oracle.com/cd/E29542\\_01/doc.1111/e14773/odi.htm](https://docs.oracle.com/cd/E29542_01/doc.1111/e14773/odi.htm)

[6] Talend Data Fabric, [Online]. Available: <https://www.talend.com/products/data-fabric/>

[7] J. Sreemathy, I. Joseph V., S. Nisha, C. Prabha I., and G. Priya R.M., "Data Integration in ETL Using TALEND," 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 1444-1448, 2020.

# Application of CNNs on mobile devices for emotion recognition based on facial expressions

S. Mihajlović\*, D. Ivetić \*\*, I. Berković\*

\* Technical Faculty “Mihajlo Pupin”, Zrenjanin, Republic of Serbia

\*\* Faculty of Technical Sciences, Novi Sad, Republic of Serbia

mihajlovcisinisa25@gmail.com, ivetic@uns.ac.rs, ivana.berkovic62@gmail.com

**Abstract** - The application of Convolutional Neural Networks (CNNs) and mobile devices together has led to significant progress in various fields. Face recognition is used daily as protection on mobile devices. Human emotion is defined by facial expression, and therefore a model for recognizing emotions based on facial expression is proposed in this paper. CNNs play a major role in the recognition and classification of facial expressions.

## I. INTRODUCTION

Nowadays, everyone uses social media, a digital camera, or a mobile phone, where face detection is widely used. There are various popular applications, which use face detection to tag friends or apply filters. Face detection is defined as computer technology used to detect human faces in digital images. In the field of facial detection, there are various algorithms used, and some of them are Convolutional Neural Networks (CNNs) and Viola-Jones algorithm. Face recognition is something that all people know about, but they have never thought about it more deeply. Work is being done daily to improve this feature to provide the best service [1].

Unlike traditional CNN approaches, models mounted on mobile devices must minimize storage requirements while maintaining high performance. Advances in deep machine learning techniques and increased availability of large datasets have led to improved performance in tasks such as predicting emotions, valence, and excitement toward facial images in real-world scenarios, not just in confined environments [2].

The task of automatic facial expression recognition is to recognize each different facial expression and classify it into appropriate classes of emotions [8].

## II. RELATED WORKS

Because of the digital race, the iPhone X has implemented new face detection and recognition software. The two most common methods of face detection are the Viola-Jones algorithm and the Convolutional Neural Network. Paul Viola and Michael Jones, in their 2001 Robust Real-Time Object Detection, proposed a visual object detection framework that was capable of rapid image detection. Convolutional neural networks (CNNs) are the product of facial detection waves. The Viola-Jones algorithm offers great accuracy for frontally placed faces but is less accurate if the faces are tilted or if the side profile is. CNNs offer not only fast detection, but also a lot of different abilities when it comes to face positioning [1].

The Emosic application prompts the user to photograph his face and predicts a visible facial expression based on the category of emotion – neutral, happy, sad, surprised, scared, horrified, angry or despised, as well as the level of valence, i.e., how positive/negative is the influence and excitement shown, i.e., how active/inactive the impact is shown using the convolutional neural network model. Based on the predicted impact on users, the Emosic application recommend several songs to the user. Emosic is the concept that emotionally intelligent user interfaces (EIUI) on mobile devices are now feasible using modern machine learning approaches and large datasets. The trained model achieves prediction performance like the reference dataset and users positively evaluate the song recommendations provided by the interface [2].

Object detection is a very important task for various applications, including autonomous driving, face detection, video surveillance, etc. A CNN-based algorithm is a great solution for detecting objects with high accuracy. Most deep learning applications run on servers or desktops. The architecture of the CNN-based object detection algorithm model on Android devices is based on SqueezeNet to obtain image feature maps and a convolutional layer to find bounding boxes for recognized objects. The size of the model is 8 MB, and the architecture of the model makes the calculation more efficient, and thus allows implementation on mobile devices. The model is based on CNN and is implemented with TensorFlow and Android. The model is trained by the KITTI standard, which contains 10 GB of well-tagged data for object detection purposes. The model can detect objects in view of camera on the Android device [3].

