RCITD 2017

Proceedings in
Research Conference In Technical Disciplines

The 5th International Virtual Research Conference In Technical Disciplines

13. - 17. November 2017

Slovak Republic

Published by: EDIS - Publishing Institution of the University of Zilina
Univerzitna 1
01026 Zilina
Slovak Republic

Editors: Ing. Michal Mokrys; Ing. Stefan Badura, Assoc. Prof. Dragan Peraković

ISBN: 978-80-554-1375-4

cdISSN: 1339-5076

eISSN: 2453-6571

DOI: 10.18638/rcitd.2017.5.1

Pages: 67

Printed in: 65 copies

Publication year: 2017

- All published papers undergone single blind peer review.
- All published papers are in English language only. Each paper had assigned 2 reviewers and each paper went through two-tier approval process.
- Papers are published as delivered by authors without content modification. All accepted papers have been formally checked by the conference Technical Committee and regularly reviewed in single blind peer reviewing process by the conference Scientific Committee (Reviewers Committee).
- The publication contents are the sole responsibility of the publication author and it does not reflect the Publishing Society, Zilina, Slovakia.

Open Access Online archive is available at: **http://www.rcitd.com/archive**
(proceedings will be available online one month after the publication release).

In case of any questions, notes or complaints, please contact us at: **info(at)rcitd.com**.

Paper Citation Example: In (Eds.) S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman, Proceedings of The 5th International Virtual Research Conference in Technical Disciplines (RCITD-2017), Vol. 5, pg #-#. Zilina, Slovakia.

International Scientific Committee and Reviewers Committee

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Sections discussed at the Conference

Information and Communication Networks and Services
Information and Communication Security
Research and Development of Information and Communication Assistive Technology
Economics of Information and Communication Systems
Telecommunications Policy
Other Information and Communication systems
Other Applied Computer Sciences
Power Systems and energy
Automotive - Safety engineering
Automotive - Ecology/emissions

eISSN: 2453-6571, cdISSN: 1339-5076

ISBN: 978-80-554-1375-4

Conference Sponsors and Partners

Publishing Society
Slovak Republic

www.The-Science.com

Science & Technology Park Zilina
Slovak Republic

Faculty of Transport and Traffic Sciences
Department of Information and Communication Traffic

Published by

EDIS
Publishing Institution of the University of Zilina

Univerzitna 1
01026 Zilina
Slovak Republic

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Summary

Dear Reader,

it is our pleasure to introduce you this Proceedings. This book contains all accepted papers from conference, which is described below in more details. We hope published papers contribute to the academic society and provide interesting information for researchers world wide.

Conference details:

- Conference full name: **Research Conference In Technical Disciplines**
- Conference short name: **RCITD**
- Conference edition: **5th**
- Conference dates: **November 13 - 17, 2017**
- Conference web page: **www.rcitd.com**
- Conference online archive: **www.rcitd.com/archive**

Conference paper approval process:

Each registered paper was evaluated in double tier approval process.

1. Scientific Committee evaluation (in average 2 reviews were prepared per paper).
2. Conference Editorial Board.

Only papers recommended by these committees were accepted for online presentation at the conference and for publication in this conference book.

Conference presentation:

All accepted papers were presented at the conference during conference dates. Asynchronous online chat was prepared for each paper, where all conference members could freely discuss the topic. During the conference, short presentation were proposed for effective conference discussion.

Conference post processing:

- Each published paper is available in **CDRom** and **online** proceedings (online archive).
- All papers in online archive are available for downloading for potential readers (**open access policy**).
- For effective linking and citation each paper is provided with **DOI**, we recommend to use DOI for referencing.
- The online archive may be possibly evaluated by several Abstracting and Indexing services in future.

Sincere thanks for:

- **Scientific Committee** for their volunteer work during reviewing.
- Conference **partners** for promotional work and their contribution.
- **Editorial Board** for enormous workload and patience.

Michal Mokrys
Conference Organizing Committee
The-Science.com community
Publishing Society

November, 2017

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Table of Contents

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Multi-criteria Decision-making in Cloud Service Selection and Adoption

Ivan Grgurević
Department of Information and Communication Traffic
Faculty of Transport and Traffic Sciences
University of Zagreb
Zagreb, Croatia
ivan.grgurevic@fpz.hr

Gordana Kordić
Department of Information and Communication Traffic
Faculty of Transport and Traffic Sciences
University of Zagreb
Zagreb, Croatia
gordana.kordic@fpz.hr

Abstract—**Cloud computing (CC) offers its unique scalable and all time available services to the user. As service provision is an important part of the Cloud computing concept, the proper choice of the needed service according to the users' needs is of most relevance. This research provides a review of Cloud service selection approach of multi-criteria decision analysis (MCDA). MCDA enables the user to choose from several available choices based on the chosen criteria. In this paper, the criteria and subcriteria for choosing Cloud services in accordance with Cloud computing service model have been analysed.**

Keywords- Cloud computing, decision-making, multi-criteria decision analysis, choice of criteria

I. INTRODUCTION

The emergence of Cloud services is one of the breakthrough points in the history of computing. Cloud services are offered in three models of computing to the end user, and those are Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) [1].

The proposed Cloud services and their applicability and adoption for the customer are described in Chapter 2. When Cloud computing customers are making decision about selecting the most appropriate service; at first, they consider the nature of their individual needs and ask themselves many questions about the Cloud adoption issue. Those questions become decision-making criteria and customers' needs become a unique use-case scenario of Cloud service demand. Using Multi-criteria Decision Methods (MCDM), the posed criteria can be evaluated and the selection of a proper Cloud service can be made.

This research aims to gather knowledge about the selection and adoption of Cloud services using MCDM methods for all three types of Cloud-services individually.

Chapter 3 gathers most popular existing MCDM methods for different user's use-case scenarios of selecting the proper Cloud service. As customer must select the best service according to their needs out of the abundance of possibilities, complex dependencies and heterogeneous sets of criteria must be well considered [2].

The provided overview of the gathered research knowledge is set in a comparative table (Chapter 3) and discussed in Chapter 4, offering conclusive facts regarding the selection process of the most appropriate service for Cloud customer.

II. APPLICABILITY AND ADOPTION OF CLOUD COMPUTING SERVICES

The following text of this Chapter provides the description of three main Cloud computing models which are the subject of research in defining the key criteria as an integral component of multi-criteria analysis.

A. IaaS

Infrastructure as a Service (IaaS) refers to the provision of raw machines (servers, storage, networking and other devices) on which the service consumers deploy their own software (usually as virtual machine images). IaaS providers host users' applications and handle tasks including system maintenance, backup and resiliency planning. IaaS platforms offer highly scalable resources that can be adjusted on demand. This makes IaaS well-suited for workloads that are temporary, experimental or change unexpectedly. Other characteristics of IaaS environments include the automation of administrative tasks, dynamic scaling, desktop virtualization and policy-based services.

IaaS customers pay on per-use basis, typically by the hour, week or month. Some providers also charge customers based on the amount of virtual machine space they use. This pay-as-you-go model eliminates the capital expense of deploying in-house hardware and software. Once the new software is tested and refined, it can be removed from the IaaS environment for a more traditional in-house deployment or to save money and free the resources for other, new projects. Leading IaaS providers include Amazon, Dimension Data, Windows Azure, Google Compute Engine, Rackspace Open Cloud, and IBM SmartCloud Enterprise, etc. [3]- [4]- [5].

B. PaaS

Platform as a Service (PaaS) is a provision service of development platform and environment providing services and storage, hosted in the Cloud. Most PaaS platforms are geared toward software development and offer developers several advantages. For example, PaaS allows developers to frequently change or upgrade operating system features. Also, PaaS is found very helpful for collaboration of development teams on projects.

Users typically access PaaS through a Web browser. PaaS providers charge for the access the users have made on a per-use basis. Some PaaS providers charge a flat monthly fee to

access the platform and the apps hosted within it. When making decision about PaaS service, it is important to evaluate service uptime and support offered by the considered PaaS provider.

Salesforce.com's Force.com is a well-known PaaS vendor that provides an enterprise customer relationship management (CRM) platform. Other PaaS platforms for software development and management are Appear IQ, Mendix, Amazon Web Services (AWS) Elastic Beanstalk, Google App Engine, Red Hat, Heroku, etc. [3]- [4]- [5].

C. SaaS

Software as a Service (SaaS) is a software distribution model in which third-party provider hosts applications and makes them available to customers over the Internet. SaaS is used in a number of common business areas, including customer relationship management (CRM), document management, accounting, human resource (HR) management, service desk management, content management and collaboration.

SaaS customers have a lot of choice among providers. When making the selection, security and all-time-availability of service are most often the demand criteria. Except for the mentioned and according to their specific needs, customers usually pose a wide set of criteria as there are thousands of SaaS vendors on the global Cloud service market, such as Salesforce.com, Oracle, Google and SAP.com, Apprenda, Cloudswitch, Marketo, Pardot, etc. to choose from [3]- [4]- [5].

III. Review of Cloud service selection and adoption

Within this chapter, a review of Cloud service selection and adoption is given through three substantial Cloud services – IaaS, PaaS and SaaS, separated in individual sub-chapters.

A. IaaS

As most IaaS Cloud computing service packages cover the storage, networking, servers, and virtualization components, and the IaaS customers are usually responsible for installing and maintaining the operating system, databases, security components, and applications, decision-making is of extreme importance. For the IaaS, there is a great number of methods, as all MCDM methods can be used. It is up to the user needs and use-case scenario which of the MCDM methods will perform better.

The research [6] states that Analytic Hierarchy Process (AHP) is the most efficient MCDM method for managing information security in Cloud computing as well as for task scheduling and resource allocation.

The main limitation of multi-criteria optimization techniques such as genetic algorithms, is that they cannot handle mixed qualitative and quantitative criteria. Therefore, research [1] applies a finite set of alternative purposes and uses AHP to resolve mixed quantitative and qualitative criteria. The research proposes a hybrid multi-goal optimization heuristic method, i.e. the use of optimization techniques such as genetic algorithm (GA) together with AHP to easily evaluate alternatives and find an optimal solution according to the AHP-based evaluation.

Research [5] proposes a Cloud service selection model based on the Analytical Network Process (ANP). According to the proposed model, the criteria and subcriteria (subfactors) for Cloud service selection are identified and evaluated by weights. ANP is one of techniques suggested to solve complex decision-making problems and therefore it is relevant for identification of the best choice of IaaS service. For the organization of choice selection model, paper [5] uses three criteria (support perspective, provider perspective and service perspective), and eight subcriteria (service brand, service price, availability, performance, scalability, security, service level agreement and service support). Based on the chosen criteria and the ANP method, the research makes a selection out of six proposed IaaS alternatives (6).

Research [6] explores use-case scenario for different multi-criteria decision-making in Cloud computing. Also, the strengths and weaknesses are provided for every multi-criteria decision-making described. The analyzed methods are AHP, TOPSIS, VIKOR, ELECTRE and PROMETHEE. This research evaluates the methods of Cloud service selection and is relevant for selecting the most efficient method for the type of Cloud service of interest.

Research [7] describes decision-making of small and medium enterprises (SME) about the choice of Cloud services, as well as their valorization criteria. The research explores the use of AHP method for decision-making in the Cloud environment. The proposed AHP model has offered financial, marketing, management and environment criteria (4), and 14 subcriteria, such as often considered payment on demand, intensive and timely information, the performance of CRM application, information security and privacy, agility and adaptability, Cloud technology reliability, etc. The method has added values (weight) for every criterion and ranked subcriteria by order of importance.

Research [8] offered a concise practical approach to choosing a Cloud provider. The authors have explored AHP and PROMETHEE (Preference Ranking Organization METHod for Enrichment Evaluations) and the Goal programming techniques for assessing the weights of the selected criteria. The research has chosen 12 criteria; security protocols enabled, file sharing capabilities, maximum file size upload, free storage space, supported operational systems, ease of use, technical support, version control, service provider reputation, additional free storage space, mobile internet support and market share. Dropbox, SugarSync, Google Drive, Microsoft SkyDrive, Apple iCloud, Amazon, MegaCloud. JustCloud and Ubundu One were evaluated according to the stated criteria. The choice of Cloud provider was finally obtained, and PROMETHEE has given SugarSync, GoogleDrive and Microsoft SkyDrive as top three providers, while AHP has given SugarSync, Dropbox and JustCloud as top three on the rank list.

Research [9] describes decision-making of a business migrating to the Cloud. The research does not propose any methodology of choice, but offers the identification of several criteria of vendor evaluation, and those are accountability, agility, assurance, costs, performance, security and privacy.

eISSN: 2453-6571, cdISSN: 1339-5076
10.18638/rcitd.2017.5.1.104

ISBN: 978-80-554-1375-4

Research [10] has chosen two (2) methods – TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) and VIKOR (VIseKriterijumska Optimizacija I Kompromisno Resenje) for IaaS Cloud service selection. The authors proposed a new decision-making model made of these two, with one method as the main method and the other as an instance method. The developed decision-making evaluation model contains three criteria, i.e. values of RAM, Bandwidth and Storage. After a 30-day evaluation of TOPSIS and VIKOR with ten (10) Cloudsim toolkit simulated Clouds, VIKOR has outperformed TOPSIS, in terms of evaluated criteria of memory and time.

Paper [11] has presented a fuzzy multi-criteria group decision-making method for evaluating the performance and the choice of Cloud services. Based on the gathered literature, the most relevant criteria were chosen, and those are security, performance, accessibility, usability, scalability, and adaptability.

Paper [12] provides an overview of different MCDM methods and their evaluation. The seven MCDM methods are used (7), and those are Outranking Methods, TOPSIS, PROMETHEE, Multi-Attribute Value Theory (MAVT), Multi-Attribute Utility Theory (MAUT), Elimination Et Choice Translating Reality (ELECTRE) and AHP. The research had also provided comparative study of the given MCDM methods with their strengths and weaknesses.

Research [13] uses multi-criteria decision-making methods to rank the service providers based on their infrastructure parameters. A combination of analytic and fuzzy method is found more trustworthy in comparison with any analytic method alone. Ranking of Cloud Service Providers (CSPs) is based on the Key Performance Indicators called Service Measurement Index (SMI) which help the organizations measure the Cloud-related business services based on their specific business and technology requirements. The indicators considered in the trust estimation of CSPs are agility, financial, performance, security and privacy, and usability (5). For the comparison of services, four (4) infrastructure CSPs were considered: Gogrid, Rackspace, Amazon EC2 and Cloudflare (alternatives).

B. PaaS

PaaS is often an IT business choice, as it can edit the operating system features and easily upgrade, and its major focus is on security, storage and database integration. PaaS services can be used from variety of international sources. PaaS Cloud solution serves as an application hosting, testing, deployment and development environment with reduction of total expenses involved.

Research [14] compares prominent PaaS Clouds: Microsoft Windows Azure (MWA), Google App Engine (GAE) and GroundOS (GOS) by several criteria of choice, and those are programming language and frameworks, databases, availability, security, services, customer care and price (7). The research has not proposed a method for the selection of PaaS provider.

Table 1 shows more significant studies that are associated with the application of multi-criteria decision-making in cloud service selection (related to cloud models).

TABLE I. STUDIES ASSOCIATED WITH THE APPLICATION OF THE MULTI-CRITERIA DECISION-MAKING IN CLOUD SERVICE SELECTION (RELATED TO CLOUD MODELS)

Studies	Year	Methods (MCDA)	Model of Cloud services	Number of criteria / subcriteria	Note
B. Do Chung and K.-K. Seo [5]	2015	ANP	IaaS	3 criteria and 8 subcriteria	Evaluation of 6 Cloud services
D. Rai and V. P. Kumar [10]	2016	TOPSIS and VIKOR	IaaS	3 criteria	Evaluation with 10 Cloudsim toolkit simulated Clouds
M. Supriya, K. Sangeeta, G.K. Patra [13]	2016	AHP, Fuzzy-AHP, TOPSIS and Fuzzy-TOPSIS	IaaS	5 criteria	Evaluation of 4 Cloud services: Gogrid, Rackspace, Amazon EC2 and Cloudflare
S. Lee and K.-K. Seo [21]	2013	AHP	IaaS	3 criteria and 8 subcriteria	-
T. Devi, R. Ganesan [14]	2015	-	PaaS	7 criteria	Evaluation of 3 PaaS Clouds: Microsoft Windows Azure (MWA), Google App Engine (GAE) and GroundOS (GOS)
J.J.H. Park, H-Y. Jeong [3]	2013	A new QoS-based MCDM system is proposed	SaaS	6 criteria and 25 subcriteria	Evaluation of 3 different types of SaaS ERP systems
P. Costa [15]	2013	MACBETH	SaaS	19 criteria	Evaluation of 2 Cloud services: Google Apps and Microsoft Office 365
M. Godse, S. Mulik [16]	2009	AHP	SaaS	5 criteria and 16 subcriteria	Evaluation of 3 leading SaaS products (as A, B, and C instead of using their real names)
N. Boussouali m and Y. Aklouf [18]	2014	AHP	SaaS	7 criteria and 25 subcriteria	Evaluation of 3 SaaS offers (A, B, and C)

10.18638/rcitd.2017.5.1.104

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

O. Boutkhoum, M. Hanine, T. Agouti and A. Tikniouine [20]	2016	Fuzzy AHP and PROMETHEE	SaaS, PaaS and IaaS	3 criterias and 10 subcriteria	Evaluation of 5 Cloud computing products

Research [6] proposes different MCDM methods for different Cloud services. The authors emphasize the use of all MCDM methods for decision-making on all Cloud services, regardless of the model, but they suggest the use of TOPSIS for PaaS decision-making.

C. SaaS

Decision-making problems in SaaS environment are of great importance. For example, SaaS vendor selection is one of the scenarios where multi-criteria decision-making is necessary. The method that is most commonly used to resolve SaaS-related decision-making problem is AHP, which is used for SaaS vendor selection, SaaS product selection, customer centered Cloud service selection and ranking of Cloud computing services [6].

The author of research paper [15] uses MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique) method to simplify the decision-making process in organizations adopting Cloud services. The proposed solution model is based on 19 chosen criteria that facilitate decision-making among two SaaS Cloud services: Google Apps and Microsoft Office 365 (2). Some of those criteria are commonly used in MCDM, such as availability, data integrity, maintainability, interoperability, service response time (SRT), cost, confidentiality and data loss, reliability, compliance with standards, laws and regulations, service level agreements capacity (SLA), innovation, client support, risks, acquisition and transaction cost, adaptability, elasticity, etc. Also, some are more specific in nature, like service response time, service level agreement, acquisition and transaction cost, confidentiality and data loss. The proposed evaluation criteria are individually described with performance reference levels, as either good or neutral.

Paper [16] presents an approach of using the AHP technique for prioritizing Cloud service features and ranking (three Cloud services). Based on the experience and the opinion of experts, the research proposes the following criteria for the selection of SaaS, and those are functionality, architecture, usability, vendor reputation, and cost (5).

Research [3] proposes the use of QoS model to make a service selection. The model consists of six (6) criteria: functionality, reliability, usability, efficiency, maintainability and portability, and 25 subcriteria: suitability, accuracy, interoperability, compliance, sustainability, maturity, recoverability, fault tolerance, elasticity, security, learnability, easy to use, operability, scalability, adaptability, time behavior, resource behavior, throughput and efficiency, stability, analyzability, changeability, testability, updateability, service use cost and update cost. The proposed QoS model is in fact a multi-criteria decision-making (MCDM) model formed to find the best SaaS ERP (Enterprise Resource Planning) in the Cloud computing environment and provide recommendations to customer in provided priority order.

Research [17] proposed a primitive cognitive network process (P-CNP) for SaaS evaluation and selection problem. Further research lies in planning to integrate the PROMETHEE and fuzzy technique into the P-CNP to form a new enhanced hybrid decision-making model.

Research [18] used 7 criteria and 25 subcriteria for the selection of SaaS Cloud service. The method used in research is AHP. Proposed criteria were CRM, collaboration, reputation, cost, usability, structure, configurability and personalization, and subcriteria included commonly mentioned safety, scalability and success control among the rest.

Research [19] used a Petri net multi-criteria decision-making (MCDM) framework to evaluate SaaS. The strategy, quality, performance, security and economic criteria were evaluated.

IV. Discussion & Conclusion

Based on the gathered, there are several multi-criteria decision methods that are used most for resolving the selection problems of Cloud environment. The methods that are found as the most efficient ones are AHP, ANP, Fuzzy-AHP, PROMETHEE, TOPSIS, Fuzzy-TOPSIS, VIKOR and ELECTRE. The often mentioned and used decision-making criteria are security, performance, accessibility, usability, scalability, and adaptability. Those are the criteria that most precisely correspond to the customer (i.e. user) needs when making the Cloud service selection. For every business decision of Cloud service selection, the costs seem to be highly evaluated, followed by privacy and security matters. It is also noticeable that when deciding about SaaS, the future customer imposes more criteria and subcriteria in number to decision-making (Table 1). Also, the majority of organizations adopting SaaS are interested in agility and innovation; those adopting PaaS are interested in scalability, and the companies that adopt IaaS Cloud service most often expect capital and operational expenditures and therefore high demand of agility.

The research has found that not every decision-making method is the best for every Service of Cloud computing environment. Depending on the nature of Cloud computing selection problem there are options that are found most efficient. For example, for SaaS AHP is recommended, for PaaS TOPSIS and for IaaS all the MCDM methods are found equally appropriate. With the use of several basic MCDM methods, more precise hybrid methods are being proposed. Hybrid, i.e. expanded and combined, methods strengthen the customer confidence in the method and facilitate the final decision.

Further research will include a more practical approach regarding Cloud service selection and adoption. The following research will provide the evaluation results of criteria that are found as relevant within this research, all with the use of AHP and PROMETHEE MCDM methods (using the appropriate program tools). The emphasis shall be on the ranking procedure of user requests (criteria and subcriteria) regarding the required Cloud services. Accordingly, appending the priorities and preferences to certain criteria and alternatives

eISSN: 2453-6571, cdISSN: 1339-5076
10.18638/rcitd.2017.5.1.104

ISBN: 978-80-554-1375-4

(Cloud services and/or product) will be of the utmost importance.

ACKNOWLEDGMENT

This research has been carried out as part of the project "Evaluation of Existing Systems to Support Sharing of Information to Carpooling Users based on Cloud Computing" (No. 303-02/17-01/19) supported by the Faculty of Transport and Traffic Sciences, University of Zagreb, Programme for Research, 2017.

REFERENCES

[1] M. Zhang, R. Ranjan, A. Haller, D. Georgakopoulos and P. Strazdins, "Investigating decision support techniques for automating Cloud service selection", 4th IEEE International Conference on Cloud Computing Technology and Science Proceedings, Taipei, 2012, pp. 759-764, DOI: 10.1109/CloudCom.2012.6427501

[2] M. Whaiduzzaman, A. Gani, N. B. Anuar, M. Shiraz, M. N. Haque, and I. T. Haque, "Cloud Service Selection Using Multicriteria Decision Analysis", The Scientific World Journal, vol. 2014, Article ID 459375, 10 pages, 2014. DOI:10.1155/2014/459375

[3] J.J.H. Park, H-Y. Jeong, "The QoS-based MCDM system for SaaS ERP applications with Social Network", *The Journal of Supercomputing*, vol. 66, no. 2, pp. 614-632, 2013, DOI.org/10.1007/s11227-012-0832-4

[4] D. Linthicum, "TechTarget Network", TechTraget, 26 August 2016. [Online]. Available at: searchcloudcomputing.techtarget.com [Accessed 20.08.2017]

[5] B. Do Chung and K.-K. Seo, "A Cloud Service Selection Model based on Analytic Network Process", Indian Journal of Science and Technology, vol. 8, No. 18, August 2015, DOI: 10.17485/ijst/2015/v8i18/77721

[6] R. K. Gavade, "Multi-Criteria Decision Making: An overview of diferent selection problems and methods", *International Journal of Computer Science and Information Technologies*, vol. 5, no. 4, pp. 5643-5647, 2014.

[7] Y.-H. Lee, "A Decision Framework for Cloud Service Selection for SMEs: AHP Analysis" *SOP Transactions on Marketing Research*, vol. 1, no. 1, pp. 51-57, 2014

[8] J. Papathanasiou, V. Kostoglou and D. Petkos, "A comparative analysis of cloud computing services using multicriteria decision analysis mehodologies", International Journal of Information and Decision Sciences, vol. 7, no. 1, pp. 23-35, 2014, DOI: 10.1504/IJIDS.2015.068117

[9] A. Alkhalil, R. Sahandi, D. John, "Migration to Cloud Computing: A Decision Process Model", Central European Conference on Information and Intelligent Systems - CECIIS-2014, Varaždin, Croatia, pp. 154-163, 2014

[10] D. Rai and V. P. Kumar, "Instance based Multi Criteria Decision Model for Cloud Service Selection using TOPSIS and VIKOR", International Journal of Computer Engineering and Technology, vol. 7, no. 1, pp. 78-87, 2016

[11] S. Wibowo, H. Deng and W. Xu, "Evaluation of cloud services: A fuzzy multi-criteria group decision making method", *Algorithms*, vol. 9, no. 84, pp. 1-11, 2016, DOI: 10.3390/a9040084

[12] S. Saroj, V. K. Dileep, "A Review Multi-criteria Decision Making Methods for Cloud Service Ranking", International Journal of Emerging Technologies and Innovative Research, vol. 3, no. 7, pp. 92-94, July 2016, Available at: http://www.jetir.org/papers/JETIR1607019.pdf

[13] M. Supriya, K. Sangeeta, G.K. Patra, "Trustworthy Cloud Service Provider Selection using Multi Criteria Decision Making Methods", *Engineering Letters*, vol. 24, no. 1, pp. 1-10, 2016

[14] T. Devi, R. Ganesan, "Platform-as-a-Service (PaaS): Model and Security Issues", International Journal of Advances in Applied Sciences (IJAAS), vol. 4, no. 1, pp. 13-23, 2015, DOI: 10.11591/telkomnika.v15i1.8073

[15] P. Costa, "Evaluation Criteria for Cloud Infrastructure as a Service", Instituto Superior Técnico, Portugal, 2013

[16] M. Godse, S. Mulik, "An Approach for Selecting Software-as-a-Service (SaaS) Product", Proceedings of the 2009 IEEE International Conference on Cloud Computing, Bombay, September 2009, pp. 155-158, DOI 10.1109/CLOUD.2009.74

[17] K. K. F. Yuen, "Software-as-a-Service evaluation in cloud paradigm: Primitive cognitive network process approach", 2012 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC 2012), Hong Kong, 2012, pp. 119-124, DOI: 10.1109/ICSPCC.2012.6335719

[18] N. Boussoualim and Y. Aklouf, "An Approach based on user preferences for selecting SaaS product", 2014 International Conference on Multimedia Computing and Systems (ICMCS), Marrakech, Morocco, 2014, pp. 1182-1188, DOI: 10.1109/ICMCS.2014.6911278

[19] M. Ribas, A. S. Lima, N. Souza, A. Moura, F. R. C. Sousa and G. Fenner, "Assessing cloud computing SaaS adoption for enterprise applications using a Petri net MCDM framework", 2014 IEEE Network Operations and Management Symposium (NOMS), Krakow, 2014, pp. 1-6, DOI: 10.1109/NOMS.2014.6838372

[20] O. Boutkhoum, M. Hanine, T. Agouti and A. Tikniouine, "Selection Problem of Cloud Solution for Big Data Accessing: fuzzy AHP-PROMETHEE as a proposed methodology", Journal of Digital Information Management, vol. 14, No. 6, December 2016

[21] S. Lee and K.-K. Seo, "A Multi-Criteria Decision-making Model for an IaaS Provider Selection Problem", International Journal of Advancements in Computing Technology, vol. 12, No. 5, pp. 363–7, 2013

An Overview of the Cyber Security Strategic Management in Republic of Croatia

Ivan Cvitić, Dragan Peraković, Marko Periša
Department of Information and Communication Traffic
Faculty of Transport and Traffic Science
University of Zagreb
Zagreb, CROATIA
ivan.cvitic@fpz.hr, dragan.perakovic@fpz.hr,
marko.perisa@fpz.hr

Mate Botica
OiV, Transmitters and Communications Ltd.
Zagreb, CROATIA
mate.botica@oiv.hr

Abstract—**Development of information and communication system and its frequent use in various sectors (private and business) requires strategic approach in protection of users' privacy and data. This paper analyzes strategic development of cyber security in Republic of Croatia and gives a comparison with EU member's strategy. Also, this paper identifies deficiencies and gives the improvement guidance of National cyber security strategy according to guidelines prescribed by ENISA.**

Keywords- information security, critical communiation, cyber crime, data protection, crisis management

I. INTRODUCTION

Information and communication (IC) systems development over the last four decades has led to a continuous increase in the need for remote communication, transmission and exchange of data, information and other IC resources. The benefits that IC systems provide to users have led to the development of new services, technologies and various terminal devices (traditional computers, smartphones, tablet devices, portable devices, etc.) as well as new concepts such as Internet of Things (IoT) and Cloud Computing. Such a dynamic environment is characterized by an increasing amount of data traffic generated by users and devices. Traffic content often contains confidential and sensitive information from private and business users that are necessary to protect while transferring, processing, and storing from unauthorized access, modification, and deletion.

The trend of continuous growth in the number of users and the volume of data traffic is also accompanied by the growth of security incidents. Security incidents can be intentional or accidental and may undermine the delivery of key services such as water supply, health services, electricity and mobile communication services. Security threats and incidents in the business sector also have an impact on the global economy of the European Union (EU). Additionally, the application of various, sophisticated attack methods for the purpose of unauthorized access to the IC system and stealing confidential information, industrial espionage and unauthorized activities supported by other states is steadily increasing.

The aforementioned factors are a major challenge for many countries around the world, but also a trigger for the development of security strategies that observe cyber space as a growing international problem. The European Commission and EU High Representatives for Foreign Affairs and Security Policy in February 2013 highlighted the EU's role in the problem area, clarified its roles and responsibilities and defined a set of necessary activities based on robust and effective civil rights protection with a goal of creating a secure cyber EU environment.

Republic of Croatia has been striving to plan and manage cyber security since 2005 when it publishes its first strategic document in the problem area under the National Program for Information Security (National Program). As an output of mentioned document the numerous legal, by-law acts, and national bodies responsible for the implementation of the legislative framework has been established. The rapid and dynamic development of the IC environment requires continuous improvement of the strategic, legislative and executive framework. Accordingly, by the proposal of the Cyber Security Strategy of the European Union, the member states have begun the development of a national strategy specific to each member state. On October 7, 2015, the Government of the Republic of Croatia adopted a Decision on the adoption of the National Cyber Security Strategy (National Strategy) and the Action Plan for the implementation of the National Cyber Security Strategy (Action Plan), and thus continue the strategic management of this problem area.

The aim of this research is to review the current development of strategic cyber security management in the Republic of Croatia and identify the potential disadvantages of the National Strategy. The purpose of the research is to point the importance of strategic planning and management in cyber security.

II. STRATEGIC PLANNING OF CYBER SECURITY AT THE NATIONAL LEVEL

The strategic planning of cyber security in the Republic of Croatia began on March 30, 2005 by the publication of the National Program by the Central Office for e-Croatia (SDUeH). The document defines the strategic objectives of information security at the national level as well as the competences and duties of certain stakeholders regarding information security. The objectives set out in the document are linked to the definitions of the 2005 e-Europe Program, the Stability Pact for Europe and the NATO Action Plan for NATO membership in which the Republic of Croatia was an active participant [1].

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

The National Program is primarily intended to define the organizational and management aspects of information security introduction, as well as the prerequisites for systematic development of laws, regulations, methods, procedures and technical systems in the area of information security. The strategic goal of the National Program was its gradual expansion to the country as a whole through the introduction of minimum security criteria in the public and private sector and through education and development of information security culture in all user structures [1].

The National Program defines the basic terminology of information security, and terms used in the National Program. Through the basic concepts it is defined:

- Security personnel check,
- Physical security,
- Data security,
- Security of information systems and
- Security of third party access and external cooperation.

The document also covers information security requirements set by international institutions such as NATO and the EU. Each strategic document is necessary to be based on an analysis of the current situation which is shown through the review of the current legislative, institutional and infrastructural framework of information security in the Republic of Croatia [2].

The implementation of the National Program is planned through three phases that precede certain pre-emptions. Some of defined preliminary work that needed to be carried out include the following [2]:

- Establishing a new expert group to monitor the implementation of the National Program,
- Adoption of the framework for Information Security law,
- Amendments to the Law on Security Services of the Republic of Croatia in the field of information security,
- Establishment and start of the Office for Information Security and Crypto Technology,
- Establishment of the Agency for Support of Information Systems,
- Establishment of National CERT, etc.

Numerous activities have been carried out through the three phases of implementation of the National Program, most of which involve the adoption of new or harmonization of existing legal and subordinate provisions, rules, guidelines and information security recommendations. Along with this, a number of educational programs and catalogs of knowledge have been developed and implemented at all educational levels, and new state authorities responsible for information security at the state level have been established [2], [3].

National Program and the Plan of Implementation of the National Program defines the implementation of measure "Adoption of the National Strategy and Information Security Policy". Although the measure had to be implemented by December 2005, publicly available information, according to the author's findings, does not exist, and it is concluded that this measure has not been implemented although, according to the National Program, the mentioned measure represents the core document and is prerequisite for all other regulations [2], [3].

III. The Development of the Cyber Security at European Union Level

The need to preserve the openness and availability of IC space implies the application of all norms, principles and values of the European Union relating to the domain of physical space. This implies the protection of fundamental rights, democracy and law in the IC environment. For that purpose the European Commission started in 2013 with the development of a strategy aimed at providing guidelines and principles needed for defining the cyber security policy at EU and international level. The basic EU values that need to be protected in the IC environment are relate to [4]:

- Protection of fundamental rights, freedom of expression, personal data and privacy,
- Equal access to Internet infrastructure and information to all citizens and social layers,
- Democratic and efficient management of Internet resources by a large number of stakeholders and
- The division of responsibility in providing protection among the relevant actors.

The trigger for cyber security strategy development [5] was a cyber-attack on Estonia. The attack was politically motivated and took place over a period of three weeks. The target of the attack was government institutions, websites of political parties, banks, media, Internet service providers and private organizations. During the attack, the DDoS (Distributed Denial of Service) method were used [6]. The attack has highlighted the numerous vulnerabilities of the IC system at national level, resulting in the development of the first Cyber Security Strategies in Estonia (2008), Australia (2009), Canada (2010), the Czech Republic (2011), France (2011), Germany (2011), the Netherlands (2011) and the United Kingdom (2011) [7] - [14].

Generally, the strategy is a document that defines the goals to be achieved, but also measures to achieve set goals as well as institutional responsibilities. The cyber security strategy primarily seeks to ensure the fulfillment of the confidentiality, integrity and availability of IC resources. Priority of such strategy is the protection of national critical infrastructure [5].

A. Strategic Priorities and Measures

The EU's strategic goal is to protect the IC environment by ensuring the highest degree of freedom and security for all users. It is the responsibility of member states to address security challenges in the IC environment, and the European Cyber Security Strategy proposes concrete measures that can improve overall EU efficiency.

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Defined strategic priorities in cyber security are as follows [4]:

1) Achieving resistance to cyber threats,
2) Drastic reduction of cybercrime,
3) Develop cyber protection policy in line with Common Security and Defense Policy (CSDP),
4) Development of industrial and technological resources for the establishment and conservation of cyber security and
5) Establish coherent international cyberspace policy at EU level and promote EU core values.

ENISA (European Network and Information Security Agency) is responsible for the implementation of these strategic priorities at EU level. ENISA is a body established by the European Commission in 2004 and is responsible for defining guidelines for managing the national cyber security strategy of the European Union member states [15]. According to ENISA, the national cyber security strategy is a tool for improving security and resilience to the threats of the national information and communication infrastructure and related services. Additionally, it is a top-down approach to cyber security whose purpose is to set up a set of national goals and priorities to be achieved within a given time period. As such, it provides a strategic framework for cyber security management at national level [16].

B. Elements of Cyber Security Strategy

According to the research [17], the generic model of the national cyber security strategy, shown in Figure 1, consists of three basic elements, inputs, bases and outcomes based on the recommendations of leading European and world bodies (ITU, NATO, OECD and EU). The input is the mission, vision and objectives of the strategy, with the need to achieve a clear understanding of the critical information infrastructure that needs to be protected.

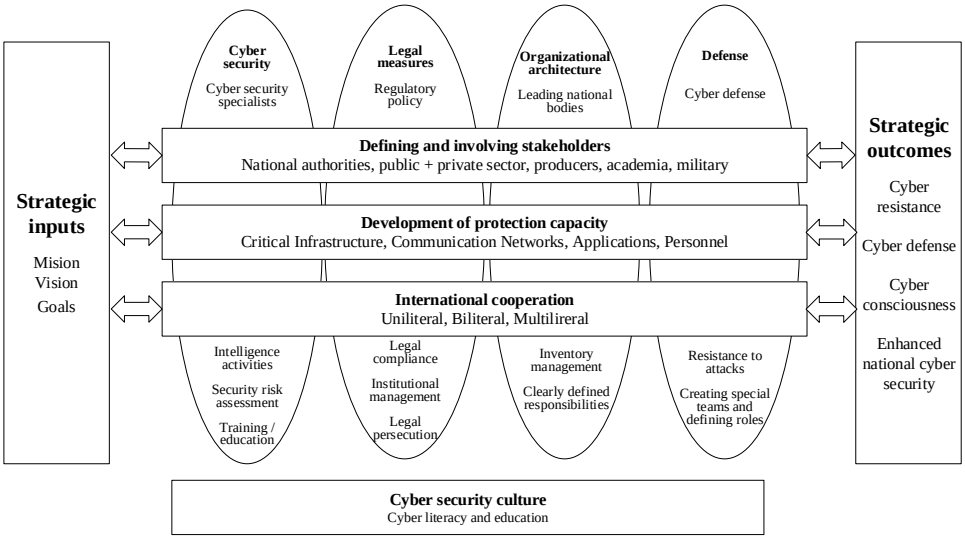


Figure 1. General model of national cyber security strategy [17]

The central supporting element can be further divided into eight sub-elements [16], [17]:

1) The culture of cyber security - is the underlying foundation of the national cyber security strategy, and refers to the way in which society accepts and uses cyber protection measures.
2) Stakeholders - national bodies are a key for creating and implementing the strategy. In addition to national bodies, it is necessary to identify other stakeholders and to clearly define their roles and responsibilities.
3) Development of protection capacity - all necessary measures must be taken to ensure protection against cyber threats. Minimum security requirements must be defined for each sector. It is necessary to apply specific security standards and to select appropriate working frameworks and involve expert staff in developing the strategy.
4) International co-operation - states must cooperate with each other as well as with the leading organizations associated with cyber security due to the international nature of cyber threats.
5) Cyber security - helps in achieving an effective cyber security framework. Includes definition of procedures such as risk management, security policy definition, technical measures tailored to existing systems and communication networks for the purpose of administration, threats identification, audit and monitoring functions.
6) Legal measures - states must continually adopt laws that monitor the dynamics of cyber security with goal of cybercrime protection.
7) Organizational architecture - it is necessary to define the bodies responsible for coordinating the development and implementation of the strategy as well as bodies that will

10.18638/rcitd.2017.5.1.107
eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

participate in the development and implementation of the strategy at the national level. Participating bodies are responsible for conducting activities in all economic sectors.

8) Defense - National security authorities must have developed defense mechanisms and responses to cyber forms of warfare.

The National Strategy outcomes require continuous evaluation by applying key performance indicators (KPI's). Indicators may be quantitative or qualitative and may relate to the overall strategy or to individual activities [16].

IV. Development of the National Cyber Security Strategy in the Republic of Croatia

The dynamics of development and comprehensiveness of IC technology has resulted in the rapid development and distribution of new services and products as a basis for the economic and social development of the country. In such a dynamic environment there is an increasing number of vulnerabilities and threats that they are trying to exploit for the purpose of unauthorized access to IC resources [18]. According to the above, the protection of IC resources in the function of achieving and preserving the required level of security is a continuous technical and technological process that needs to be implemented in all structures of society such as the public, academic, economic sector and citizenship as a whole [19].

The growth trend of adopting IC technologies and services in the public sector is visible through statistical indicators such as penetration of broadband Internet access. Figure 2 shows statistical data on penetration of broadband Internet access in European Union and Croatian companies, with a continuous trend of growth.

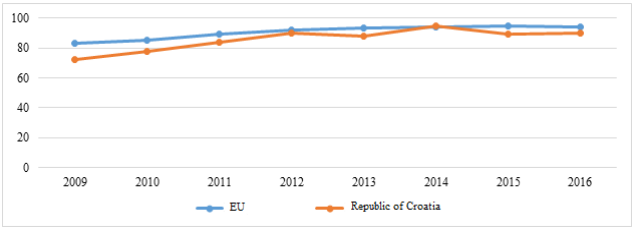


Figure 2. The growth trend of Internet access in companies within the EU and the Republic of Croatia [22]

Figure 3 shows statistical data on the penetration of broadband Internet access by private users of the European Union and the Republic of Croatia, where, as in the previous case, there is continuous growth.

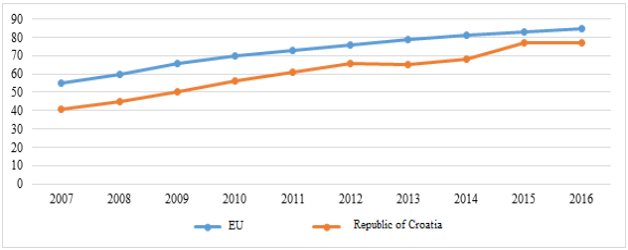


Figure 3. The growth trend of Internet access by private users within the EU and the Republic of Croatia [23]

With the growth of broadband access penetration to the Internet network in the Republic of Croatia and the EU, statistical data also show the rise in the use of the Internet in a function of public communication with public institutions. According to data from 2016, 48% of EU citizens used the Internet as the main communication channel with public institutions, which is an increase of 13% compared to 2008. In the Republic of Croatia, this share is 36% for the same year, an increase of 20% compared to 2008. Impact on growth is the development of a large number of services through the e-Citizens system that serves as a citizen interface in G2C (Government-2-Customer) communication for the use of a large number of public services.