A CNN architecture called ShuffleNet is an extremely efficient computing architecture. It is specifically designed for mobile devices with very limited computing capabilities. The architecture uses two new operations, point group convolution and channel mixing, to reduce computational costs while maintaining accuracy. Experiments on ImageNet classification and detection of MS COCO objects show superior performance of ShuffleNet compared to other structures. Maintaining comparable accuracy, ShuffleNet achieves 13 times greater acceleration than AlexNet [4].

To place deep CNN on mobile devices, there is an efficient whole CNN compression scheme, called one-time compression of the entire network. The scheme consists of three steps: rank selection with Bayesian matrix variation factoring, Tucker decomposition to the kernel tensor, and fine-tuning to recover accumulated loss of accuracy, and each step can be easily performed using

publicly available tools. A significant reduction in model size operating time and energy consumption is obtained, at the cost of a small loss in accuracy [5].

The popularity and portability of mobile devices with high-quality cameras is influencing the development of real-time face recognition using CNNs for ordinary users. CNNs are not easy to implement for real-time applications on mobile devices, because storage, memory, and computing power are relatively limited. It is designed to respond to the CNN architecture for face recognition on mobile devices. The main challenge is processing images of users' faces on the devices itself without uploading images to an external cloud server for processing. The advantages are the security of user information and privacy, and the processing speed on mobile devices, because data does not have to be loaded over the network [6].

A model with less than a million parameters was created by adjusting the width and depth of the bottleneck. The proposed CNN model is fine-tuned to use less than 0.5 million parameters. Fine-tuned CNN model achieved an accuracy of 90.3% for 5 classes and 86.8% for 7 classes in the Realworld Affective Faces (RAF) database composed of facial images [7].

Two approaches to facial expression recognition are unit detection and facial point detection. The first approach was implemented using a framework called FACS (Facial Action Coding System). The framework quantifies a person's facial expression by observing changes in facial muscles when an emotion is triggered. FACS is characterized by the movement of muscles around 44 areas of the face or so-called action units (AUs). Facial expression can be recognized by the existence and intensity of several AUs. Facial expressions have two main steps, AU detection and AU recognition. Performing this task requires the use of the Deep Convolutional Neural Network which has an architecture consisting of a filter layer and a classification layer. The filter stage includes a convolutional layer, followed by a temporal pooling layer and a soft max unit [8].

Face recognition is a technique used to verify or identify a person's identity by analyzing and linking patterns based on a person's facial features. The variance of expression and the intensity of light affect the speed of face recognition. The most effective technique is based on machine learning that has solved these problems [9].

### III. PROPOSED METHOD OF RECOGNIZING EMOTIONS BASED ON FACIAL EXPRESSIONS

There are different facial expressions. Each facial expression represents one emotion. The FER (Facial Expression Recognition) dataset consists of 6 sets of images. Facial expressions are divided into six classes: happy, sad, surprised, fear, disgust, angry, and in addition neutral expression were added [10].

An expression of a smile is an expression that can show that someone feels happy or that they like something. A happy expression is on the upward movement of the cheek muscles and the side or edges of the lips to form a smile [11].

The facial expressions of anger stem from what is expected and reality. The expression is shown on both sides of the inner eyebrows that join and lean down, while the lips are narrowed, and the eyes are sharp when looking [11].

A face that shows sadness appears when there is disappointment or a feeling that something is missing. Based on the characteristics of a sad facial expression when the eye loses focus, the lips are pulled down and the upper lid is lowered [11].

A form of expression that occurs when someone experiences an inability to cope with any event or in a terrifying atmosphere, then that person is said to be afraid. The expression of fear on the man's face can be seen from the two eyebrows that are raised at the same time, the eyelids are clenched, and the lips are open horizontally [11].

A person who expresses his face in a state of disgust because he sees something that is not usual or listens to information that is not worth hearing. The expression of disgust will be read when a person's face in the area of the nasal bridge folds and the upper lip is raised [11].