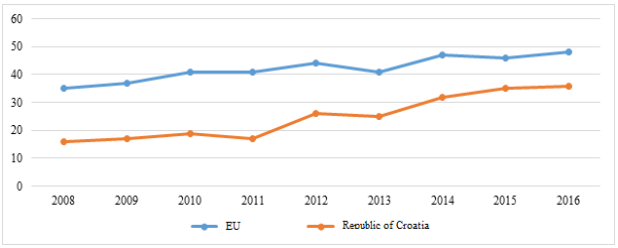


Figure 4. Communication of citizens and public institutions via the Internet network [21]

The presented indicators reflect the dynamics of the ICK environment, whereby the continuous protection of all elements of the IK system plays a key role in further use and the level of acceptance of such a form of communication, business and data exchange of all society structures.

A. *Motives and Reasons for the Adoption of the National Cyber Security Strategy in Republic of Croatia*

The National Cyber Security Strategy of the Republic of Croatia was issued on October 07, 2015, prompted by the European Commission document entitled "Cybersecurity Strategy of the European Union: An Open, Safe and Secure Cyberspace" published on 07 February 2013. This document established that all EU members must have developed a flexible and dynamic cyber security strategy to respond in time to global threats [16]. In addition, the Republic of Croatia has recognized the importance of cyber space security as a common responsibility of all social segments. Accordingly, the National Strategy has been developed for the purpose of systematic and coordinated implementation of activities necessary to increase the level of security of the Republic of Croatia. The primary goal is to identify organizational problems in the implementation of cyber security and to spread and raise awareness of the importance of problems in all social segments. The fundamental role of the National Strategy is to connect different social sectors and to understand complex issues such as cyber security [19].

B. *Principles and Main Goals of the National Strategy*

The National Strategy is based on basic principles such as the comprehensiveness of cyber security, the integration of activities and measures, a proactive approach through continuous adaptation of activities and measures, and strengthening of resistance, reliability and adaptability by establishing basic security principles (confidentiality, integrity and accessibility). Apart from the aforementioned, additional

principles, which are the basis for modern society, such as law enforcement, the development of a harmonized legislative framework, the application of the principle of subsidiarity and the principle of proportionality [19].

The general objectives of the National Strategy are reflected through [19]:

1) A systematic approach to the implementation and development of the national legislative framework,
2) Implementation of activities and measures to increase the security, resilience and reliability of the cyber space,
3) Establishing a more efficient system for data exchange, transfer and access,
4) Strengthening awareness of the security for all users of cyber space,
5) Encouraging the development of harmonized education programs at all educational levels,
6) Encouraging the development of e-services through the development of user trust in e-services by defining minimum security requirements,
7) Encouraging research and development and
8) Systematic approach to international cooperation.

Defined goals are important for the previously identified IC environment dynamics and for monitoring global cybernetic security trends. In this way, by fulfilling the set objectives, harmonization of the legislative framework with international obligations will be achieved, as well as application of appropriate and harmonized IK resource protection standards. In addition, as a product of the National Strategy, an increase in security awareness in all social structures is expected, as well as the activation of the potential and the coordinated work of the academic, public and economic sectors.

C. Functional areas covered by cyber security

The guidelines for development of the National Strategy have been defined and published by ENISA in [16]. Among other, the document also defines a proposal for activities to be undertaken to achieve the ultimate goals of the strategy. One of the activities mentioned is identifying the priority needs of the Republic of Croatia in terms of security. Accordingly, the National Strategy has defined a total of nine cyber security areas [20]:

1) Public electronic communications,
2) Electronic administration,
3) Electronic financial services,
4) Critical communication, information infrastructure and crisis management,
5) Cyber crime,
6) Data protection,
7) Technical coordination in the processing of computer and security incidents,
8) International co - operation and
9) Education, research, development and raising cyber security.

For each area, one or more goals are defined. The defined goals are achieved by implementing one or more measures set out in the Action Plan for Implementation of the National Strategy. Each measure is defined holder and optional co-holder authorities, the beginning and end of implementation and implementation indicators.

D. Identification of the Deficiencies and Improvement Suggestions

In the development of the National Strategy and the Action Plan, the guidelines prescribed by ENISA [28] have been followed. Despite that, review of documents have identified certain shortcomings and possibilities for their improvement Although certain measures, according to defined deadlines, at the time of writing the text are completed, publicly available documents confirming the implementation of the same were not found.

According to ENISA guidelines, a clear methodology of risk assessment or adaptation of existing risk assessment at national level needs to be developed when implementing the Nation Strategy. An overview of existing and available documents (National Strategy and Action Plan) it is noted that there is no clearly defined mentioned activity. Nonetheless, the National Strategy has been evaluated in each area to assess the current state and highlighting the shortcomings that the National Strategy needs to correct or improve. A clearly defined risk assessment methodology presents a set of clearly defined steps and methods applied when identifying, analyzing and evaluating security risk. It is suggested that the used methodology presents an integral part of the National Strategy document. This approach was also applied in the National Cyber Security Strategy of the United Kingdom, based on risk-oriented approach [14].

The success of the National Strategy depends on the valid and effective co-operation of all stakeholders, where ENISA guidelines propose the mandatory inclusion and co-operation of the public and private sector [16]. The National Strategy of the Republic of Croatia does not include private sector stakeholders, but only public sector bodies (DORH, HZZO, MORH, MUP, etc.) or legal persons with public authority (e.g. HAKOM, ZSIS, CARnet, etc.) [20]. Private sector stakeholders should be part of the strategy development process as they are owners of a large number of critical information and communication infrastructure and related services. An example of such a way of involving private sector stakeholders can be seen in the National Cybersecurity Security Strategy of Estonia, whose development and implementation plan was based on the proposals of public bodies and working groups made up of private sector representatives [12].

V. CONCLUSION

The research presents an overview of the development of strategic planning and cyber security management in the Republic of Croatia. Through the work is visible considerable international influence (set requirements for access to NATO

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

and the EU) on the development of strategy from year 2005, whose product is a strategic document entitled National Program for Information Security of the Republic of Croatia. This document has resulted in the development of the legislative framework of the Republic of Croatia on issues of cyber security and many state bodies responsible for the implementation of adopted laws. After accession to the EU it is visible its impact on the development of a new strategic document entitled National Strategy for Cyber Security of the Republic of Croatian and the accompanying Action Plan for the implementation of the National Strategy in 2015. The mentioned document seeks to monitor the dynamics of the development of the IC environment and the growth of the related vulnerabilities and threats, and systematic approaches seek to cover the complex issues of cyber space security.

The strategic planning of cyber security in the Republic of Croatia is guided by the guidelines envisaged by ENISA as the leading EU body for cyber security issues. Despite this, certain deficiencies have been identified in the form of non-compliance and failure to implement the proposed guidelines. In the future, a continuous upgrading of the National Strategy is expected in line with the changes taking place in the domain of issues such as increasing the number of users, developing new concepts such as IoT and cloud computing, and thus the emergence of new vulnerabilities as well as threats to IC assets.

REFERENCES

[1] Systems and Sygnals Laboratory, "Information secrurity in Croatia (Informacijska sigurnost u Hrvatskoj)," Zagreb, Croatia, 2010.

[2] Expert group on information security, "National Information Security Program in the Republic of Croatia," Zagreb, Croatia, 2005.

[3] Expert group on information security, "Implementation plan of the National Information Security Program in the Republic of Croatia," Zagreb, Croatia, 2005.

[4] European Commission, "Cybersecurity Strategy of the European Union: An Open, Safe and Secure Cyberspace," Brussels, Belgium, 2013.

[5] A. Seger, "Cybercrime Strategies," Strasbourg, France, 2012.

[6] C. Czosseck, R. Ottis, and A.-M. Talihärm, "Estonia after the 2007 Cyber Attacks," in International Journal of Cyber Warfare and Terrorism, 2011, pp. 24–34.

[7] National Coordinator for Security and Counterterrorism, "National Cyber Security Strategy," Den Haag, Nederland, 2013.

[8] Ministry of Public Safety, "Canada's Cyber Security Strategy," Canada, 2010.

[9] Department of the Prime Minister and Cabinet, "Australia's Cyber Security Strategy - Enabling innovation, growth and prosperity," Australia, 2016.

[10] National Security Authority, "National Cyber Security Centre of The Czech Republic for the Periode from 2015 to 2020," Czech Republic, 2015.

[11] Federal Ministry of the Interior, "Cyber Security Strategy for Germany," Berlin, Germany, 2011.

[12] Ministry of Economic Affairs and Communication, "Estonia's Cyber Security Strategy (2014-2017)," Estonia, 2014.

[13] SGDSN, "French National Strategy Digital Security Strategy," Paris, France, 2015.

[14] Her Majesty's Government, "National Cyber Security Strategy 2016-2021," London, UK, 2016.

[15] ENISA, "An evaluation Framework for National Cyber Security Strategies," Heraklion, Greece, 2012.

[16] ENISA, "National Cyber Security Strategies: Practical Guide on Development and Execution," Heraklion, Greece, 2012.

[17] R. Sabillon, V. Cavaller, and J. Cano, "National Cyber Security Strategies: Global Trends in Cyberspace," Int. J. Comput. Sci. Softw. Eng., vol. 5, no. 5, pp. 2409–4285, 2016.

[18] H. Elkhannoubi and M. Belaissaoui, "A framework for an effective cybersecurity strategy implementation: Fundamental pillars identification," in International Conference on Intelligent Systems Design and Applications, ISDA, 2016, pp. 1–6.

[19] Croatian Parlament, "National Cyber Security Strategy," Zagreb, Croatia, 2015.

[20] Croatian Parlament, "Action Plan for the Implementation of the National Cyber Security Strategy," Zagreb, Croatia, 2015.

[21] Eurostat, "E-government activities of individuals via websites," 2017. [Online]. Available: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_ciegi_ac&lang=en. [Accessed: 05-Mar-2017].

[22] Eurostat, "Use of mobile connections to the internet in Enterprise," 2017. [Online]. Available: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_cimobe_use&lang=en. [Accessed: 08-Mar-2017].

[23] Eurostat, "Households - level of internet access," 2017. [Online]. http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_ci_in_h&lang=en [Accessed: 01-Mar-2017].

10.18638/rcitd.2017.5.1.107
eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Conceptual informing system architecture for drivers with reduced communication skills

Marko Periša*, Rosana Elizabeta Sente*
*Department of Information and Communication Traffic
University of Zagreb, Faculty of Transport and Traffic Sciences
Zagreb, Republic of Croatia
marko.perisa@fpz.hr, rosana.sente@fpz.hr

Dijana Beganović**, Josip Dankić**, Matija Keleković**, Ana Mežnarek**
**Graduate Student at University of Zagreb
Faculty of Transport and Traffic Sciences
Zagreb, Republic of Croatia
dijana.beganovic.fpz@gmail.com, jdankic440@gmail.com, matijakelekovic31@gmail.com, ana.meznarek@gmail.com

Abstract—**The main goal of the research is to design a system conceptual architecture for adaptive service with the aim to deliver information to drivers in traffic environment which form targeted user group (users with hearing impairment, speech impairment, dyslexia, dysgraphia, color blindness and physical impairment). The purpose of informing service is to offer a real-time information in traffic environment, road assistance service and to raise the quality of life of users and their involvment in traffic environment. The suggested service with its functionalities is based on contemporary technologies for delivering information in traffic environment and architecture is based on Cloud Computing concept. The proposed service is realized in the form of an application solution that provides the user with the necessary information adapted to their disability. With the suggested architecture, it is possible to raise the quality of life and mobility of the users in traffic environment.**

Keywords– information system, quality of life, cloud computing, mobility, assistive technology

I. INTRODUCTION

With the continuous development of information and communication technologies (ICT) there is an increasing supply of mobile applications and services tailored for all types of mobile devices (MD). Development of service for informing users which are moving in traffic environment doesn't affect only the quality of everyday life and mobility, but also the independence of using MD. While offering informing service to the users, it is necessary to take care of simplicity of using that service, functionalities, and the information adaptation for the targeted user group (TUG). This research proposes the conceptual system architecture for delivering informing service to the TUG which are moving through the traffic environment. The functionalities of the suggested service are defined by users' requirements. For this research, the questionnaire has been conducted to obtain users' needs. The results of the questionnaire were used for defining user's requirements of new service and defining the elements of the system architecture that are needed for its deliverance.

II. RESEARCH METHODOLOGY

Materials and methods which are used in this research are currently available scientific literature and questionnaire which collected data about TUG needs and their lifestyle, coping in traffic environment, possibility of education and use of new technologies. The research was conducted in cooperation with the Center for research, education and application of new technologies, UP2DATE. In the city of Zagreb there are 1.784 users with hearing impairment, from which 471 of them have severe hearing loss, complete hearing loss and loss of hearing which is lower than 60 [dB]. According to the Ministry of Interior, Croatian Autoclub, the Ministry of Health and associations of people with mentioned disabillities, the problem has been pointed out in terms of missing the database of total number of drivers who are TUG of this research. The data were collected through the questionnaire in a period from 03.02.2017. to 14.03.2017. and a total of 85 users were questioned.

A. An overview of current research and application solutions for informing users

The existing technologies are oriented on adaptation of the vehicles, but without special dedication to the informing system which is necessary for safe and comfortable ride. Nowadays, most of the drivers and passengers rely on application solutions to access information adapted to MD. Communication for people with hearing and talking impairment is based on application solutions for smartphones. In a case of accident, they send a message using smartphone device. The work of application is based on pictograms or icons in the shape of the signs for danger. To use such application, user must register, give basic information like name, address, gender, age, e-mail, and types of impairment and anamnesis so it can be seen which type of impairment user has [1]. Big challenge for persons with visual impairment is to detect and recognize objects in their environment, and inability to face such a challenge is making them stressed and uncomfortable. Systems based on RFID (Radio-frequency identification) technologies and computer vision are presented

10.18638/rcitd.2017.5.1.103
eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

like solutions for recognizing objects for users with visual impairment [2]. Automatic configuration of application for users with physical disabilities proves that caregiver does not have to be with them to use the application, but the user can use it independently. Configuration is accomplished by using Cloud Computing (CC) technology and it is based on Windows Azure platform [3]. In this research, the main goal was to show positive aspects in which technology, which is constantly upgrading, can be used in the way of overcoming obstacles which persons with disability are facing everyday. Conclusion of that research was that technology would drastically increase quality of life of persons with disability [4]. The study about deaf drivers presents how they get along in traffic environment despite their disability, and what kind of habits they have while driving. Results show that regardless of having hearing impairment, drivers or passengers will adjust their communication methods to different conditions of driving [5]. Assistance, which is presented with mobile technologies for deaf users and users with hearing impairment is implemented on MD as multimodal device for alert named Vibe. Vibe receives information from environment and transfers them to the users, depending on their special needs through visual, tactile or hearing aids [6]. Further research is based on hypothesis that hearing canals are limited for deaf persons and according to that, hearing data must be presented in a visual shape, originally in a sign language [7]. The following research was conducted to determine the audible alarms of emergency service vehicles when they are near drivers who use tactile devices for displaying information. The study shows the speed of driver detection of a horn sound and his response time. Results show that using external tactile display for showing information can provide constant access to the sound of the siren of emergency vehicles to drivers who have hearing impairment [8].

B. Application solutions and informing systems

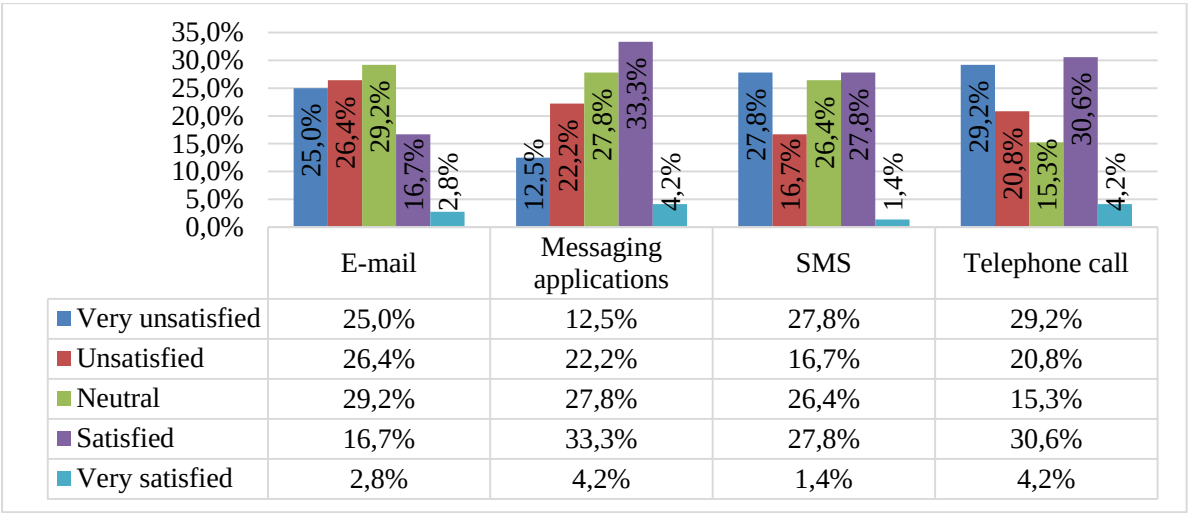
Some of the currently available application solutions that have the purpose of informing users have been analyzed. The analyzed solutions could be useful for TUG in traffic environment, but their functionalities do not meet all their needs. *Dragon Speech Recognition Software* application allows the user to convert their tablet or smartphone into a wireless microphone. The disadvantages of this application are inadequacy for drivers with hearing impairment, difficulty in voice to voice communication and dyslexia. The application is insufficient for all user needs and requirements in terms of traffic informing [9]. *It's Accesible* application supports people with mobility problems who are looking for places like restaurants, hotels, and parking. The lack of application is inaccessible informing system for TUG [9]. *Red Panic button* is an application that, by pressing the red button, alarms all the contacts that the user mentioned in the application. The alarm arrives in the form of a text message, e-mail service, Facebook or Twitter. In addition to the message, a link to user location on Google Maps is also available. People with hearing or speech impairment can not use voice dialing, and alerting is not even adapted for the users with dyslexia, dysgraphia or color blindness. People with physical disabilities can not even leave the vehicle for photographing [9]. *Automobile Association* (AA) is an application for an English auto club that has the possibility to report a crash, Route Planner with travel information (only major traffic jams) and information about parking spaces in the center of the city. The advantages of this application are the ability to report a vehicle breakdown and one type of travel information. The disadvantage is that it is unsuitable for all TUG. The application provides information about parking spaces for standard vehicles but not information about parking spaces for disabled. Travel information provides information about traffic jams, but does not include possible road work and temporary traffic regulations on a specific route [10]. The *Avto-moto zveza of Slovenia* (AMZS) is a Slovenian automobile club that had application solution tailored to the needs of drivers with hearing impairment. The application was only available to members of the Slovenian AMZS. The application worked in the way that the call was converted to SMS, which was then sent to the Association for the Deaf and Hard of Hearing people who contacted the Roadside Contact Center, and later the Road Service. The disadvantage of this application is that when sending SMS request for help, it can not specify the location of the vehicle in malfunction, and in the case of travel information the operator must enter all the information manually and send it via MD to user [11]. In 2014., Hrvatski autoklub (HAK) has offered an application in which, besides the mParking service, the location of the nearest gas stations and other services important for traveling offers also roadside service and quick call for roadside assistance. The application has many advantages for providing traffic information and sending the location of users who need roadside assistance. Although it has numerous advantages for traffic users, it is still not fully adapted to all TUG. HAK has recognized the need to provide adaptive alert service for TUG and provided unique SMS number for them to get help on the road. The disadvantages are that user relies only on his ability to position himself on the map while calling for help which can cause imprecise positioning, manually entering of needed information and financial insignificance [12].

C. Availability of technology for targeted user group

The survey methodology was conducted on the use of contemporary ICT in the field of travel informing for TUG. Total number of respondents was 85, 48 of male and 37 of female respondents. Out of the total number of respondents, 84.7% are drivers. In the case of a vehicle breakdown on a road, 43.1% of the respondents would inform the family first, 34.7% would inform friends, and 30.6% the professional staff. Respondents most frequently reported vehicle malfunction via SMS 52.8%, via messaging applications 38.9%, via telephone call 33.3% and 11.1% via e-mail service. Half of the respondents do not use the informing system while driving. The biggest reason is the inadequacy of the content (58.3%) to the degree of disability, and often the problem is that the respondents are not sufficiently informed about the assistive technologies or the content is not technically accessible (41.7%). Figure 1 shows the overall level of user satisfaction with the system informing services.

eISSN: 2453-6571, cdISSN: 1339-5076
10.18638/rcitd.2017.5.1.103

ISBN: 978-80-554-1375-4



	E-mail	Messaging applications	SMS	Telephone call
Very unsatisfied	25,0%	12,5%	27,8%	29,2%
Unsatisfied	26,4%	22,2%	16,7%	20,8%
Neutral	29,2%	27,8%	26,4%	15,3%
Satisfied	16,7%	33,3%	27,8%	30,6%
Very satisfied	2,8%	4,2%	1,4%	4,2%

Figure 1. User satisfaction with the system with information services

In most cases roadside assistance was needed by 66.7% of respondents, 54.2% needed travel information, 44.1% technical advice, while 11.1% of respondents did not need any help. The accuracy of the information received via SMS, 25% of the users rated 1, while grade 5 gave 5.6% of respondents. Technical accessibility of SMS services was rated with grade 5 by 29.2% of respondents. With a phone call service, 34.7% of respondents rated the accuracy of the received information with grade 5, while grade 1 gave 4.2% of the respondents. Technical accessibility of the call service with grade 1 was rated by 29.2% of respondents, while 2.8% of respondents rated this service with grade 5. Respondents evaluated the accuracy of the information received, the speed of information received, and the technical accessibility of the Messaging application. The accuracy of the information obtained with grade 1 was rated by 15.3% of respondents, while grade 5 gave 6.9% of the respondents. The speed of receiving information 22.2% of respondents rated 1, and 5. was rated by 5.6% of respondents. Technical accessibility for messaging applications 25% of respondents rated 1, and 2.8% of respondents with grade 5. Figure 2 shows the preferred type of information while waiting for roadside assistance.