Expressions of surprise are obtained when someone does not know in advance an event or a received message that is sudden, unexpected, or important. the expression is a shocked face represented by raised eyebrows with wide open eyes and a reflex of opening the mouth [11].

A neutral face is an empty expression that implies a lack of noticeable emotions. Most of the time, an emotionless face is defined by a flat mouth, unfocused eyes, and relaxed cheeks. Although it conveys negativity to some, others see it as a reflection of calm. Figure 1 shows a sample of dataset containing images of facial expression [11].

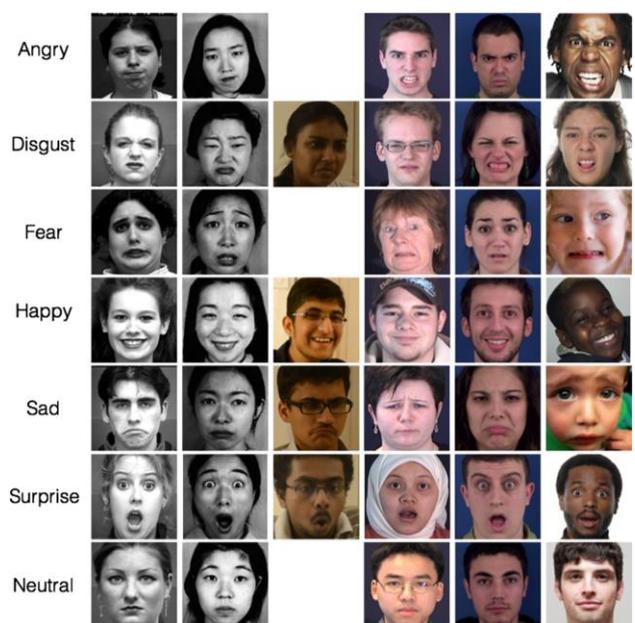


Figure 1. Sample of dataset containing images of facial expression [10]

The facial expression recognition method consists of three main steps: face detection, CNN feature extraction,

and modeling. The method takes as input  $n$  images from different datasets [10].

The face contained in the picture is found and cropped. These images represent the input to the convolutional neural network to extract relevant characteristics. The task of recognition is through modeling performed by a machine learning algorithm. Face detection is performed using the Viola-Jones algorithm, which shows good compromise performance between accuracy and speed of face detection. Each image is pre-processed, and the output is a vector with three pairs of dots indicating the position of the face. The images are then cropped covering the entire face and transformed into a gray scale for the purpose of standardizing for the next step [10].

The second step uses the VGG – face descriptor, a specialized model trained to recognize faces using a deep learning approach. This model consists of 2.6 million faces obtained from YouTube and with minimal manual tagging. The MatConNet framework enables the integration of the VGG face model. Each image is prepared with the final size, scaling, or resizing the image as needed. Grayscale images are used to reduce the variability in brightness in images. Changing images to grayscale does not reduce performance, and CNN models allow images to be represented through neuronal activations, where each neuron acts as a filter that matches and highlights features such as edges, contours, saturation, or gradients, giving CNN features a large capacity to learn correspondence. The final step aims to teach the model to distinguish facial expressions that represent CNN characteristics. Achieving this step is done by comparing between different traditional machine learning algorithms [10].

On mobile devices, it is necessary to create an application, which would recognize the face via the phone's camera, and determine the emotion based on the facial expression. The classification of facial expressions is done using CNN and depends on how well trained she is. Trained CNN should contain as few parameters as possible, for easier and more efficient execution on mobile devices. Mobile devices are not powerful enough, and therefore it is necessary to customize the application to be able to use it.

#### IV. DISCUSSION

Face recognition is used daily, and one of the applications is on mobile devices. It is most used for the purpose of unlocking mobile devices. The face is something that defines a person because it is unique. A person has different facial expressions, and each facial expression represents the emotion of whether he is happy, sad, angry, etc. Some of the challenges are recognizing facial expressions and determining emotions based on facial expressions. Small variations in facial expressions can differentiate emotions. All this leads to the difficulty of determining emotions based on facial expressions.