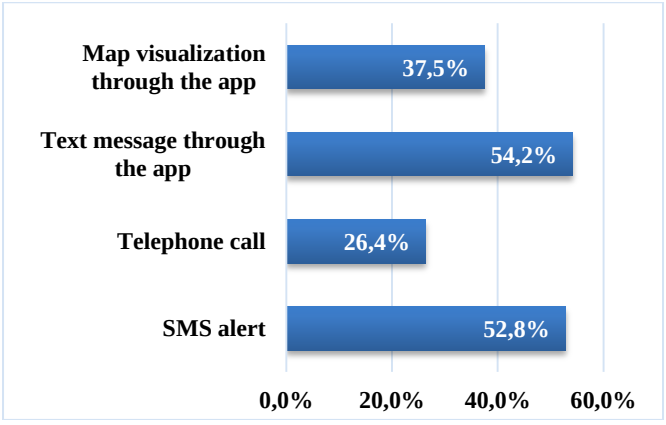


Figure 2. Desired type of information in a time of waiting for road assistance

According to obtained data from the conducted survey, 91.7% of people use a Smartphone device, of which 68.1% use Android operating system, so the service should be based on it. The survey results show that currently available solutions do not meet the TUG needs. The results point to the need for designing new ICT services and they are basis for designing it.

III. Conceptual Informing System Architecture for the Delivery of Information Services

The elements of the system architecture for providing informing service to TUG were defined based on the collected questionnaire results. The proposed system architecture is shown in Figure 3 and consists of relevant elements for delivering information services: computer, mobile device, GPS, mobile application, web application, user database, information database, and functionality of the service. ICT for delivering informing services to user can be divided into: technologies to determine location, communication technologies, and technologies for data processing and storage (CC concept). GPS technology enables accurate location of the user's location with 95% location accuracy (location errors are 8.5 [m]) [13]. These locating errors can have negative affect in navigating pedestrians and for this reason GPS technology is used in combination with other technologies such as Automatic Identification and Data Capture (AIDC) and Bluetooth. 3G and 4G technologies affect location enhancement with the help of the base station and the signal between it and the mobile device. By using LTE technology, it is possible to provide a better quality of real-time services and services based on the user's location [14]. For the purposes of data processing and storage, a CC concept is used, which ensures the availability of information in real time. The user database is in the CC environment and contains user's data and their content accessibility requirements. The information database contains all the data that are used for creating the information for end users. System stakeholders have the option to enter and view the data.

10.18638/rcitd.2017.5.1.103

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

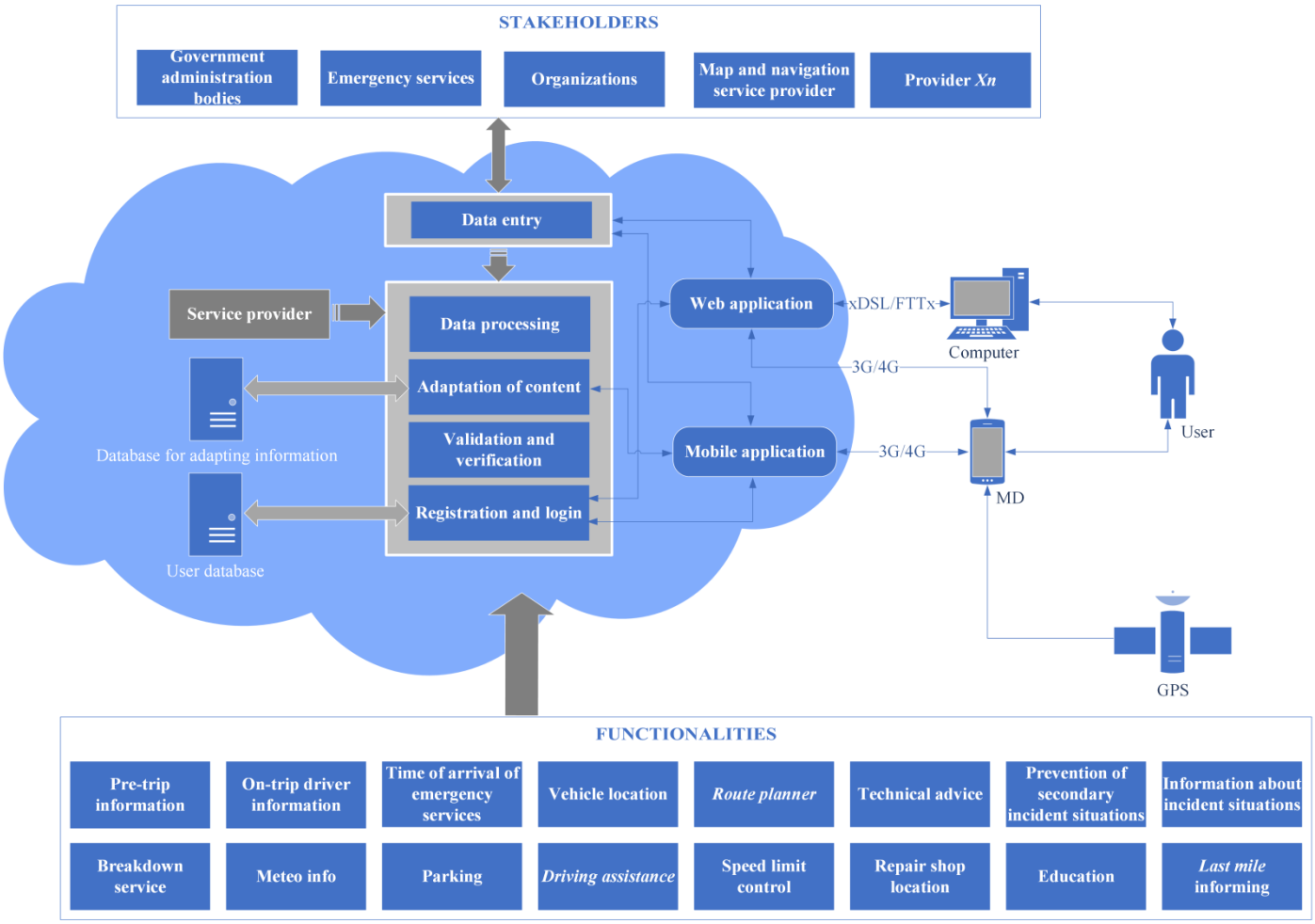


Figure 3. Conceptual informing system architecture for informing targeted user group

The user accesses the mobile application through MD, while access to the web application is possible via a browser on the computer or MD. A mobile application with its functionality and universal design provides customized content to all forms of TUG disabilities. User registration is required to use a mobile application. The registration process includes the entry of basic user information and information on the type of disability. After registration, user data is stored in the user database. When user signs in the application, a validation and verification process is performed to check the accuracy of the entered data. After the successful checkup, the content gets adjusted to user requirements for accessible information. By defining the degree of user's disability, stakeholders are provided with this information to know how to communicate with the user. The role of stakeholders is to participate in the exchange of relevant information and information required for TUG in the terms of accurate and real-time information.

A. *Functionalities of the informing system and the application solution*

The informing system functionalities are divided into two groups: basic functionalities and additional functionalities. The functionalities of real-time information cover incoming and outgoing travel information and information about the time of arrival of a service provider in case of roadside assistance. Additional information service reporting in weather conditions is closely related to road conditions and guides users to adjust speeds during unfavorable weather conditions. Vehicle breakdown entry is possible via SMS, Messaging applications, calls or e-mail services that include vehicle location information, service provider arrival time, technical advice and breakdown service and locations for repairing vehicles. Users who use voice to voice communication have the SOS call option that has priority over other calls, ensuring fast and efficient help in case of vehicle breakdown or incident. Additional functionalities of the system include: last mile information, information about free parking slots adapted for people with disabilities, education about driving through architectural barriers, weather services, Route planner and driving assistance. To use the before mentioned functionalities, it is necessary to use the mobile application, and it is necessary to perform a user registration beforehand. The registration process involves entering basic user data and defining forms of disability and difficulty. Figure 4. shows the appearance of application solution (user registration, breakdown service, home screen).

10.18638/rcitd.2017.5.1.103

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

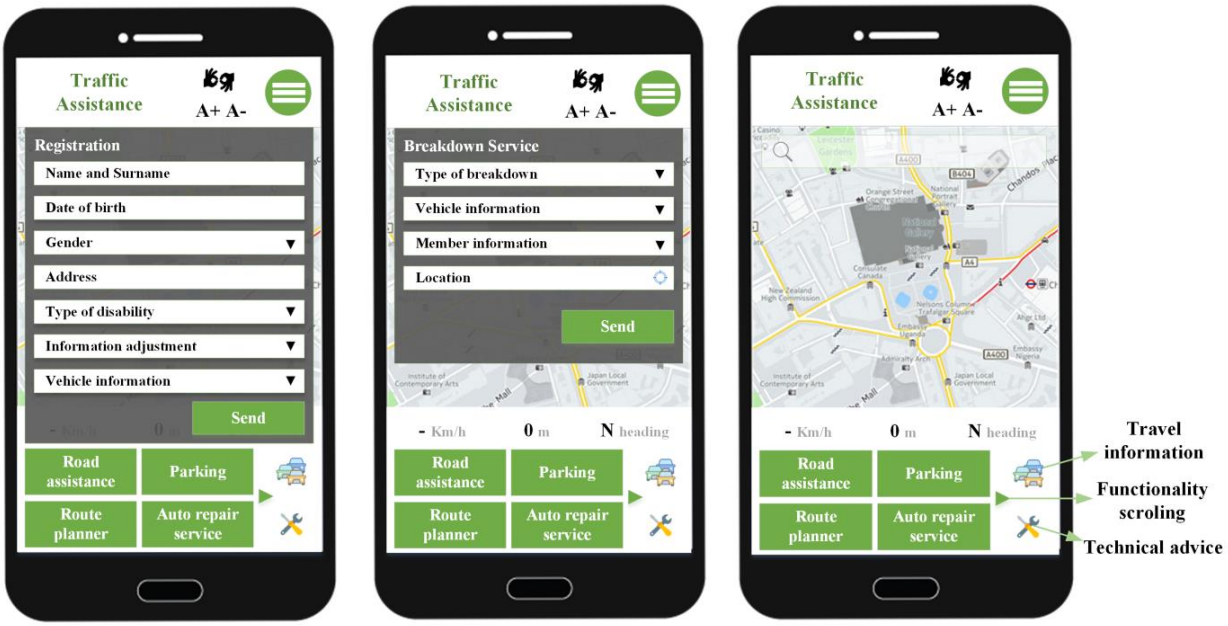


Figure 4. Proposition of application solution

When registering, the user defines the type of disability that is used to adapt the information available within the application solution. The home screen shows a map showing the location of the user, and below the map are some of the functionalities that the user can use. If user chooses functionality that is based on a vehicle location, it is immediately applied to the map. Stakeholders participating in the delivery of information receive a user request for a breakdown report. Users can enter detailed information such as type of breakdown, vehicle information, and member information.

B. Additional functionalities and stakeholders

Last mile information is an additional functionality that has the purpose of informing TUG before the end of their travel. It provides drives with information about locations close to their destination such as hotels, resorts, sights, venues, restaurants, etc. One of the additional functionalities is information about free parking space. Service on the number of free parking spaces is intended for users who need a parking space that is adapted for people with disabilities. It provides a locating service for free parking spaces that are the closest to the user's destination and are adapted to their needs. Another additional functionality that this application solution would provide is education. It is intended for users with physical disabilities that use wheelchair to obtain information on how to drive safely on a road that has architectural barriers. It's is also intended to inform them about weather conditions and how to drive during the bad weather. Users would be educated in safe conditions (on a polygon, in safe driving center, etc.). Meteo info is another additional functionality that provides the TUG with real-time information about weather conditions. It provides them with suggestions to lower the speed during the rain, turn on the lights during the mist, etc. It also includes the speed control system to prevent incident situations caused by unattended speed in adverse conditions. Route planner is the functionality that allows the TUG to determine which route will be the fastest to reach the service for repairing vehicles. It is related to the functionality of locating auto repair services. It also provides them with the information on how to reach the destination by showing the travel route. This functionality has the ability for providing trip information to the user only for the route they chose so that they are not distracted with other, unnecessary information. Driving assistance is a feature that warns the TUG in case they exceed the speed limit. It also provides them with short-term information as a reminder of road signs such as the possibility of animals on the road, close proximation to a school, etc. Functionality offers a service of memorizing short-term information and greatly helps drivers with dyslexia, dysgraphia, and color blindness which, due to the slower flow of information, do not pay much attention to these warnings. In case of incident situations, service provider informs TUG about accidents on their travel route. By providing this information, it increases the driver's attention and lowers the possibilities of secondary incidents, thus eliminating chain collisions or traffic accidents in which the TUG would potentially participate. If it's possible, the drivers are redirected to another route to prevent the secondary accidents. The relationship between TGU, system stakeholders and additional functionalities is shown in Figure 5.

10.18638/rcitd.2017.5.1.103
eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

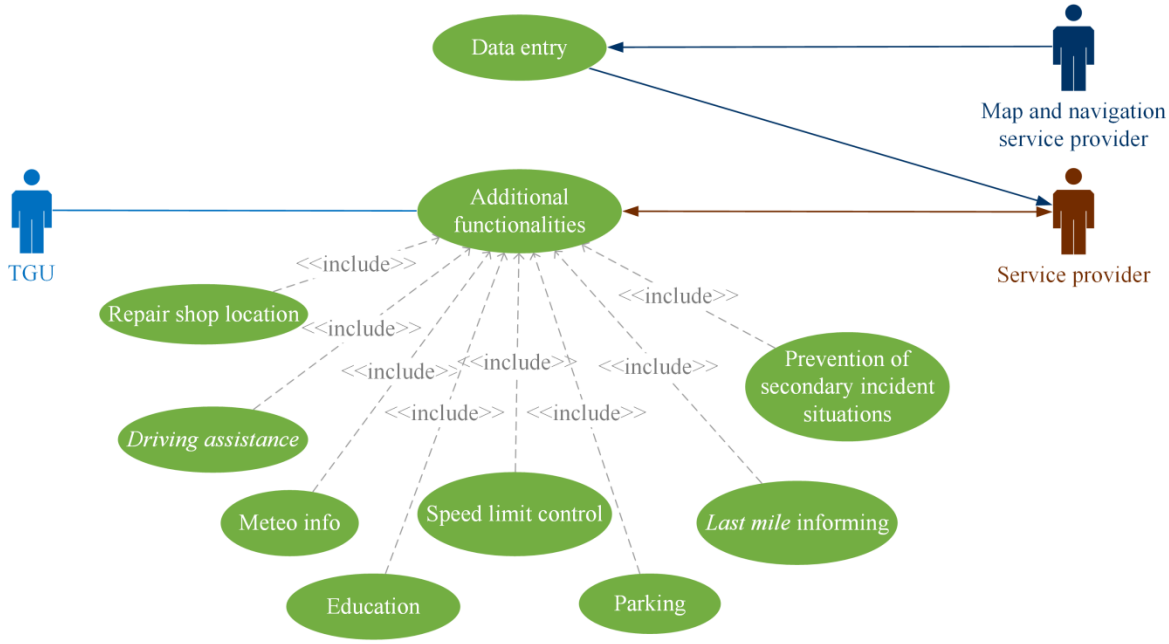


Figure 5. Relation between stakeholders and additional functionalities

When TUG logs in the application, mentioned additional functionalities are available for usage. By picking each functionality, user initiates information exchange with service provider. For some functionalities (repair shop location, parking, prevention of secondary incident situations etc.) there is a need for locating and navigating information. This information is obtained from map and navigation service provider and delivered to the user through service provider.

IV. Discussion

The conducted research has determined that TUG requirements can be met by proposal in the form of an application solution for MD. The described application solution is a part of an information and communication system that offers a service of informing users in traffic environment that is adapted for TUG. Research shows a steady growth and development of information and communication systems for informing users in traffic environment. It also shows that all available solutions with their functionalities does not meet TUG requirements and needs. It has also been pointed out that there is no TUG database, so all user requirements were obtained through a conducted survey. Parameters of simplicity, adaptation of information delivery and desirable functionalities were determined based on the results of questionnaire. They need to be taken into consideration during the development of informing system for users with reduced communication skills in traffic environment. Regarding to currently available market solutions and existing modern technologies, proposed architecture of informing system is based on CC concept. Great advantage of this informing system is the ability to adapt for all TUG needs and requirements, especially for users with multiple types of disabilities. This eliminates the need for installing multiple applications because this application solution would have all adapted information for the TUG needs.

V. Conclusion

Each TUG has different requirements and needs for information delivery in traffic environment. Because of that there is a need for creating a simple application solution that would contain functionalities that would adapt for each group requirements and needs. Currently available application solutions analyzed in this paper does not meet all TUG requirements with its functionalities, neither do they provide adapted information delivery. Due to the many forms of disabilities, the biggest disadvantage is that users need to install more application solutions. There is also a great disadvantage that there are no guidelines for designing such informing systems and that there is no available TUG database. Database like that could be used as a source for future research. Functionalities of the service are divided into basic functionalities and additional information. The implementation of this kind of application solution would increase the availability of information for all users in traffic environment (24/7). This research is a good basis for creating guidelines for informing TUG in the traffic environment. It is also the basis for future researches that aim to develop such services with the goal of raising the quality of life of users. Future research should focus on testing, developing and actual deployment of proposed application solution and system for informing TUG. Testing should concentrate on data safety, realibility and accuracy of informing system. Additionaly, the future research should focus on the interoperability of the system.

References

[1] N. Hosono, H. Inoue, M. Nakanishi, Y. Tomita, "Urgent Communication Method for Deaf, Language Dysfunction and Foreigners," Proceedings of the 14th International Conference on Computers Helping People with Special Needs, ICCHP, pp. 397-407, Paris, France, 2014.

[2] S. A. Ali, R. Jafri, "A Multimodal Tablet-Based Application for the Visually Impaired for Detecting and Recognizing Objects in a Home Environment," Proceedings of the 14th International Conference on

eISSN: 2453-6571, cdISSN: 1339-5076
10.18638/rcitd.2017.5.1.103

ISBN: 978-80-554-1375-4

Computers Helping People with Special Needs, ICCHP, pp. 356-359, Paris, France, 2014.

[3] P. Heumader, R. Koutny, K. Miesenberger, K. Kaser, "Automated Configuration of Applications for People with Specific Needs," Proceedings of the 14th International Conference on Computers Helping People with Special Needs, ICCHP (2), pp. 234-237, Paris, France, 2014.

[4] M. Hersh, "Inclusive settings, pedagogies and approaches in ICT based learning for disabled and non-disabled people: introduction to the special thematic session," Proceedings of 16th International Conference on Computers Helping People with Special Needs, ICCHP, pp. 470-474, Linz, Austria, 2016.

[5] P. Hamilton, "Communicating through Distraction: A Study of Deaf Drivers and Their Communication Style in a Driving Environment," Thesis, Rochester Institute of Technology, 2015.

[6] U. Ravid, P. Cairns, "The Design Development of a Mobile Alert Device for the Deaf and Hard of Hearing," UCL Interaction Centre, University College London, 2008.

[7] M. Hilzensauer, "Information Technology for Deaf People," Intelligent Paradigms for Assistive and Preventive Healthcare. Studies in Computational Intelligence, vol. 19, N. Ichalkaranje, A. Ichalkaranje, and L. Jain, Eds. Springer, pp. 183-206, 2006.

[8] M. Karam, "Evaluating Tactile-Acoustic Devices for Enhanced Driver Awareness and Safety: An Exploration of Tactile Perception and Response Time to Emergency Vehicle Sirens," Universal Access in Human-Computer Interaction. Aging and Assistive Environments, vol. 8515, C. Stephanidis, M. Antona, Eds., Springer, pp. 729-740, 2014.

[9] Disability Horizons (2016, June 05). Top 10 apps for disabled people. Available at: http://disabilityhorizons.com/2016/09/top-10-apps-disabled-people/

[10] The Automobile Association, Available at: http://www.theaa.com/apps

[11] Zveza Društev Gluhih in Naglušnih Slovenije (2014, February 03) Available at: http://www.zveza-gns.si/novice/slovenija/okvara-vozila-uporabite-amzs-pomoc/

[12] Hrvatski Autoklub, Available at: http://www.hak.hr/smartphone/hak http://www.hak.hr/

[13] M. Periša. "Dynamic guiding and routing of disabled and visually impaired persons in traffic", PhD Thesis. Faculty of Transport and Traffic Sciences, University of Zagreb, 2013.

[14] R. A. Amineh, A. A. B. Shirazi, "Estimation of User Location in 4G Wireless Networks Using Cooperative TDoA/RSS/TDoA Method," Fourth International Conference on Communication Systems and Network Technologies (CSNT), pp. 606-610, Bhopal, India, 2014.