The advantage is that CNNs can be used because they can classify images and recognize facial expressions. It is necessary to train a model based on CNN. A trained model should have as few parameters as possible, and

this is extremely difficult to achieve. The number of parameters determines the execution speed of CNN. Speed is very important, as well as accuracy and efficiency. Mobile devices are powerful, but they have their drawbacks such as storage, power, and the ability to calculate quickly.

Using the phone's camera, face recognition can be performed, and emotion can be determined based on facial expressions. It is necessary to create a prototype application for the computer, and then it can be adapted for mobile devices. In further research, a prototype application will be created and tested.

#### V. CONCLUSION

Mobile devices are evolving every day and have a wide application. Convolutional neural networks are becoming more prevalent, and it brings innovations every day, and deficiencies are eliminated. The paper describes the applications of CNNs on mobile devices for recognizing emotions based on facial expressions. A method for recognizing emotions based on facial expressions has been proposed.

It is necessary to create an application that will be able to recognize emotion based on facial expressions. Recognizing facial expressions and determining emotions will be done using convolutional neural networks. CNNs are powerful, accurate and fast. The disadvantage is storage, because CNNs require a large capacity, as well as a powerful configuration that will be able to do it all. It is necessary to modify the application to be used on mobile devices, because mobile devices are not as powerful as computers. This is one of the biggest challenges that will be solved in further work and research.

#### ACKNOWLEDGMENT

Ministry of Education, Science and Technological Development, Republic of Serbia financially supported this research, under the project number TR32044: "The development of software tools for business process analysis and improvement", 2011-2021.

#### REFERENCES

- [1] K. Enriquez, Faster face detection using Convolutional Neural Networks & the Viola-Jones algorithm, 2018.
- [2] C. Hewitt & H. Gunes, CNN-based Facial Affect Analysis on Mobile Devices, 2018.
- [3] Z. Li, R. Zhang, Object Detection and Its Implementation on Android Devices, cs231n
- [4] X. Zhang, X. Zhou, M. Lin, J. Sun, ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices, 2018.
- [5] K. Yong-Deok, P. Eunhyeok, Y. Sungjoo, C. Taelim, Y. Lu, S. Dongjun, Compression of Deep Convolutional Neural Networks for Fast and Low Power Mobile Applications, 2016.
- [6] G. Yuanyuan, X. Yifan, W. Jing, Y. Hui, C. Rung-Ching, Real-Time Facial Affective Computing on Mobile Devices, Sensors 20, 2020.
- [7] C. Yoon, D. Kim, Mobile Convolutional Neural Networks for Facial Expression Recognition, 2020 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1315-1317, 2020.

- [8] D. Y. Liliana, Emotion recognition from facial expression using deep convolutional neural network, *J. Phys.: Conf. Ser.* 1193 012004, 2019.
- [9] S. Khan, M. H. Javed, E. Ahmed, S. A. A. Shah, S. U. Ali, Facial Recognition using Convolutional Neural Networks and Implementation on Smart Glasses, 2019 International Conference on Information Science and Communication Technology (ICISCT), pp. 1-6, 2019.
- [10] S.M. González-Lozoya, J. de la Calleja, L. Pellegrin, et al., Recognition of facial expressions based on CNN features, *Multimedia Tools Applications*, Vol. 79, pp. 13987–14007, 2020.
- [11] L. Zahara, P. Musa, E. Prasetyo Wibowo, I. Karim and S. Bahri Musa, The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi, 2020 Fifth International Conference on Informatics and Computing (ICIC), pp. 1-9, 2020.

# APPENDIX

## Conference Chairs

**Višnja Ognjenović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Dalibor Dobrilović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**co-chairman Evgeny Cherkashin**, *Institute of High Technologies, Irkutsk, Russia*

**co-chairman Andrijana Bocevska**, *Faculty of Information and Communication Technologies - Bitola, North Macedonia*

## Organizing Committee

**Višnja Ognjenović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Dalibor Dobrilović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Dragica Radosav**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Miodrag Ivković**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Biljana Radulović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Ivana Berković**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Duško Letić**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Branko Markoski**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Željko Stojanov**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Ljubica Kazi**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Eleonora Brtka**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Vladimir Brtka**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Zoltan Kazi**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Siniša Mihajlović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Velibor Premčevski**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Aleksandra Stojkov**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Dijana Karuović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Vladimir Šinik**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Nadežda Ljubojev**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Vladimir Karuović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Evgeny Cherkashin**, *Institute of System Dynamic and Control Theory SB RAS, Russia*

**Anastasia Popova**, *Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences*

**Andriana Bočevska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*

**Blagoj Ristevski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*

**Natasa Blazeska-Tabakovska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*

**Tome Dimovski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*

**Nikola Rendeovski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*

**Maja Gaborov**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Milica Mazalica**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Igor Vecštejn**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Marko Blažić**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

**Filip Tsvetanov**, *South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria*

**Goran Gecin**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

## **Program Committee**

- Višnja Ognjenović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia - president*
- Dalibor Dobrilović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Ljubica Kazi**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Dragica Radosav**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Dragana Glušac**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Borislav Odadžić**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Duško Letić**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Miodrag Ivković**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Biljana Radulović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Ivana Berković**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Ljubica Kazi**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Vladimir Brtka**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Branko Markoski**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Željko Stojanov**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Ilija Hristoski**, *Faculty of Economics - Prilep, North Macedonia*
- Zoltan Kazi**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Eleonora Brtka**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Jelena Stojanov**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Vesna Makitan**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Nadežda Ljubojev**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Dijana Karuović**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Vladimir Šinik**, *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*
- Igor Nedelkovski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Aleksandar Markovski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Violeta Manevska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Pece Mitrevski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Ilija Jolevski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Dragan Gruevski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Konstadina Veljanovska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Sonja Mančevska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Snežana Savoska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*

- Ramona Markoska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Mimoza Bogdanoska-Jovanovska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Božidar Milenkovski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Zoran Kotevski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Nikola Rendeovski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Andrijana Bocevska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Blagoj Risteovski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Monika Markovska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Tome Dimovski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Marina Blažekovik Toševski**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Nataša Blažeska Tabakovska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Lela Ivanovska**, *Faculty of Information and Communication Technologies, University "Sv. Kliment Ohridski", Bitola, North Macedonia*
- Andrey Gachenko**, *Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences. Irkutsk, Russia*
- Andrey Mikhailov**, *Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences. Irkutsk, Russia*
- Anastasia Popova**, *Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences. Irkutsk, Russia*
- Alexey Daneev**, *Irkutsk State Transport University, Irkutsk, Russia*
- Denis Sidorov**, *Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia*
- Viacheslav Paramonov**, *Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia*
- Andrey Dorofeev**, *Institute of High Technologies, Irkutsk National Research Technical University, Irkutsk, Russia*
- Gogolák László**, *Subotica Tech - College of Applied Sciences, Subotica, Serbia*
- Zlatko Čović**, *Subotica Tech - College of Applied Sciences, Department of Informatics, Subotica, Serbia*
- Zora Konjović**, *University Singidunum, Centar Novi Sad, Serbia*
- Siniša Nešković**, *Faculty of organizational sciences, University of Belgrade, Serbia*
- Nataša Gospić**, *Faculty of transport and traffic engineering, Belgrade, Serbia*
- Željko Trpovski**, *Faculty of technical Sciences, Novi Sad, Serbia*
- Branimir Đorđević**, *Megatrend University, Belgrade, Serbia*
- Slobodan Jovanović**, *Faculty of Information Technology, Belgrade, Serbia*
- Željko Eremić**, *College of Technical Sciences - Zrenjanin, Serbia*
- Rajnai Zoltán**, *Obuda University, Budapest, Hungary*