10.18638/rcitd.2017.5.1.103

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

The role of Entrepreneurial Orientation towards Innovativeness, Pro-activeness and Competitive Aggressiveness in the performance of the SME's in Albania

ADELA SHESHI*, DONIKA KËRÇINI
Faculty of Economy and Agribusiness
Agricultural University of Tirana
Tirana, Albania

Abstract—**The main issue of the study is the relationship between entrepreneurial orientation and how it impacts the performance of the business in Albania. Promoting entrepreneurship and creating new businesses, are key factors of this excessive working population. Moreover, knowing that entrepreneurship, in particular (SMEs), tends to be innovative, this may help in finding solutions to several social problems of Albania, such as: economic growth and innovation, higher standards in Education system, more efficient in managing production and trade of products. This means that, pushing for a better performance, based on diversified activities, enables the enterprise to offer a unique way of joint values. This study was designed as descriptive and its data gathering method is questionnaire. The questionnaires were addressed to 100 (one-hundred) shareholders of the companies in various industries that are operating in Albania. These findings of the study will allow us to develop a model of business performance of SMEs and to test the hypotheses proposed. The study will have a positive impact because will bring important contributions such as: determining the role of each attribute of entrepreneurial orientation (innovativeness, pro-activeness and competitive aggressiveness) towards business performance of SMEs in Albania.**

Keywords- entrepreneurship, entrepreneurial orientation, and business performance

I. INTRODUCTION

The orientation of Albanian economy is more focus in the private sector particularly towards SMEs. However, still the state continues to control the main business companies, mostly those in the energy sector. After 90' property rights, freedom and corruption remain sensitive issues. Government intervention and their regulatory control continue to restrict the dynamic investments and economic efficiency in general. Despite recent reforms, inefficient business environment obstructs the development of the economy. Since the labor market is incapable of operating well, informal labor activity continues to glow.

Small and medium enterprises exert a strong influence on the economies of all countries, particularly in the fast-changing and increasingly competitive global market [5]. During these years the definition of entrepreneurial orientation had evolved a lot. The first pioneer was Mintzberg in 1973 by stating: "The entrepreneurial mode, strategy-making is dominated by the active search for new opportunitiesas well as 'dramatic leaps forward in the face of uncertainty"[6].Entrepreneurial Orientation is related to innovation strategy of a company entreprenurial oriented.

In general research made earlier on entrepreneurial orientation were focused on Miller theory. He emphasized: An entreprenurial firm is one that engages in product-market innovatioin, undertakes somewhat risky ventures, and is first to come up with proactive innovations, beating competitors too the punch"[7]. Meanwhile, in this study the focus was on Lumpkin & Dess theory becuase it is more appropriate for Albanian market.

This study is intended to identify and determine the effect of entrepreneurial orientation in the performance of SMEs in Albania. Data collection was conducted through surveys to SMEs businesses operating in Albania. The main issue of this study is the importance of entrepreneurial orientation to sustain competitiveness and higher performance. Combining empirical methods with theoretical definitions of entrepreneurial orientation, will help to determine the impact of it in business performance. The study will bring an important contributions indetermining the role of each attribute of entrepreneurial orientation (innovativeness, pro-activeness and competitive aggressiveness) towards the performance of SMEs in Albania. The paper is divided in 5 (five) sessions; session one is the introduction; session two is literature review; session three is methodology used in this; session four explain the data analysis and session five is referring to the conclusion of the study.

II. LITERATURE REVIEW

A. Entrepreneurship/Entrepreneurial Orientation

Entrepreneurship was studied by Schumpeter in 1930. According to him an entrepreneur is willing and able to convert a new idea or invention into a successful innovation [9]. Entrepreneurship is the gale of creative destruction to replace in whole or in part inferior offerings across markets and industries, simultaneously creating new products and new business models. In entrepreneurship literature, the entrepreneurial orientation is one of the most important concept, and is used to measure it [8]. Entrepreneurship is presented as a multidimensional concept an aggregate variable

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

consisting of the composite weighting of these three variables, innovatinevess risk-taking and proactiveness. Miller in his research in 1983 thinks that firms can reach high levels in two of three dimensions may be performing well but not necessarily entrepreneurial [7]. Moreover Covin and Slevin thinks that: Entrepreneurial firms are those in which the top managers have entrepreneurial management styles, as evidence by the firms'strategic decisions and operating management philosopies. Non-entreprenerial or consercative firms are those in which the top management style is decidedly risk-averse, non-innovative and pasise or reactive [1].

If a company change the technology, develop a new product or incorporate innovative processes becuase the competitors do so in most of the theories they agree that the company is not entrepreneurial oriented if the proactiveness is not always presented. However, Lumpking and Dess thinks that the processes, practices ans decisions making activities, to create new ventures, it is a strategy making process [4].

There are a lot of theories about Entrepreneurial Orientation, however, our study is focused on 5 (five) dimensions of Lumpkin &Dess [4]. These multi-dimensions are: (a) Innovativeness (b) Risk taking (c) Pro-activeness (d) Autonomy and (e) Competitive Aggressiveness. Lumpkin &Dess are stating: Entrepreneurial orientation refers to the processes, practices, and decision-making activities that lead to "new entry" characterized by one, or more of the following dimensions: "a propensity to act autonomously, a willingness to innovate and take-risks and a tendency to be aggressive toward competitors and proactive relative to marketplace opportunities". [4] The questionnaire was designed based on 3 (three) out of 5 (five) dimensions that are: innovativeness, pro-activeness and competitiveness aggressiveness.

B. Business Performance

The SMEs performance has a multi-dimensional structure; based on the combination of quantitative and qualitative variables can be measured the success of the business. Some of these attributes that measures the success are: profitability of the business; quality of the products/ services; image of the company; customer satisfaction; employee satisfaction/ motivation; efforts to innovate; number of employees; productivity of the company.

Factors that have positive impact in the performance are: motivation, education, partnership. Many authors had tried to explain the success of SMEs. One of these theories state that human capital, social capital and financial capital are essential factors in improving the performance of business. Performance is directly related to innovation. Increase business performance affects the expansion, competitiveness and chances of survival. Improved performance has an impact on the economy of the country or region [10]. Competitive organizations, constantly supporting and implementing transformative changes [2]. Organizational performance refers to ability of an enterprise to achieve such objectives as high profits, qualities product, large market share, good financial results and survival at pre-determined time using relevant strategy for action [3].

III. METHODOLOGY

The Purpose of Study is to determine the relationship between entrepreneurial orientation and business performance of small- and medium enterprise. Hypothesis of the study proposed are:

H1: *Entrepreneurial Orientation towards innovativeness affect positively the business performance (SME's).*

H2: *Entrepreneurial Orientation towards pro-activeness affect positively the business performance (SME's).*

H3: *Entrepreneurial Orientation towards competitive aggressiveness affect positively the business performance (SME's).*

In this research was used the quantitative data collection methods. Data collection was conducted during the period April - June 2016. The collection process began with the identification of potential candidates in Albanian business environmental. The number of valid questionnaires were 100 (one-hundred). The data collection was done by searching for information about the business and the collection of information through direct survey.

The questionnaire was segmented into 2 (two) core parts:

1. Measuring 3 (three) out of (five) dimensions of entrepreneurial orientation; the independent variables are: a. innovativeness; b. pro-activeness; c. competitive aggressiveness (using 5- point Likert scales)
2. the performance of the business within market- is measured by 8 factors (profitability of the business, quality of the products/ services, image of the company, customer satisfaction, employee satisfaction/ motivation, efforts to innovate, number of employees, productivity of the company) (using 5- point Likert scales).

In addition, the questionnaire was collected data for the type of industry, number of employees, number of years operating in the market and annual sales.

IV. DATA ANALYSIS

The sample is limited to 100 surveys and in this study was used Multi- Regression Analysis to test proposed hypothesis. This approach is designed to develop a regression model with the fewest number of variables which are statistically independent. Analysis of Variance (ANOVA) was used to determine the differences between the models of leadership and performance. Was used the SPSS statistical program to analyze the data based on multiple regression analysis to test the proposed hypothesis. To measure the reliability of a model and significant that the components / questions is used Cronbach alpha level. If Cronbach Alpha values are greater than 0.70 are considered acceptable. In this study the level of reliable values are above 0.70.

Kaiser-Meyer-Test Olkin (KMO) measure the adequacy of the data collected, it should be greater than 0.5. This correlation indicates that the models are relatively compact and analysis factor should provide adequate and reliable factors the result is showed in Table 1.

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

TABLE I. VALIDITY AND RELIABILITY (KMO AND BARTLETT'S TEST)

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.567
Bartlett's Test of Sphericity	Approx. Chi-Square	184.116
	Df	21
	Sig.	.000

Factor analysis method is applied to determine the size of a scale. Factor analysis is a technique to identify groups or groups of variables and to understand the structure of any hidden variable. Factors will be presented in the form of linear equation (1) to explain the variables measured and their importance as factors.

$$Pkk = a + b1OSak + b2OSi + b3OSp \quad (1)$$

Note:

Pkk= Business Performance

OSak= Entrepreneurial Orientation toward Competitive aggressiveness

OSi= Entrepreneurial Orientation toward Innovativeness

OSp= Entrepreneurial Orientation toward Pro-Activeness

The factor analysis method had identified three main dependent variables as it is showed in Table 2, and other variables' items and statements as it is showed in Table 3.

Using SPSS regression model to test the correlation between each Entrepreneurial Orientation (EO) models and the business performance of SMEs in Albania. It was first investigated the relationship between the model of EO towards competitive aggressiveness and performance. In Table 6 the output shows the results of regressive analysis performed with SPSS regression model where the dependent variable is as performance and as the independent variable is EO towards competitive aggressiveness.

TABLE II. TOTAL VARIANCE/ PRINCIPAL COMPONENT ANALYSIS

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.423	34.609	34.609	2.423	34.609	34.609	2.056	29.368	29.368
2	1.544	22.063	56.672	1.544	22.063	56.672	1.688	24.115	53.483
3	1.264	18.052	74.724	1.264	18.052	74.724	1.487	21.241	74.724
4	.618	8.826	83.550						
5	.465	6.646	90.196						
6	.443	6.330	96.526						
7	.243	3.474	100.000						

TABLE III. VARIABLES AND STATEMENTS

Statement	Components/ Variables		
	Competitive Aggressiveness	Innovativeness	Pro-activeness
OSKA2 Overall leaders in my company operate: by choosing risky way to potential opportunities.	.894		
OSKA3 Generally the leaders in my company when take decisions that are ambiguous they are: in decision making the attitude iscourageous and aggressive.	.862		
OSKA1 Overall leaders in my company tend: to favor high-risk projects.	.684		

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

OSI2 Main focus of my company is: developing new products or services.		.872	
OSI1 In general, my company executives favor: research & development and leadership towards technological innovation.		.842	
OSP1 My company in relation to competitors: takes initiatives to respond to competitors, or react later.			.902
OSP2 My company relative to its competitors: They are often the first to introduce new products, services, or other administrative techniques and technologies.			.782

TABLE IV. MODEL OF EO TOWARDS COMPETITIVE AGGRESSIVENESS

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.097[a]	.010	-.001	4.782

a. Predictors: (Constant), EO towards Competitive Aggressiveness

TABLE V. CORRELATION ANALYSIS (ANOVA)REGARDING DEPENDENT VARIABLE PERFORMANCE AND INDEPENDENT VARIABLE EO COMPETITIVE AGGRESSIVENESS.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	21.501	1	21.501	.940	.335[b]
	Residual	2240.939	98	22.867		
	Total	2262.440	99			

a. Dependent Variable: Performance; b. Predictors: (Constant), EO towards Competitive Aggressiveness

TABLE VI. REGRESSION ANALYSIS REGARDING EO TOWARDS COMPETITIVE AGGRESSIVENESS

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	28.334	1.449		19.554	.000
	Competitive Aggressiveness	.140	.144	.097	.970	.335

a. Dependent Variable: Performance

From the ANOVA Table 5, as p = .024 <α = .05, we can say that: "we can say that: "There is no a significant correlation between the model of EO towards competitive aggressiveness and performance of SMEs businesses in Albania.

Moreover, it was investigated the relationship between the model of EO towards innovativeness and performance. Output in Table 9 shows the results of regressive analysis performed with SPSS regression model where the dependent variable is as performance and as the independent variable is EO towards innovativeness.

From the ANOVA Table 8, as p = .024 <α = .05, we can say that: "There is a significant correlation between the model of EO towards innovativeness and performance of SMEs businesses in Albania".

To conclude, it was investigated the relationship between the model EO towards pro-activeness and performance. Output

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

in Table 12 shows the results of regressive analysis performed with SPSS regression model where the dependent variable is as performance and as the independent variable is EO towards pro-activeness.

TABLE VII. MODEL OF EO TOWARDS INNOVATIVENESS

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.453[a]	.205	.197	4.284

a. Predictors: (Constant), EO towards Innovativeness

TABLE VIII. CORRELATION ANALYSIS (ANOVA)REGARDING DEPENDENT VARIABLE PERFORMANCE AND INDEPENDENT VARIABLE EO TOWARDS INNOVATIVENESS.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	464.037	1	464.037	25.287	.000[b]
	Residual	1798.403	98	18.351		
	Total	2262.440	99			

a. Dependent Variable: Performance; b. Predictors: (Constant), EO towards Innovativeness.

TABLE IX. REGRESSION ANALYSIS REGARDING EO TOWARDS INNOVATIVENESS.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	22.453	1.496		15.010	.000
	Innovativeness	1.028	.204	.453	5.029	.000

a. Dependent Variable: Performance

TABLE X. MODEL OF EO TOWARDS PRO-ACTIVENESS.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.378[a]	.143	.134	4.448

a. Predictors: (Constant), EO towards Pro-activeness.

TABLE XI. CORRELATION ANALYSIS (ANOVA)REGARDING DEPENDENT VARIABLE PERFORMANCE AND INDEPENDENT VARIABLE EO TOWARDS PRO-ACTIVENESS.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	323.772	1	323.772	16.367	.000[b]
	Residual	1938.668	98	19.782		
	Total	2262.440	99			

a. Dependent Variable: Performance

b. Predictors: (Constant), EO towards Pro-activeness.

TABLE XII. REGRESSION ANALYSIS REGARDING EO TOWARDS PRO-ACTIVENESS.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	22.052	1.932		11.411	.000
	Pro-activeness	1.017	.251	.378	4.046	.000

Dependent Variable: Performance

From the ANOVA Table 11, as p = .024 <α = .05, we can say that: "There is significant correlation between the model of and performance of SMEs businesses in Albania".

The correlation between the model of EO towards competitive aggressiveness and performance is negative, which this issue has been emphasis on previous studies. The examination of entrepreneurial orientation was done by evaluating separately the dimensions of: competitive aggressiveness, innovativeness and pro-activeness. The result showed that 2 (two) our of 3 (three) dimensions were positively related to business performance. In terms of impacing the business performance the competitive aggressiveness was not statistically significant. As Lumpking and Dess stated: company's leaders that are willing to improve competitiveness must act anticipating future problems, needs or change that is proactiveness [4]. In order to perform better the companies should motivate and support the new ideas.

V. CONCLUSIONS

In this study concepts likeentrepreneurial orientation towards competitive aggressiveness, innovativeness and pro-activeness are very important regarding the performance of the business. Analysis was conducted on three types of entrepreneurial orientation models as competitive aggressiveness, innovativeness and pro-activeness.

In this research was reflected the literature debate regarding the dimensions of entrepreneurial orientation. Some scholars agreed that EO construct is one-dimensional concept and some other agreed that different EO dimensions combinations not necessarily needs to be all presence to support EO.

In conclusion, there is a lack of information in this specific area in Albania, so these findings can be used as guidelines in practice by business owners, managers, directors in improving the performance of the company and being more competitive in the market. In additional, this study can help the policymaker to look for new approaches or update the old ones so SMEs can improve more their performance and being competitive.

REFERENCES

[1] J. Covin, & D. Slevin, "Strategic Management of Small Firms in hostile & benign environments", Strategic Management Journal, p. 75-87, 1989.

[2] M. Cohen, "Commentary on the organization science special issue on complexity",Organization Science, p.373-376, 1999.

[3] H. Koontz, C. Donnell, "Introduction to Management", New York: McGraw-Hill Inc: 1993.

[4] G.T. Lumpkin, G.G. Dess, "Clarifying the entrepreneurial orientation construct and linking it to performance", Academy of Management Review, 21 (1), p.135-172, 1996.

[5] W.M. Ladzani, and J.J. Vuuren, "Entrepreneurship training for emerging SMEs in South Africa", Journal of Small Business Management, 40 (2), p. 154-161, 2002.

[6] H. Mintzberg, "The nature of managerial work", New York: Haper& Row, 1973.

[7] D. Miller, "The correlates of entrepreneurship in three types of firms", Management Science, 29, p. 770-791, 1983.

[8] M.H. Morris, and D.F. Muratko, "Corporate Entrepreneurship", Fort Worth, TX: Harcourt College Publishers, 2002.

[9] J. Schumpeter, "Capitalism, socialism, and democracy", New York: Harper and Brothers, 1942.

[10] J. Wolff,T. L. Pett, "Small firm performance modeling the role of product and process improvement",Journal of Small Business Management; p. 268-284, 2006.

[11] World Trade Organization, "Levelling the trading field for SMEs",World TradeOrganization,2016.

Exploring Digital Divide in Mobile Phone Ownership: Evidence from Nigeria

Ivan Forenbacher, Siniša Husnjak, Ivan Cvitić
Department of Information and Communication Traffic
Faculty of Transport and Traffic Sciences
University of Zagreb
Zagreb, CROATIA
ivan.forenbacher@fpz.hr, sinisa.husnjak@fpz.hr, ivan.cvitic@fpz.hr

Abstract—**In 2012, as part of its response to mobile digital divide, the Nigerian government introduced mobile phone subsidies to rural farmers and did not seem to have achieved the expected results. It was argued that no necessary analysis of socio-economic factors affecting mobile phone ownership had been conducted previously. The paper took advantage of this period to examine socio-economic factors affecting the probability of mobile phone ownership in Nigeria. In order to estimate a logit model we used national representative data from *DataFirst* on households and individuals ICT access and usage in 2011-2012. In contrast to what had previously been thought, we found out that poverty may not be correlated with the probability of owning a mobile phone. Type of electricity source, education, and activity has the greatest effect on the probability of owning a mobile phone. The findings may help to improve more coordinated digital divide policy, and serve as a complement to ICT Roadmap 2017 to 2020 in Nigeria and similar countries.**

Keywords- Mobile, Digital Divide, Information Communication Technology (ICT), Nigeria, Africa

I. INTRODUCTION AND LITERATURE REVIEW

Digital divide today is a well-known phenomenon and reflects the differences in accessing and using Information Communication Technology (ICT). The understanding of digital divide is not comprehensive and is present in both developed and developing countries, including the ones on the African continent [1]. One option to bridge digital divide is through mobile phones. In particular, they can enhance pro-poor development in various important sectors, including health, education, agriculture, and employment and bridge digital divide [2], [3].

Africa has the largest and fastest growing base of mobile phone users [4]. Lots of users, due to cost-effectiveness, leapfrog the traditional technologies like computers and fixed-broadband, and accept wireless technologies [5]. This implies formulation of a coherent and coordinated ICT policy at the national, regional and local level by including a strategic framework.

Performing such analysis for the African continent during the past has had certain difficulties. This is primarily because the quality and credible data from trustworthy sources was rarely available for developing African countries, thus making it harder to perform necessary analysis. Without the right approach and in-depth analysis it may result in inefficient policies and methods for confronting development challenges, including digital divide.

This actually occurred in Nigeria in 2012 as a part of its response to bridge mobile urban/rural digital divide [5]. The government subsidized the mobile phones to rural farmers through Growth Enhancement Support (GES), despite the fact that they didn't prove any *ex-ante* causation between poverty and mobile phone ownership [6]. Moreover, a farmer will not buy mobile phone if he does not see a direct benefit from it, for example, through the use of application for the purchase of seed [6].

One explanation is that policy makers should take a more holistic approach. Instead of focusing just on mobile phone ownership (first-order digital divide), they should also focus on the use of technologies, such as Internet (second-order digital divide) [7]. As a result, much remains to be done to support further growth of mobile markets and services in Nigeria by building integral ICT policy, especially in the advent of the country's ICT Roadmap 2017 to 2020.

At the moment, despite various researches on digital divide in African countries, including Rwanda, Ghana, Uganda, Gabon, Kenya, Somalia, Cameroon, and Ethiopia [8-13], to the best of authors' knowledge, there is lack of research examining digital divide in Nigeria as the largest African economy [14].

Recent findings on mobile digital divide in Nigeria are only based on various consulting reports, stating basic digital divide indicators. Authors in [14] claim that since 2004, the number of mobile subscribers has been growing linearly slow. First significant increase of around 50 million subscribers occurred in the period from 2006 to 2010. It still remains unclear to what extent particular socio-economic factors may have led to such increase.

This suggests the need to gain knowledge about particular socio-economic factors and their magnitude of effect on Nigerian mobile subscriber probability to adopt a mobile phone as the first step. This question is important to investigate because understanding about the reasons behind proliferation of mobile subscribers can lead to efficient future policies.