**Mirjana Pejic Bach**, *University of Zagreb, Croatia*

**Androklis Mavridis**, *Aristotel University of Thessaloniki, Greece*

**Madhusudan Bhatt**, *R.D. National College, University of Mumbai, India*

**Amar Kansara**, *Parth Systems LTD, Navsari, Gujarat, India*

**Narendra Chotaliya**, *H. & H.B. Kotak Institute of Science, Rajkot, Gujarat, India*

**Zeljko Jungic**, *ETF, University of Banja Luka, Bosnia and Hercegovina*

**Saso Tamazic**, *Univerisity of Ljubljana, Slovenia*

**Marijana Brtko**, *Centro de Matemática, Computação e Cognição, Universidade Federal do ABC, São Paulo, Brazil*

**Zoran Cosic**, *Statheros, Split, Croatia*

**Istvan Matijevics**, *Institute of Informatics, University of Szeged, Hungary*

**Slobodan Lubura**, *Faculty of electrical engineering, University of East Sarajevo, Republic of Srpska, Bosnia and Hercegovina*

**Edit Boral**, *ASA College, New York, NY, USA*

**Dana Petcu**, *West University of Timisoara, Romania*

**Marius Marcu**, *"Politehnica" University of Timisoara, Romania*

**Aleksej Stevanov**, *South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria*

**Petar Apostolov**, *South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria*

**Filip Tsvetanov**, *South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria*

**Francesco Flammini**, *School of Innovation, Design and Engineering, Division of Product Realisation, Mälardalen University, Eskilstuna, Sweden*

**Deepak Chahal**, *Jagan Institute of Management Studies (JIMS, Rohini Sector-5), New Delhi, India*

**Abdel-Badeeh M. Salem**, *Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt*

**Dragan Peraković**, *University of Zagreb, Faculty of Transport and Traffic Sciences, Croatia*

**Gordana Jotanović**, *University of East Sarajevo, Faculty of Transport and Traffic Engineering, Doboje, Republic of Srpska, Bosnia and Herzegovina*

**Goran Jauševac**, *University of East Sarajevo, Faculty of Transport and Traffic Engineering, Doboje, Republic of Srpska, Bosnia and Herzegovina*

## List of Reviewers

**Željko Stojanov**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Vladimir Brtko**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Eleonora Brtko**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Vladimir Šinik**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Ivana Berković**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Dalibor Dobrilović**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Ljubica Kazi**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Vesna Makitan**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Zoltan Kazi**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Višnja Ognjenović**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Dijana Karuović**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Tome Dimovski**, *Faculty of Information and Communication Technologies, University “Sv. Kliment Ohridski”, Bitola, North Macedonia*

**Evgeny Cherkashin**, *Institute of High Technologies, Irkutsk, Russia*

**Deepak Chahal**, *Jagan Institute of Management Studies (JIMS, Rohini Sector-5), New Delhi, India*

**Zoran Kotevski**, *Faculty of Information and Communication Technologies, University “Sv. Kliment Ohridski”, Bitola, North Macedonia*

**Blagoj Ristevski**, *Faculty of Information and Communication Technologies, University “Sv. Kliment Ohridski”, Bitola, North Macedonia*

**Biljana Radulović**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Duško Letić**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Jelena Stojanov**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Goran Jauševac**, *University of East Sarajevo, Faculty of Transport and Traffic Engineering, Doboje, Republic of Srpska, Bosnia and Herzegovina*

**Gordana Jotanović**, *University of East Sarajevo, Faculty of Transport and Traffic Engineering, Doboje, Republic of Srpska, Bosnia and Herzegovina*

**Dragana Glušac**, *University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia*

**Snežana Savoska**, *Faculty of Information and Communication Technologies, University “Sv. Kliment Ohridski”, Bitola, North Macedonia*

**Tibor Szakall**, *Subotica Tech - College of Applied Sciences, Department of Informatics, Subotica, Serbia*

**Zlatko Čović**, *Subotica Tech - College of Applied Sciences, Department of Informatics, Subotica, Serbia*

**Andreja Samčović**, *University of Belgrade, Faculty of Transport and Traffic Engineering, Belgrade, Serbia*

**Dragan Peraković**, *University of Zagreb, Faculty of Transport and Traffic Sciences, Croatia*

## Index of Authors

### A

Ali, Ilker 104  
Arsovski, Saša 138

### B

Babić, Bojan 180  
Bakator, Mihalj 36, 41, 46  
Berković, Ivana 126, 175, 203  
Bhatt, Mm 133  
Blazheska Tabakovska, Natasa 50, 104  
Blazić, Maroko 9  
Bocevska, Adrijana 50  
Brtka, Eleonora 116, 180  
Brtka, Vladimir 148