The present paper uses national representative dataset from [15] on household ICT access and usage from 2011-2012 to estimate logit model which had been previously used by [11] or [18] for measuring digital divide. This, in turn, allowed us to

perform initial analysis of the effect of socio-economic factors on the probability of owning a mobile phone in Nigeria at the population level. The results may have important implications for integral digital divide policies in Nigeria and may complement other studies dealing with the use of various technologies and services in other similar countries.

The rest of the paper is organized as follows. In Section 2 the data and methodology used are described. In section 3 the main results are presented. In Section 4 we discuss the results. Finally, in Section 5 we summarize the results and conclude.

II. Data And Methodology

A. Data sources

We had at our disposal dataset from *DataFirst*, South Africa [15]. The data were collected by Research ICT Africa (RIA) as a part of RIA Household and Small Business Access and Usage Survey 2011-2012 [16].

Data collecting mode was face-to-face. Units of the analysis used in the survey were both households and individuals aged 15 or older.

Sampling procedure consisted of random sampling which was performed in five steps for individuals [16]:

- Step 1: The national census sample frame was stratified to urban and rural Enumerated Areas (EA).
- Step 2: EAs were then sampled for each stratum using Probability Proportional to Size (PPS).
- Step 3: For each EA, lists that served as a sample frame for simple random sections were created.
- Step 4: 24 households were sampled using a simple random sample for each selected EA.
- Step 5: One individual was randomly selected based on a simple random sampling from all of household members or visitors staying in the household at the moment survey was performed, and aged 15 years or older.

Sample consisted of total 1552 observations for Nigeria. Therefore, the sample was weighted in order to gross up the data to the population level. The weights were calculated for households and individuals. They were based on the inverse selection probabilities. Applying weights during calculation ensures grossing up the sample of 1552 observations to 90 595 137 individuals which makes the data national representative of the target population aged 15 years or older.

B. Relevant variables to be used in the model

Relevant variables reflecting various socio-economic factors (Table 1) were adopted from previous research, such as [11] and [17] which examined factors affecting probability of mobile phone ownership.

TABLE I. Relevant Socio-Economic Factors Used to Generate Logit Model

Concept	Socio-economic factor	Variable	Description
Digital divide	Mobile phone ownership	mpo	Mobile phone owner (1 = yes, 0 = No)
Geographic area	Location	rural	Rural area (1 = yes, 0 = no)
Quality of life	Electrical infrastructure	main electricity grid	Household is connected to main electricity grid (1 = yes, 0 = no)
		generator	Household has generator (1 = yes, 0 = no)
		other	Household uses other electricity sources, e.g. solar (1 = yes, 0 = no)
Personal characteristics	Gender	female	Individual is a female (1 = yes, 0 = no)
	Age	age	Number of years of an individual
Education status	Schooling	primary	Individual has primary degree (1 = yes, 0 = no)
		secondary	Individual has secondary degree (1 = yes, 0 = no)
		tertiary: diploma /certificate	Individual has diploma degree (1 = yes, 0 = no)
		tertiary: bsc/ba	Individual has BSc/Ba degree (1 = yes, 0 = no)
Activity	Activity last 6 months	unpaid house work	Individual is at home, e.g. housewife (1 = yes, 0 = no)
		retired	Individual is retired (1 = yes, 0 = no)
		unemployed	Individual is unemployed (1 = yes, 0 = no)
		disabled and unable to work	Individual is disabled and unable to work (1 = yes, 0 = no)
		employed	Individual is employed (1 = yes, 0 = no)
		self-employed	Individual is self-employed (1 = yes, 0 = no)
Literacy	Reading skills	with difficulty	Reading difficulties (1 = yes, 0 = no)
		not at all	Doesn't know to read (1 = yes, 0 = no)
	Writing skills	with difficulty	Writing difficulties (1 = yes, 0 = no)
		not at all	Doesn't know to write (1 = yes, 0 = no)
	English reading / writing skills	eng_yes	Knows to read and write in English (1 = yes, 0 = no)
Economic status	Earnings	disposable income	Amount of money at disposal each month (1 = yes, 0 = no)

The dependent variable is *mpo* (Mobile Phone Ownership). This is a dichotomous variable indicating whether individual owns a mobile phone (*mpo = 1*) or not (*mpo = 0*).

The explanatory variables were split in two groups – ratio and categorical. The first ratio variable included in the model is *age*. In contrast to other studies [11] and [17], we have not categorized the variable *age*. The second ratio variable is *disposable income* which refers to the amount of money an individual had at free disposal each month.

Categorical variables referred to the type of household location (rural or urban); type of electricity (no electricity, main electricity grid, generator, and other, e.g. solar); gender (male or female); highest level of schooling completed (none, primary, secondary, tertiary with diploma, tertiary with BSc/Ba degree); main activity during last six months (student/pupil, unpaid housework, e.g. housewife, retired, unemployed, disabled and unable to work, employed, and self-employed);

reading and writing skills in the mother tongue and English (easily, with difficulties, and not at all).

Table 1 and table 2 present basic descriptive statistics of the variables used in the model for Nigerian market.

TABLE II. NATIONAL REPRESENTATIVE DESCRIPTIVE STATISTICS FOR CATEGORICAL VARIABLES

Socio-economic factor	Variable	Proportion		
Mobile Phone Ownership (MPO), D. V.	no	0.3363		
	yes	0.6637		
Location	rural	0.498		
	urban	0.502		
Electrical infrastructure	no	0.2993		
	main electricity grid	0.5747		
	generator	0.1025		
	other	0.0235		
Gender	male	0.5315		
	female	0.4385		
Schooling	none	0.2872		
	primary	0.1867		
	secondary	0.3777		
	tertiary: diploma certif.	0.0972		
	tertiary: bsc/ba	0.0512		
Activity	student/pupil	0.1548		
	unpaid housework	0.2093		
	retired	0.0116		
	unemployed	0.0603		
	disabled - unabled to work	0.0016		
	employed	0.1469		
	self-employed	0.4156		
Reading skills	easily	0.4619		
	with difficulties	0.2108		
	not at all	0.3273		
Writing skills	easily	0.473		
	with difficulties	0.2014		
	not at all	0.3256		
English reading/writing skills	eng_yes	0.5171		
	eng_no	0.4829		
Number of strata	2	**Number of obs.**	1552	
Number of PSUs	63	**Population size**	90 595 137	
		Design df	61	

TABLE III. NATIONAL REPRESENTATIVE DESCRIPTIVE STATISTICS FOR RATIO VARIABLES

Variable	Mean	Std. error	[95% Conf. Interval]		Min	Max
age	34.263	0.646	32.970	35.556	15	99
disposable income	6062.294	681.510	4699.53	7425.058	0	200000
Number of strata		2	**Number of obs.**		1552	
Number of PSUs		63	**Population size**		90 595 137	
			Design df		61	

C. *The empirical model*

The econometric binary logit model to be estimated can be derived from the latent variable model:

$$mpo_i = \begin{cases} 1 & y_i^* = \alpha + \beta x + \varepsilon > 0 \\ 0 & otherwise \end{cases}. \quad (1)$$

where y_i^* is unobserved, latent variable indicating utility of owning a mobile phone. If y_i^* is greater than zero, meaning there is positive usefulness from mobile phone for observation i, we get to observe an individual i owning a mobile phone, i.e. $mpo_i = 1$, and $mpo_i = 0$ otherwise. The mpo is, therefore, our limited dependent variable (LDV). The error term ε is unobserved and distributed by the standard logistic distribution.

The α is a constant term and βx refers to a vector of m explanatory variables presented in Table 1 and Table 2 for a single observation i from a sample of total n observations, denoted as x_i with corresponding coefficient β_k. This can be rewritten as $\beta x = \beta_1 x_{ik} + \cdots + \beta_k x_{ik}$ for $k = 1,2,\ldots,m$. From this, the primary interest is to determine probability that

$$Prob(mpo_i = 1|x) = Prob(mpo_i = 1|x_1, x_2, \cdots x_k) \quad (2)$$

To avoid limitations of the linear probability model (please see [17] for details), we must consider binary response model

$$Prob(mpo_i = 1|x) = \mathrm{L}(\alpha + \beta x) = L(z) \quad (3)$$

where z is a linear function of our explanatory variables and L is a non-linear function taking values in the range $0 < L(z) < 1$, for all real numbers z. In our case the non-linear function L is from the family of logistic functions defined as follows:

$$Prob(mpo_i = 1|x) = L(z) = \frac{\exp(z)}{1+\exp(z)} = \Lambda(z) \quad (4)$$

which lies in the interval between 0 and 1 for all real numbers z. $\Lambda(z)$ is a cumulative logistic function for a standard logistic random variable. The Equation (4) can be interpreted as the probability of *mpo* equaling "1" (success or owning a mobile phone). From this we can now derive the inverse of the logistic function $\Lambda(z)$ *logit* or *ln (odds)* to get the linear expression:

$$logit[\Lambda(z)] = ln\left(\frac{\Lambda(z)}{1-\Lambda(z)}\right) = \alpha + \beta x. \quad (5)$$

By rewriting Equation (5), we get our final econometric logit model to be estimated

$$\ln(odds)_i = \alpha + \beta_k x_{ik} + \cdots + \beta_k x_{ik}, \quad k = 0,1,\ldots,m \;\; i = 0,1,\ldots,n\,. \quad (6)$$

Equation (6) states that dependent variable (DV) refers to the logit of mobile phone ownership for a particular observation i in the sample. The coefficient β_k measures *ceteris paribus* effect of one-unit change of x_{ik} on the DV. Predicted probabilities can be calculated by using equation (4).

III. RESULTS

Our strategy was to perform logit model across whole sample for Nigerian market to determine the coefficients values. Coefficients of the logit model were estimated with Maximum Likelihood Estimation (MLE) using described equations (1)-(6) from our statistical software. Table 3 presents the main results.

Final model consisted of total 22 variables and a constant term. The weighted analysis was performed on total of 1552 observations.

It is important to note that, since our dependent variable *mpo* is at the individual level, we had to use individual weights in our calculations in order to gross up the data to national level. This resulted with the target population of 90 595 137 individuals which were 15 years old or older in 2011-2012.

The critical *F value* is 16.79 with (22, 40) degrees of freedom. Probability of observing as extreme or more extreme F value is 0.000, given that the null hypothesis is true.

Main results suggest that most variables are statistically significant based on *p* values either at 1%, 5% or 10% level. However, this is not the case with certain variables. For example, variable *location* (urban/rural) is not statistically significant ($p = 0.865$) and has no effect on probability of mobile phone ownership. Similarly, other electricity sources (e. g. solar) does not contribute to the model significantly ($p = 0.958$).

TABLE IV. ESTIMATED LOGIT MODEL FOR NIGERIAN MARKET

MODEL	Coef.	Lin. Std. Err.	t	P>t	[95% Conf.	Interval]
intercept	-1.291	0.708	-1.82	0.073	-2.708	-0.126
rural	-0.069	0.407	-0.17	0.865	-0.884	0.745
main electricity grid	0.850	0.329	2.58	0.012	0.190	1.509
generator	1.817	0.434	4.18	0.000	0.947	2.687
other	-0.028	0.536	-0.05	0.958	-1.100	1.044
female	-0.534	0.273	-1.95	0.055	-1.080	0.012
age	-0.017	0.008	-2.04	0.046	-0.034	-0.0003
primary	0.808	0.496	1.64	0.109	-0.184	1.801
secondary	1.720	0.542	3.17	0.002	0.636	2.804
tertiary: diploma /certificate	1.803	0.471	3.83	0.000	0.860	2.745
tertiary: bsc/ba	1.879	0.934	2.01	0.049	-0.010	3.747
unpaid housework	0.993	0.385	2.57	0.012	0.222	1.765
Retired	2.289	0.860	2.66	0.010	0.569	4.009
unemployed	1.048	0.518	2.02	0.047	0.012	2.085
disabled and unable to work	0.708	0.977	0.72	0.472	-1.247	2.663
employed	1.436	0.424	3.38	0.001	0.587	2.284
self-employed	1.179	0.314	3.75	0.000	0.550	1.808
reading_with difficulty	0.822	0.512	1.61	0.113	-0.201	1.846
reading_not at all	0.851	0.866	0.98	0.330	-0.880	2.583
writing_with difficulty	-0.088	0.421	-0.21	0.835	-0.930	0.754
writing_not at all	-1.763	0.829	-2.13	0.038	-3.422	-0.105
eng_yes	0.430	0.433	0.99	0.324	-0.436	1.298
disposable income	0.0008	0.00001	4.52	0.000	0.00004	0.0001
Model summary						
Number of strata	2		**Numb. Of obs.**	1552		
Number of PSU	63		**Population size**	90 595 137		
			Design df	61		
			F (22, 40)	16.79		
			Prob > F	0.000		

In addition, variables regarding literacy (reading with difficulties, reading not at all, writing with difficulties, and English reading and writing skills) and are not statistically significant.

Considering other statistical significant variables, there is a gender gap in mobile phone ownership with *female*'s coefficient of -0.534 meaning they are less likely than males to own a mobile phone.

With respect to education, the higher degree an individual has, they are more likely to own a mobile phone. Interestingly, results report that monthly disposable income does not contribute significantly to the propensity of owning a mobile phone.

Furthermore, results report that *age* (-0.017) doesn't have significant effect on probability of mobile phone ownership. From the group of variables regarding main activity during last six months, *retired* people are more likely (2.289) to have mobile phones than student/pupils. They are followed by *employed* (1.4363) and *self-employed* (1.179) individuals.

IV. DISCUSSION

Some previous digital divide policies in Nigeria, such as *Growth Enhancement Support* (GES) when government introduced mobile phone subsidy to farmers in 2012 may not yield desired results. It was argued by [6] that they should first explore main socio-economic factors responsible for mobile phone ownership and then decide when and to whom to subsidize mobile phone. Identification of these factors may, in turn, facilitate the development of more coherent policies.

In this study we explored the impact of socio-economic factors on the probability of mobile phone ownership in Nigeria. We used national representative dataset from the *DataFirst*, based on RIA Africa ICT access and usage survey for period 2011-2012 [15], [16].

We used a binary logit model that had been used in previous research for similar purposes, such as [11] and [18]. The model's "quality" indicators suggest that the model shows good fit to the data. Additionally, the type of variables included in the model is justified. The variables representing socio-economic factors included in the model were adopted from [11] and [18] and most of them are statistically significant.

Another important aspect is the ratio of the number of observations and number of variables included in the model. The literature suggests that the minimum number of observations per variable for logit model should be 10 – 20 in order to achieve empirical validity [19]. This was indeed the case with our model resulting with approximately 67 observations per each variable.

The direction and magnitude of estimated coefficients seems reasonable. Importantly, a variable *location* denoting geographic area of residence for an individual (rural or urban) is statistically significant and does not contribute to the model [6]. Namely, this is consistent with the theory from analysis [6] arguing that the farmer will not buy a mobile phone with no obvious direct benefit. In other words, the farmer will not

spend a certain amount of money to buy a mobile phone and on recurring costs that follow if the benefit doesn't outweigh the cost. Therefore, poverty cannot be taken exclusively as a cause of not owning a mobile phone in Nigeria.

Results suggest presence of gender gap in mobile phone ownership. Nigerian females are less likely to own mobile phones than males. Such results are consistent with the studies [11], [20], [21], and [22]. These studies argue that mostly men are early-adopters when new technology is being introduced. Additionally, gender gap decreases eventually as technology becomes more and more prevalent.

Another important socio-economic factor affecting the probability of mobile phone ownership is *age*. The results show that age has negative and mild effect on the *log odds* of mobile phone ownership. Namely, several studies [11], [23], and [24] suggest that the sign of correlation between odds of mobile phone ownership and age should be positive. This is something that can be implied from the variables referring to the main activity last 6 months. Coefficients values suggest that retired are more likely to own mobile phones than the youngest - pupils and students.

One more important socio-economic factor to consider is education, and its certain aspects deserve attention. Our results have confirmed previous claims from [11] that individuals with higher education degree are more likely to own a mobile phone. This can be interpreted by the fact that if someone is more educated, they will have less training costs and will be able to see the benefits of having a mobile phone more quickly [11].

Variable *income* suggests an interesting thing. Despite its positive sign which seems reasonable, the magnitude of its effect on the log odds of owning a mobile phone is very moderate. This can be interpreted from three aspects. First, in our research, we did not use a variable that would refer to the total monthly income but exclusively to the monthly amount which remains available to the individual. Second, it could be argued that the effect of variable *income* is moderate because monthly amount available is insufficient to own a cell phone. Thirdly, there is always a possibility that a person is moonlighting and receiving a salary, thus not reporting it in the survey.

Socio-economic factor *English reading/writing skills* suggest that knowing how to read and write in English is not statistically significant. In other works, log odds of owning a mobile phone and English literacy are not correlated. This aligns with similar study conducted in Gabon [11] where authors obtained statistical insignificance of the same variable. This may suggest that there isn't enough content in English that would attract potential mobile phone owners.

Similarly, if individuals do not know how to write it is very unlikely that they will own mobile phones. This may suggest the need to work harder on writing skills which are inevitable when using a mobile phone.

V. CONCLUSION

Careful *ex-ante* evaluation of socio-economic factors affecting the mobile phone ownership can be used to improve effectiveness of digital divide policies. This is because, as shown, certain socio-economic factors may have different effects from those that seemed reasonable at first and may be specific to individual countries, such as Nigeria.

As our results demonstrate, urban or rural *location* is not correlated with the probability of mobile phone ownership, as well as other electricity sources (e.g. solar), disability and inability to work, literacy problems. On the other hand, the main aspects of socio-economic factors which stimulate the most mobile phone ownership in Nigeria are housings that are connected to the main electrical grid or have a generator, individuals with tertiary level of education, retired, employed, and self-employed.

Further research is needed to avoid inefficient policies. An important issue to resolve for future studies is not only socio-economic factors affecting the first-order digital divide, but also the second-order digital divide. Another aspect would be to repeat such analysis with newer data set. This, in turn, would enable to analyze if certain policies yielded desired results and measure evolution of individual's behavior over time.

Our results may pave the road to more extensive studies in the future. The results may help authorities and policy makers to make coherent and efficient ICT policies and strategies in Nigeria and other similar countries, especially on African continent.

ACKNOWLEDGMENT

Present paper has made use of data "Africa - RIA Household and Small Business Access and Usage Survey 2011-2012", produced by *Research ICT Africa* and provided by *DataFirst*, Cape Town, South Africa.

REFERENCES

[1] L. Hosman, E. Fife, "The use of mobile phones for development in Africa: top-down-meets-bottom-up partnering," The Journal of Community Informatics, vol. 8(23), 2012.

[2] United Nations. (2012). *Annual report 2011/2012*. Available at: http://www.undp.org/content/dam/undp/library/corporate/UNDP-in-action/2012/English/UNDP-AnnualReport_ENGLISH.pdf

[3] H. Geser, (2004). *Towards a sociological theory of the mobile phone*. Available at: http://socio.ch/mobile/t_geser1.pdf

[4] J. Patterson, (2016). *Africa's mobile subscriptions grow fastest globally*. Available at: http://www.thisisafricaonline.com/News/Africa-s-mobile-subscriptions-grow-fastest-globally?ct=true?utm_campaign=TIA+e-newsletter+March+2016+1st+issue+-+post-download+version&utm_source=emailCampaign&utm_medium=email&utm_content=

[5] M. Ragnedda, G. W. Muschert, "The Digital Divide: The Internet and social inequality in international perspective (Routledge Advances in Sociology)", 1st edition, 2015.

[6] E. Ojameruaye (2013, January, 28). *A commentary of mobile šhones subsidy for poor rural farmers in Nigeria*. Available at: http://chatafrik.com/articles/nigerian-affairs/a-commentary-of-mobile-phones-subsidy-for-poor-rural-farmers-in-nigeria#.VtAUtOY70kl

[7] E. Hargittai, "Second-Level Digital Divide: Differences in People's Online Skills," First Monday, vol. 7(4), 2002.

[8] N. Mumporeze, M. Prieler, "Gender digital divide in Rwanda: A qualitative analysis of socioeconomic factors," Telematics and Informatics, doi: https://doi.org/10.1016/j.tele.2017.05.014, in press 2017.

[9] C. Fuchs, E. Horak, "Africa and the digital divide," Telematics and Informatics, vol. 25(2), pp. 99-116, 2008.

[10] M. Kasusse, "Bridging the digital divide in Sub-Saharan Africa: The rural challenge in Uganda," The International Information & Library Review, vol. 37(3), pp. 147-158, 2005.

[11] T. Penard, N. Poussing, G. Zomo Yebe, P. Nsi Ella, "Comparing the determinants of Internet and cell phone use in Africa: Evidence from Gabon," Digiworld Economic Journal, no. 86, pp. 65-83, 2012.

[12] I. Brännström, "Gender and digital divide 2000–2008 in two low-income economies in Sub-Saharan Africa: Kenya and Somalia in official statistics," Government Information Quarterly, vol. 29(1), pp. 60-67, 2012.

[13] A. Chabossou, C. Stork, M. Zahonogo, "Mobile Telephony Access and Usage in Africa," Proceeding ICTD'09, Proceedings of the 3rd international conference on Information and communication technologies and development, pp. 392-405, Doha, Qatar — April 17 - 19, 2009.

[14] GSMA (2013). *Sub-Saharan Africa mobile economy 2013*. Available at: https://www.gsma.com/mobileeconomy/archive/GSMA_ME_SubSaharanAfrica_2013.pdf

[15] Statistics South Africa, "Africa - RIA Household and Small Business Access and Usage Survey 2011-2012", [dataset]. Pretoria. Statistics South Africa [producer], 2012. Cape Town. DataFirst [distributor], 2012.

[16] Research ICT Africa. Household and Small Business Access & Usage Survey 2011. Available at: http://www.researchictafrica.net/docs/Survey%20Methodology%202011:12.pdf

[17] J. M. Wooldridge, "Introductory Econometrics: A Modern Approach,", 6th Edition, 2015.

[18] M. C. Tran, A. B. Labrique, S. Mehra, H. Ali, S. Shaikh, M. Mitra, P. Christian, K. Jr West, "Analyzing the mobile "Digital Divide": Changing determinants of household phone ownership over time in rural Bangladesh," JMIR Mhealth Uhealth, vol.3(1), 2015.

[19] D. Hosmer, S. Lemeshow, R. Sturdivant, "Applied logistic regression," 3rd Edition, Hoboken, NJ: John Wiley & Sons, 2013.

[20] A. Gillwald, A. Milek, C. Stork, "Towards evidence-based ICT policy and regulation gender assessment of ICT access and usage in Africa," Research ICT Africa , vol. 1, Policy Paper 5, 2010.

[21] P. Schumacher, J. Morahan-Martin J. "Gender, Internet and computer attitudes and experiences," Computers in Human Behavior, vol. 17(1), pp. 95-110, 2001.

[22] B. Bimber, "Measuring the Gender Gap on the Internet," Social Science Quarterly, vol. 81(3), pp. 868-876, 2000.

[23] B. Oyelaran-Oyeyinka, C. N. Adeya, "Dynamics of adoption and usage of ICTs in African universities: A study of Kenya and Nigeria," Technovation, vol. 24(10), pp. 841-851, 2004.

[24] R. E. Rice, J. E. Katz, "Comparing Internet and mobile phone usage: Digital Divides of usage, adoption, and dropouts," Telecommunications Policy, vol. 27(8-9), pp. 597-623, 2003.

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Specifics of operating systems of terminal devices in generating data traffic

Siniša Husnjak
Department of Information and Communication Traffic
Faculty of Transport and Traffic Sciences
University of Zagreb
Zagreb, Croatia
sinisa.husnjak@fpz.hr

Jelena Karlović, Matija Kolarić
Department of Information and Communication Traffic
Faculty of Transport and Traffic Sciences
Graduate students at University of Zagreb
Zagreb, Croatia
jkarlovic2@fpz.hr, mkolaric@fpz.hr

Abstract—**This paper presents the differences between operating systems (OS) of terminal devices and their characteristics in generating data traffic. An important factor in generating data traffic is also the OS of terminal device because each OS does not generate equal amount of data traffic. The aim of this paper is to show how much data traffic generate terminal devices that use different OSs. By accessing different networks, the same device also detects differences in generating data traffic. Obtained data from the study also shows the differences and characteristics of each OS. Data traffic measurements were performed by viewing the same video track using the YouTube application on all terminal devices and using NetWorx, SurPlusMeter, My Data Manager and Daily Data Sense applications. Access to the data network was achieved using Wi-Fi, UMTS and LTE networks. Based on obtained data by this study, a comparison of data traffic was made for each OS.**

Keywords– terminal devices, operating system, data traffic, measurements, comparison

I. INTRODUCTION

Today, there is a large number of OSs in the world and they are all intended for the same purpose, which is to enable the connection between the hardware and the user programs. However, OSs are different and each of them has certain characteristics that make it more or less quality depending on what the user requires. Without OS, no terminal device could work.

Generating data traffic is an inevitable segment of any modern terminal device, and the main reason for this is the implementation of a large number of applications and other device functionalities. A factor that greatly influences the generation of data traffic of a terminal device is an OS. In addition to the OS, factors such as screen size and resolution, device settings, device categories, web browser capabilities, individual cellular network communications technology, system and application updates, and complementary access networks also affect the usage of data traffic.

A. Previous research

The amount of data traffic generated by users of different smartphones varies considerably from the researches shown in [1] [2] [3]. As noted by the authors in [1], mobile Internet usage is shown on all mobile devices that enable the generation of mobile data traffic (MDT), also there are the factors that influence the generation of data traffic.

The measurements performed on mobile devices shown in [4] were supplemented by users who were participants in the research itself. According to [5] measurements on mobile devices were made to gather information related to the period of use of these mobile phones. The data was collected by test users who had pre-installed software/application that serves to monitor data traffic. The gathered data were used to describe the context of using mobile devices that help understand the user's behavior. The research presented in [6] focuses on measuring data traffic using mobile networks, and enables the application of market intelligence to different subjects.

The recent research shown in [7] demonstrates that the highest level of data traffic is generated by Mac OS, and the least Linux. Other research carried out by other telecommunications companies, such as Ericsson, has shown that Android devices generate more traffic than iOS due to larger screens, resulting in downloading larger files and video records with more pixels.

B. Research methodology

This paper compares the most frequently used OS, namely: Windows and Mac OS that are used on computer devices, and Android, iOS, and Windows Phone that are used on mobile devices. The measurements were done using different applications. The NetWorx and SurPlusMeter applications were used on the computer OS, which have the same function, respectively, they enable the measurement of data traffic generation. NetWorx is a program that allows to measure data traffic generation on Windows computers, and SurPlusMeter serves the same purpose on Mac OS of computers.

My Data Manager and Daily Data Sense were used on mobile devices. When data generation was measured, the My Data Manager application on Android and iOS was used, and the Daily Data Sense application on the Windows Phone. Both applications have the same purpose which is to measure generation of data traffic on mobile devices.

A video track was viewed for metrics purposes and data traffic measurement. The video track is available at "https://www.youtube.com/watch?v=FnzDneTj9YY" link and has a duration of 11:10 minutes. The video track was viewed at 480p quality, except for the Windows Phone that did not allow automatic quality change or display of the same.

II. DISTRIBUTION OF TERMINAL DEVICES

In today's time there are distinguished a large number of terminal devices. Every terminal device has its own characteristics such as size, version of the device, manufacturer, price of the device and appearance. The common characteristic of all terminal devices is that they all have access to the Internet. Figure 1 shows the use of various terminal devices in the world.

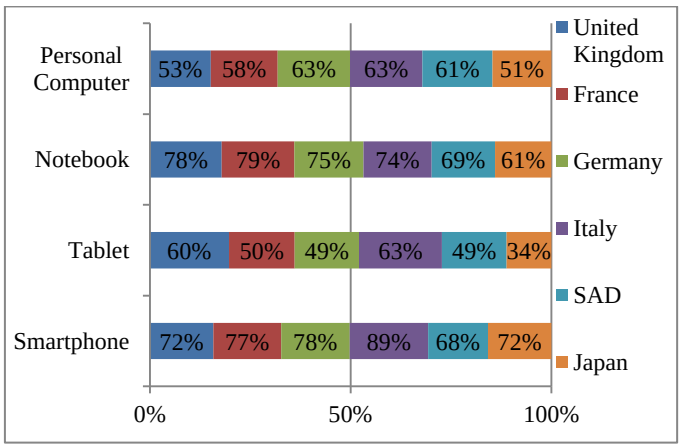


Figure 1. Use of various terminal devices in the world [7]

The Figure 1 shows that smartphones have the largest increase in usage of all of the above mentioned terminal devices. The reason is that smartphones provide almost identical services as other mentioned devices, but because of their size, they are more suitable for users to use.

Although these devices offer similar services, there are differences between them. Not all devices have the same price. Computers have different OSs than tablets and smartphones. Also, not all devices achieve the same rate of data transfer, or have the same resolution and screen size. All of these factors, expect the price of the device, affect the usage of data traffic. As the device has a higher screen resolution and aspect ratio, the quality of the content that is displayed on the screen will be higher, which will generate a larger amount of data traffic.

III. ANALYZED OPERATING SYSTEMS

OS is a software program that acts as a connection between computer users and computer hardware. OSs have been developed to allow the use of different types of applications on terminal devices. OSs differ in different types of devices, as well as their capabilities and specifics. In this paper are described OSs of mobile devices as well as OSs of computers.

A. *Computer operating systems*

Today there is a large number of OSs which differ in details, where some are more significant than others. However, this paper will cover only two types of computer OSs, as well as their variants and specifications of some, and these are: Windows and Mac OS. Figure 2 shows that Windows is the most widely used OS today.

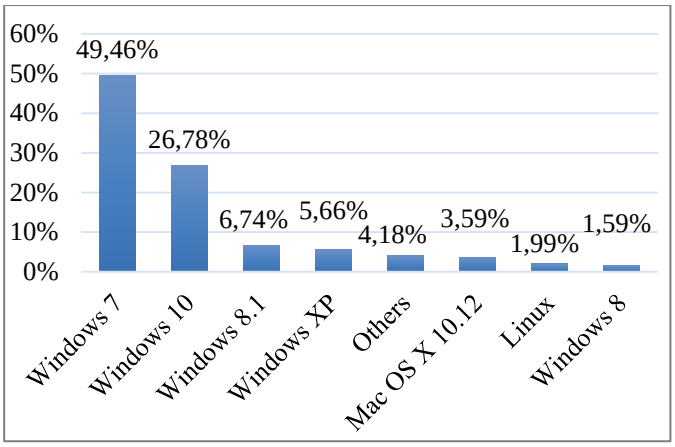


Figure 2. View of usage of different types of computer OSs [8]

Also, according to the picture, it is possible to see that Windows 7 is the most widely used version of Windows that is built on the core of Windows Vista. The most widely used version has become very popular, and the users used it most often for playing games.

1) *Windows operating system*

Most used Microsoft's OS is a Windows OS that can be used on almost all computers. It belongs to the category of open OS[1], and it is the most popular OS.

In this research, the Acer Aspire E1-532 notebook was used with OS Windows 10, which enables the "Start" menu that was not enabled in the previous version. What this system has is a quick start, embedded security, and returning the "Start" menu in an expand format [9].

2) *Mac operating system*

The MacBook Pro notebook used in this work has a built-in version of "Sierra" that is the last but one version of Mac OS X with new features: adding Siri to enable voice commands, optimized data storage, permissions for photos, messages, and iTunes. Also there is no credit card number entry through Apple Pay and transactions are protected [10].

B. *Operating systems of mobile devices*

The development of OS on mobile devices has been ongoing for a long time, and because of their complexity, it is still developing. Mobile OS combines OS of personal computer features with features that are required for manual or mobile use. There are different types of OS for mobile devices, and some of them are Android, Bada, Blackberry, iOS, Windows Phone, MeeGo, Palm, Symbian, etc.

Nowadays, most people use the Android OS which is more popular than the Windows OS used on computers. Accordingly, users on their mobile devices, even 80.7% of them use Android. In Figure 3 below we can see the percentage of usage of certain mobile OSs.

[1] An operating system developed for execution on various hardware platforms.

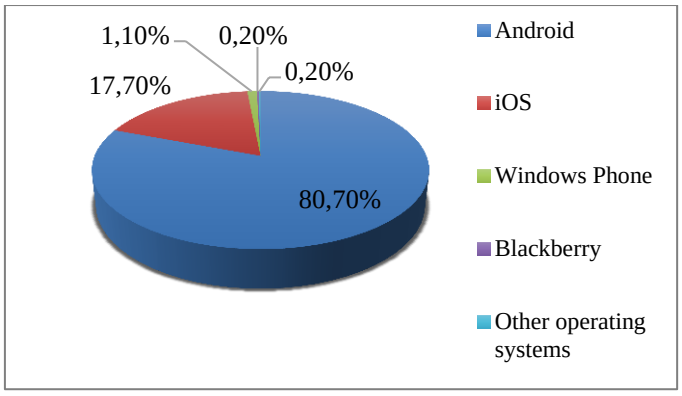


Figure 3. Percentage of OS usage on mobile devices [11]

1) Android operating system

Android is the first open OS for mobile devices launched by Google and run by the Open Handset Alliance. This system is based on the Linux core, and provides a platform that allows to create different applications and games for your users [12]. Shortly after entering the market, Android gains great popularity on the basis of an acceptable and beautiful look, as well as its efficient and productive work. The openness that Android offers makes it a favorite to different types of users, and that has marked a large increase in the demand for certain applications that can be found on their official Google Play store.

During the research, Samsung Galaxy A5 2015 device was used, and the OS version on it is 6.0.1 which is called Marshmallow. This version has a USB type C support, and fingerprint authentication support. In addition, it also provides Android Pay and improvements to Google Now [13].

2) iOS operativni sustav

iOS is actually abbreviation for the iPhone OS (iPhone Operating System) and runs only on Apple's mobile devices, such as: iPhone, iPad, and iPod Touch. Apple iOS includes the following features:

- Wi-Fi, Bluetooth and cellular connectivity, along with VPN support;
- Integrated search support, which enables simultaneous search through files, media, applications and email;
- Gesture recognition supports – for example, shaking the device to undo the most recent action;
- Safari mobile browser;
- Integrated front – and rear-facing cameras with video capabilities;
- Direct access to the Apple App store and the iTunes catalog of music, podcasts, television shows and movies available to rent or purchase;
- Compatibility with Apple's cloud service, iCloud
- Siri personal assistant
- *Apple Pay* which stores users credit card data and allows them to pay for goods and services directly with an iOS device [14].

The research was performed on the mobile device iPhone 6 with OS version 10.3.3, which is also the latest version of the system. Version 10 is currently supported and supports the iPhone 5 models, and so on. With this version, it was announced that it was the biggest release ever, and has completely redesigned features to lock the screen with 3D touch-sensitive information, and a very simple camera and screen. SMS messages have been updated and enhanced, and now there are options for sending animations, different effects, and so on.

Also, the voice command Siri has also been enhanced in a way that enables developers to build support in their applications. In addition, Maps and Apple Music are also redesigned with a simpler interface, and Maps now have different proactive suggestions, and music has the ability to find songs better and easier, as well as better focus on the same [15].

3) Windows Phone operating system

Windows Phone is the OS for mobile terminal products developed by Microsoft. The design used in this operating system is Metro. As with the previous two operating systems, Windows can also install different applications and they are available through the application store called Microsoft Windows Marketplace for Mobile [16]. Windows Phone has only three versions and they are: Windows Phone 7, Windows Phone 8, and Windows Phone 8.1. Windows also has the fourth version called Windows Mobile 10, which is the latest version of the system.

The research in this paper was performed on a Nokia Lumia 535 mobile device that has a version of Windows Mobile 10 OS which is the latest version of the Windows system, and which focuses more on the user's experience and functionality. The OS on mobile devices is very similar to those on computers. Many apps have been enhanced, so the app of settings is completely reorganized and universal for all Windows 10 devices.

Also, top-of-the-list applications appear on recently-used apps, and it is making it easier to use a mobile device, and speeds up the time which takes to find a particular app [17]. It is possible to establish voice interaction by dragging the names from the phone book and knowing how to pronounce them. Consequently, it represents the best version so far, and allows all users who have Windows 8.1 ability to update.

IV. GENERATION OF DATA TRAFFIC

Data traffic is generated by the transmission of information in data form between the source and destination terminal device while using part of the capacity of the common resources of the public telecommunications network [18].

Devices generate a large amount of data traffic that users are not aware of, so when we load the Facebook page ten times on the home page, we are generating around 2 MB of data traffic. This amount of data traffic depends on the number of notifications on the home page.

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

The Google Maps app is linked to a huge system of informations about companies, roads and offers. With the access to the app and a search for a location, the app will generate around 150-200 kB at the searched location. Google provides voice over voice services. Google transmits voice samples to their end servers for transcription, and after transcription, they send the text back. When performing voice search and the average length of words, Google reduces the sound by generating about 20 kB for the entire operation, excluding loading the results page [19].

In 2016 data traffic increased by 63% compared to 2015. The reached level of data traffic on a monthly level for 2016 is 7.2 EB (*Exabyte*)[2], while for the year 2015 is 4.4 EB per month [20].

Over the past five years, data traffic has increased 18 times. Fourth generation networks have the highest usage of data traffic, or four times more data traffic than third-generation networks. 4G networks have generated 69% of total data traffic last year, which is a large number given that 4G networks last year accounted for only 26% of mobile connections. 3G networks compared to 4G networks in 2016 generated 24% data traffic to 33% of mobile connections. 60% of total mobile data traffic was generated through Wi-Fi networks. In 2016, 429 million new mobile devices and connections were added, and most of these devices are smartphones and M2M (*Machine To Machine*) devices [16].

On a global level smartphone represent 46% of all mobile devices and connections, and with the use of 3G and 4G networks accounted for 89% of mobile data traffic. In 2016 it was noted that the average smartphone generated 13 times more data traffic than a feature phone, but it should be noted that mobile network speeds in 2016 increased three times compared to 2015, a speed of 6.8 Mbit/s compared to 2 Mbit/s. More than half, about 60 percent of data traffic is generated by browsing video tracks.

The level of generated data traffic is different from device to device, but the highest data traffic is generated by the use of smartphones, so in 2016, their usage increased by 38%. The average amount of data traffic on a smartphone for the past year was 1 164 MB per month, which is an increase of 70% compared to 2015. Each year there is an increase in the number of accomplished connections to the terminal device, so in the last year there was an increase in the connection of tablet devices by 26%, and personal computers by 8% [20].

By 2021, an even greater increase in data traffic is anticipated. Mobile data traffic will reach 49 EB per month, and it will reach an amount of 0.5 ZB (*Zettabyte*)[3]. Based on this, technology development will continue, and because of the further growth in technology it is anticipated that each resident will have an average of 1.5 mobile devices, which is a number of 11.6 billion worldwide mobile devices. It will also record the development of M2M devices that will exceed the world population.

[2] The unit of 1 EB is one billion GB (*Gigabyte*).
[3] One zettabyte amounts toone trillion GB.

By 2021, also the development and use of 5G network is foreseen. It is assumed that the 5G network will have a theoretical 10 Gbit/s transmission rate, while the actual transfer rates will be greater than 20 Mbit/s [20]. Figure 4 shows the increase in the use of terminal devices in 2017.

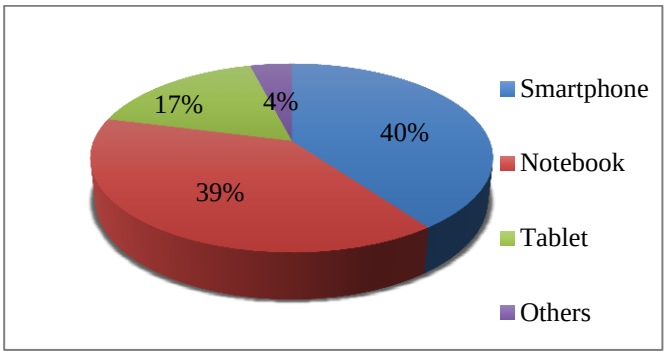


Figure 4. Overview of increased use of terminal devices in 2017. [21]

Also, large capacity networks and advanced devices enable the growth of data-intensive applications. Users migrate from existing 2G networks to 3G and 4G connections to benefit from advanced mobile devices with greater functionality and higher data rates offered by these networks. LTE network users tend to almost double the data transmission compared to users of other mobile network data traffic generations [22].

LTE network has a higher data transfer rate than other mobile network generations. The higher the data transfer rate is, the higher the amount of generated data traffic will be.

The amount of data traffic generated by a user who has access to a larger network generation with higher speeds significantly exceeds the amount of generated data of a user who had access to previous generation networks. This is proven because Android smartphones with the ability to access 4G networks have downloaded 13.1 GB of data traffic per month, while Android users with only 3G access can generate 5 GB of data traffic per user in the same month [22].

Assuming this 4G network growth, it is expected that 2/3 of the total amount of MDT in 2019 will generate 4G networks, as can be seen in Figure 5.

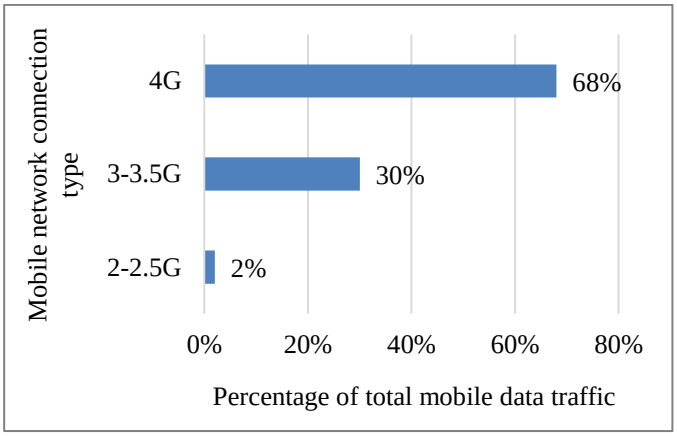


Figure 5. Prediction of generated MDT per connection type in 2019, globally [22]

V. RESULTS OF RESEARCH

In recent years, there has been a huge increase in data traffic generation as well as in the use of various OSs. However, each OS differs in the generation of data traffic where each of its components already has different applications that have different usage of data traffic. Each application generates traffic differently depending on which type of content is being used.

In the research conducted on the generation of data traffic according to the OS, OS of computers and mobile OS are measured. OS of computers which were used are Windows and Mac OS, while mobile OSs are: Android, iOS, and Windows Phone.

Own OS surveys were performed using various applications to determine the generation of data traffic. There are different applications because each OS does not support the same program that allows measuring of data traffic.