### C

Čeliković, Milan 85  
Cherkashin, Evgeny 152  
Ćočkalo, Dragan 36  
Ćočkalo-Hronjec, Melita 36  
Čović, Zlatko 194

### D

Đalić, Nataša 46  
Dinu, Dragan 26, 90  
Dobrilović, Dalibor 194  
Đorđević, Dejan 36  
Đorđević, Filip 15  
Dimovski, Tome 199  
Dimova Popovska, Hristina 199

### F

Felbab, Srđan 194

### G

Gaborov, Maja 9, 20, 116  
Gajić B., Dušan 26  
Gračić, Saša 110

### H

Hristoski, Ilija 199

### I

Ilić, Luka 100  
Ilić, Miloš 70, 190  
Ivanovska, Lela 32  
Ivetić, Dragan 15, 20, 95, 175, 203

### J

Jauševac, Goran 142, 148  
Jolevski, Ilija 50, 65, 79  
Jotanović, Gordana 142, 148  
Jovanović, Jovana 26  
Jovanović, Martin 7  
Jovanovski, Damjan 60

### K

Karuović, Dijana 9,  
Kavalić, Mila 9, 41, 46  
Kazi, Ljubica 126, 133  
Kazi, Zoltan 126  
Kopylova, Viktoria 152  
Košut, Zlatko 172  
Kotevski, Zoran 120  
Krstić, Dragana 55

### L

Lukyanov, Nikita 152  
Ličina Boris 100

### M

Makitan, Vesna 148  
Markoski, Branko 138, 172  
Masliković, Dejan 126  
Mazalica, Milica 74, 95  
Mihajlović, Siniša 133, 175, 203  
Mijoska, Mimoza 65  
Mikić, Vladimir 70, 190  
Milić Keresteš, Neda 1  
Milić, Tamara 116  
Milosavljev, Dragana 9, 36, 41, 46  
Minić, Siniša 55  
Mitreska Jovanovska, Elena 60  
Moh Said, Hanaa 157

### N

Naneva, Veselina 162  
Naumoski, Andreja 60  
Nedelkovski, Igor 104  
Nejković Valentina 167

### O

Ognjenović, Višnja 116

### P

Peraković, Dragan 142  
Petrović B., Veljko 26

Petrović, Nenad	55, 167	<b>V</b>	
Pireci Sejdiu, Nora	79	Vasiljević, Aleksandar	85
Popov, Srđan	20, 85, 90	Vecštejn, Igor	9, 116, 180
Popović, Zoran	55	Veljanovska, Kostandina	2
Popovska, Katja	60	Viduka, Dejan	100
Premčevski, Velibor	138	Vuković, Vuk	110
		Vulikić, Natalija	126
<b>R</b>		<b>Z</b>	
Radenković, Maša	167	Zakić, Aleksandar	70, 190
Radišić, Rade	90	Zec, Saša	172
Radosav, Dragica	41, 46, 126	Zlatković, Dragan	70, 190
Radulović, Biljana	74, 126		
Risteovski, Blagoj	50, 65, 79, 104		
<b>S</b>			
Savoska, Snežana	50, 65		
Shijakova Kotevski, Aneta	120		
Sim Jia Wun, Angely	138		
Simon, Jozsef	194		
Sisak, Mira	148		
Sofić, Aleksandar	138		
Stanisavljev, Sanja	172		
Stanković Pešterac, Tijana	126		
Stefanova, Kremena	162		
Stojanov, Željko	142, 185		
Stojanović, Željko	172		
Stojkov, Aleksandra	74, 185		
Suljović, Suad	55		
<b>T</b>			
Tashevski, Panche	65		
Terek Stojanović, Edit	41		
Thê Trần, Đức	152		
Trajkovic, Vladimir	50		