Measuring was performed while browsing video tracks on YouTube to determine how much data was on downlink and on uplink in the period when the video track was viewed. The experiment was repeated multiple times on each operating system to determine the accuracy of the results. During the experiment, other applications did not make any communication. In addition, the app that showed the total amount of data traffic usage had possibility to show the percentage for each app separately, as well as for the YouTube where the measurement was performed.

A. *Generating data traffic on computer operating systems*

The measurements were performed on the Windows OS and on Mac OS. Figure 6 shows the obtained results on both operating systems. The results were obtained while browsing the video track on the YouTube application in a time interval of 11:10 minutes. These results are shown in the Figure 6.

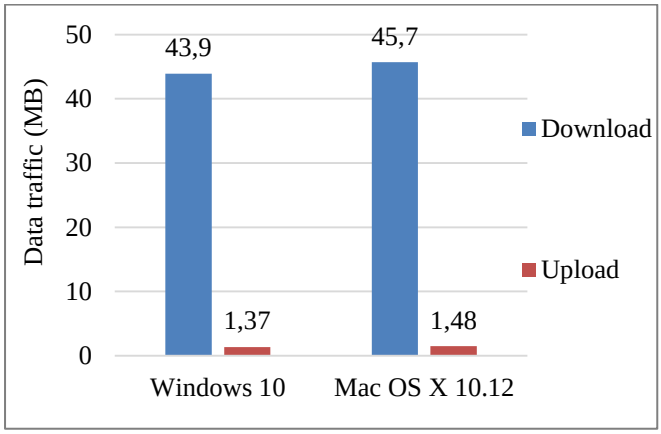


Figure 6. Overview of the generation of data traffic between the OS of computers on data in downlink and uplink

Windows OS 10 on the terminal device Acer Aspire E1-532 generated 43.9 MB (megabytes) of data in downlink, while in uplink the device generated 1.37 MB which would mean that they totally generated data traffic was 44.46 MB. MacBook Pro which was used in this measurement has Mac OS X 10.12 version and it generated a larger amounts of data in downlink and uplink than Windows OS 10. On the downlink 45.7 MB of data was generated and on uplink that number was 1.48 MB which is approximately like Windows and makes totally generated data traffic equivalent to 46.56 MB.

It has been shown that while using the same application and the same content, different OSs do not have the same usage of data traffic.

B. *Generating data traffic on operating systems of mobile devices*

On the mobile OS data traffic generation was measured using the same video track on the YouTube app in 11-minute interval. Measurements were obtained on mobile devices while they were connected to mobile networks (3G and 4G) and to the Wi-Fi network. The OSs on which data traffic is measured are: Android 6.0.1, iOS 10.3.3, and Windows Phone 10. The Android device which was used is the Samsung Galaxy A5 2015, and the iOS device is the iPhone 6 and Windows Phone is Nokia Lumia 535.

When data generation was measured, the My Data Manager application on Android and iOS was used, and the Daily Data Sense application was used on the Windows Phone. The applications have the same purpose, which is to enable data traffic measurement across different networks. Traffic generation measurement can be displayed at specific intervals or by different applications.

The measurement results can be seen in Figure 7.

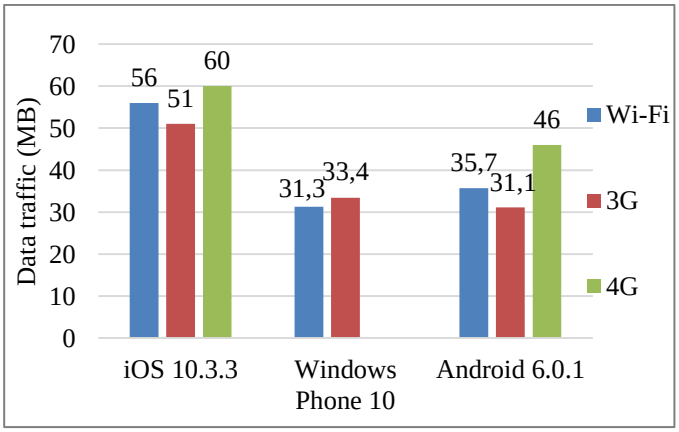


Figure 7. Overview of generated data traffic between OS on mobile devices

According to the measurements seen in the figure 7, it is apparent that the greatest amount of generated data traffic was on the Wi-Fi network by using iOS 10.3.3 or iPhone 6. That amount was 56 MB of data traffic. Android 6.0.1 device generated a total amount of 35.7 MB of data, and the last was the device using Windows Phone 10 with 31.3 MB of generated data traffic.

The order is slightly different by using 3G network. As well as on the Wi-Fi network, the iOS device generated the highest amount of data traffic, which was 51 MB, followed by Windows Phone 10, where the generation on the 3G network is similar to the Wi-Fi and it was 33.4 MB. The least device was mobile device using Android 6.0.1 with 31.1 MB of generated data traffic.

By using 4G network, only Android 6.0.1 and iOS 10.3.3 measurements were performed because device based on Windows Phone 10 did not support the connection to 4G network. A higher level of data traffic was generated on iOS 10.3.3 with 60 MB of data, while Android 6.0.1 mobile device generated 46 MB of data traffic.

C. Comparison of the gathered data

According to the own measurements in generating data traffic, a certain difference can be observed between OS of computers and OS of mobile devices. All OSs were tested by using the Wi-Fi network and because of that it is best to show comparison according to it.

Figure 8 shows how this comparison looks between OSs in generating data traffic while they were connected to the Wi-Fi network and while the same video track was playing on every of the device.

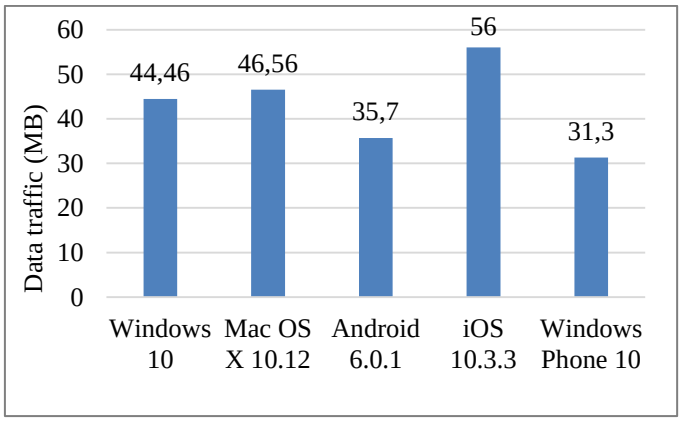


Figure 8. Comparison of OS by generating data traffic by using the Wi-Fi networks

Figure 8 shows the differences in generating data traffic while all of the devices were connected to the Wi-Fi network. It is shown that the Windows Phone 10 on the Nokia Lumia 535 generated the smallest amount of data traffic while iOS 10.3.3 on iPhone 6 generated the greatest amount of data traffic. Windows Phone 10 generated 31.3 MB of data, and iOS 10.3.3 generated 56 MB of data traffic, representing up to nearly double amount of generated traffic for the same video track.

VI. CONCLUSION

We live in a time where men's daily life without terminal devices became unthinkable. The use of terminal devices became a habit, but also a need, so during the day every human being uses at least two to three terminal devices.

Terminal devices could not provide different options without an OS who allows the functionality of all applications on them. The OS is actually used to allow the user to communicate with the desired applications. However, if the user wants to communicate with other people throughout their terminal device, he usually needs access to data networks. Today there are different networks, and the most often used are Wi-Fi, 3G and 4G networks.

The amount of generated data traffic is influenced by factors such as the OS that the terminal device applies, screen size, communication technology, etc. So, a device with a large display and resolution will generate a large amount of data traffic because it is able to show a large amount of high-quality data. Through the years, the development of mobile networks has been noted. With the development of mobile network technologies, the results show a higher amount of data traffic usage. Each mobile network generation has different transfer rates, so newer generations of mobile networks have higher data transfer rates.

Generating data traffic is not the same for all OSs, regardless of whether the same application is used on everyone. This can be supported by the results that show that the most data traffic is generated by iOS OS. While browsing the video track on Wi-Fi network, iOS has generated 56 MB of data. At the same time, using the same video track on the same Wi-Fi network, Windows Phone device generated 31.3 MB of data, which is much less than iOS.

Beside that some OSs use more data traffic when loading certain content, the overall amount of data traffic is affected by the fact that some OSs are more interesting with their applications, so the user is spending more time on the device and generates more data traffic. Also, the same applications on different OSs do not generate the same amount of data traffic. Although unaware of that, the terminal device generates background data as our application runs, or simply needs more data on one device to run than on other.

REFERENCES

[1] Gerpott T. J., Thomas S., "Empirical research on mobile Internet usage: A meta-analysis of the literature," Telecommunications Policy Journal, 2014., vol. 38, no. 3, pp. 291–310

[2] Gerpott T. J., Thomas S., Weichert M., "Characteristics and mobile Internet use intensity of consumers with different types of advanced handsets: An exploratory empirical study of iPhone, Android and other web-enabled mobile users in Germany," Telecommunication Policy Journal, 2013., vol. 37, no. 4–5, pp. 357–371

[3] Gerpott T. J., Thomas S., Weichert M., "Personal characteristics and mobile Internet use intensity of consumers with computer-centric communication devices: An exploratory empirical study of iPad and laptop users in Germany", Telematics and Informatics, 2013., vol. 30, no. 2, pp. 87–99

[4] Smura T., Kivi A., Töyli J., "A framework for analysing the usage of mobile services," The journal of policy, regulation and strategy for telecommunications, information and media, 2009., vol. 11, no. 4, pp. 53–67

[5] Soikkeli T., "The Effect of Context on Smartphone Usage Sessions," Master thesis, University School of Science, Aalto, Finland, 2011.

[6] Riikonen A., "Mobile Internet Usage - Network Traffic Measurements," Doctoral thesis, Helsinki University of Technology, Helsinki, Finland, October, 2009.

[7] D. Chaffey (2017, Sep 05). *Mobile marketing statistic compilation*. Available at: http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/.

[8] V. Beal (2017, Sep 05). *History of Windows OS*. Available at: http://www.webopedia.com/DidYouKnow/Hardware_Software/history_of_microsoft_windows_operating_system.html.

[9] M. Rouse (2017, Sep 05). *Windows OS*. Available at: http://searchwindowsserver.techtarget.com/definition/Windows.

[10] Apple Inc. (2017, Sep 05). *Mac OS X 10.12*. Available at: https://www.apple.com/macos/sierra/.

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

[11] J. Vincent (2017, Sep 05). *Operating systems of mobile devices*. Available at: https://www.theverge.com/2017/2/16/14634656/android-ios-market-share-blackberry-2016.

[12] Android Developers (2017, Sep 05). *Android*. Available: https://developer.android.com/about/index.html.

[13] (2017, Sep 05). *Android versions comparison*. Available at: http://socialcompare.com/en/comparison/android-versions-comparison.

[14] M. Rouse (2017, Sep 05). *Apple iOS*. Available at: http://searchmobilecomputing.techtarget.com/definition/iOS.

[15] J. Clover (2017, Sep 05). *iOS 10*. Available at: https://www.macrumors.com/2016/09/13/apple-releases-ios-10/.

[16] F. Stroud (2017, Sep 05). *Windows Phone*. Available at: http://www.webopedia.com/TERM/W/windows_phone.html.

[17] M. Rougeau (2017, Sep 05). *Windows Mobile 10*. Available at: http://www.techradar.com/news/phone-and-communications/mobile-phones/windows-10-for-phones-detailed-including-better-voice-commands-1281560.

[18] D. Peraković, (2017, Sep 07). "Authorized lectures from the course Terminal devices," University of Zagreb, Faculty of Transport and Traffic Sciences, 2017.

[19] R. Whitwam (2017. Sep 07). *How much data the most popular smartphone tasks and apps actually use?* Available at: http://www.tested.com/tech/smartphones/3105-how-much-data-the-most-popular-smartphone-tasks-and-apps-actually-use/.

[20] Cisco. (2017, Sep 07). *Global Mobile Data Traffic Forecast Update.* 2016–2021. (2017). [Online] Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html.

[21] Kevin C. Tofel (2017, Sep 07). *Which devices use wifi hotspots the most?* Available at: https://gigaom.com/2012/11/07/which-devices-use-wi-fi-hotspots-the-most-not-laptops/.

[22] S. Husnjak, D. Peraković, and I. Cvitić: *Relevant Affect Factors of Smartphone Mobile Data Traffic*, Promet – Traffic&Transportation, 2016., vol. 28, no. 4, pp. 435–444

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Design and Development of a Diagnosis System with a SVM-VOA Hybrid Model Approach

Utku Kose
Dept. of Computer Engineering
Suleyman Demirel University
Isparta, Turkey
utkukose@sdu.edu.tr

Serkan Ugur
Institute of Science and Technology (Graduate Student)
Turkish Aeronautical Assoc. University
Ankara, Turkey
pasha036@gmail.com

Pandian Vasant
Faculty of Science and Inform. Tech.
Universiti Teknologi Petronas
Perak, Malaysia
vasantglobal@gmail.com

Abstract—**Artificial Intelligence has a great effect on all fields of the modern life. It is currently widely applied in different types of real-world based problems and in time, it has even divided into several sub-fields regarding to its application approaches, methods, and techniques. As based on learning from a pre-data to deal with new cases, Machine Learning techniques are very remarkable in this manner and they are used in developing advanced intelligent solutions for problems. In this study, an intelligent diagnosis system, which is formed with a hybrid approach of Vortex Optimization Algorithm (VOA) trained Support Vector Machines (SVM) has been introduced. Thanks to the related techniques, the system is generally able to applied in different kinds of diagnosis processes after it has been adjusted according to key parameters of the objective problems. The study briefly focuses on details regarding to infrastructure of the system and provides some findings on its application on two diagnosis problems (medical and fault oriented ones).**

Keywords- diagnosis system, support vector machines, vortex optimization algorithm, artificial intelligence, machine learning

I. INTRODUCTION

Nowadays, humankind is in the process of rapid and revolutionary developments in a technological manner. The related developments are generally because of some remarkable technologies or cause for newer scientific fields. At this point, some scientific fields and research areas like Artificial Intelligence, Big Data, Cloud Computing has important role on shaping humankind's future. Here, especially Artificial Intelligence has more steps ahead according to other ones because of its wide multidisciplinary application ways. No one can deny that rise of the field of Artificial Intelligence (AI) has affected many fields associated with the modern life and made everything better and more practical [1-5]. In this context, roles of AI in solving real-world based problems is an important subject that should be considered in order to understand more about how this scientific field is affecting nowadays and the future.

When the associated literature is taken into consideration, some problem solution ways that AI currently employs can be stated briefly as follows [6-11]:

- Prediction,
- Recognition,
- Diagnosis,
- Data modelling,
- Optimization,
- Adaptive Control

Diagnosis as one the mentioned solution ways has a long-time popularity among AI oriented studies. That's because the role of AI approaches, methods, and techniques has made it possible to analyze previously obtained data and receive diagnosis based comments as from some kind of artificial expert knowledge. The concept of diagnosis can be defined as "identifying nature of an illness or other problem with examination of the related symptoms." [12]. As it can be understood from the definition, mechanism of the diagnosis has been very effective in especially medical problems. But of course, the concept of diagnosis gives a wider meaning to the AI when real-world based problems are considered in detail. By eliminating details regarding to application approaches, diagnosis based problems in which AI can be employed easily is examined under two titles as follows [13, 14]:

- Clinical – medical diagnosis,
- Fault diagnosis.

According to real-world based cases, clinical – medical diagnosis is associated with diagnosis of diseases appeared in living organisms (i.e. humans, animals) while fault diagnosis is related to problems occurred in machines [13, 14]. In detail, the main objective here is to benefit from known symptoms or some essential information that are useful for prediction –

detection to have early ideas about situation of a disease or fault. So, by using features – functions of AI, it is possible to develop some diagnosis systems, which can be called with some concepts like prediction environments on expert knowledge or Expert Systems [15, 16] (although the Expert Systems can be examined separately because of its some additional details, diagnosis systems can be defined with also that name according to the authors). Eventually, the applications of AI based diagnosis have taken researchers' interests for a while and it has been widely done thanks to different variations of AI techniques.

When the AI mechanism is examined in detail under the diagnosis problem, it can be seen that Machine Learning, which is as sub-field of the AI as a result of its wide applications via 'learning from pre-data', has an active role in diagnosis applications. In detail, some data has been used for training Machine Learning based techniques to run accurate diagnosis processes for newly encountered cases – symptoms. Here, employment of hybrid systems (with a core and a supportive technique for the learning) is a popular system design chosen by researchers [17-21].

In this study, an intelligent diagnosis system, which is formed with a hybrid approach of Vortex Optimization Algorithm (VOA) trained Support Vector Machines (SVM) has been introduced. Thanks to the related techniques, the system is generally able to applied in different kinds of diagnosis processes after it has been adjusted according to key parameters of the objective problems. The study briefly focuses on details regarding to infrastructure of the system and provides some findings on its application for two diagnosis problems (medical and fault oriented ones).

Considering the subject – objective of the study, remaining sections are organized as follows: The next section provides some brief information about two techniques: SVM and VOA as included for designing the diagnosis system. After that section, the third section provides some details on the hybrid diagnosis system introduced here. Next, the fourth section is devoted to two applications done with the diagnosis system and finally, the last section provides conclusions and discussion on future work.

II. BACKGROUND

The diagnosis system introduced in this study is based two AI based techniques, which are Support Vector Machines (SVM) and Vortex Optimization Algorithm (VOA) respectively. Some brief information regarding to these techniques are as follows:

A. Support Vector Machines (SVM)

As introduced by Vapnik and Lerner, SVM is a technique used for generally classification and regression [22, 23]. SVM is based on the approach of optimally separating hyper-plane so that the margin between two groups is maximized. In detail, the elements of support vectors here are known as subset of data instances (to be employed for determining the hyper-plane) while the distance seen between the nearest support vector and the hyper-plane is briefly margin [24] (Figure 1. [25]).

Under the role of SVM for classification and regression problems, the technique has two different types called as 'Linear SVM' and 'Non-linear SVM' respectively. The difference between these two types are based on using a linear or non-linear 'decision boundary' to separate the related data considered [26, 27].

As a Machine Learning technique, SVMs is based on 'supervised learning' and it tries to find the optimum hyper-plane, which can classify data points, by employing a training set storing support vectors located on the closest points of the classes [24, 28].

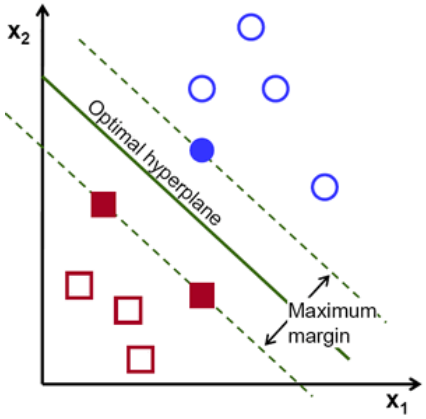


Figure 1. A SVM with hyper-plane [25].

In order to understand more about the problem of determining optimum hyper-plane, it is possible to examine it under optimization oriented definitions. Briefly, the optimum separating hyper-plane in a SVM can be defined as [24, 26, 29]:

$$w^T x_i + b \geq 1 - \xi_i \; y_i = 1 \quad (1)$$

$$w^T x_i + b \leq -1 + \xi_i \quad (2)$$

$$\xi_i \geq 0 \quad (3)$$

where ξi approximates the number of samples that were misclassified. Over a training data set (under input vector xi Rn, i = 1, ..., l and the output labels yi, y {1,-1}, a model of SVM should find solution for the optimization problem mentioned as follows [24, 26, 29]:

$$\text{Min.} \quad \frac{\|w\|^2}{2} + c\sum_{i=1}^{l} \xi_i \quad (4)$$

$$\text{subject to:} \quad y_i(w^T x_i + b) \geq 1 - \xi_i \; \xi_i \geq 0 \quad (5)$$

where c > 0 is a value - parameter between the error and the margin of hyper-plane. Thanks to Lagrange multipliers, the related optimization problem can be represented in its dual form as follows [24, 26, 29]:

$$\text{Max.} \quad W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1,\, j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (6)$$

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

$$\alpha_i > 0, \sum_{i=1}^{n} \alpha_i y_i = 0$$

subject to: (7)

The formulation can be explained in a clearer way [26, 30]:

- Each αi', except for the ones as training patterns and close to support vectors are zero.
- So, data points, which are not support vectors have no effect on the solution.
- Eventually, the hyper-plane can be formed via only information from support vectors.

As different from Linear SVMs, Non-linear SVMs are generally based on the data, which do not show a linear distribution. In this case, the data should be accessed after mapping results in a higher dimensional space. In detail, Non-linear support vector of the input space should be mapped in a higher dimensional feature space that is able to be divided into training data. Here, it is important to use a kernel function to achieve that classification approach. When dealing with Non-linear SVMs, determining which kernel function to use is an effective factor on obtaining a desired classification performance. Some kernel functions that are used in the literature are as follows [24, 26, 29]:

Linear: $K(x_i, x_j) = x_i^T x_j$ (8)

Gaussian (RBF): $K(x_i, x_j) = e^{\frac{\|x_i - x_j\|^2}{2\sigma^2}}$ (9)

Polynomial: $K(x_i, x_j) = (1 + x_i^T x_j)^p$ (10)

By taking the related kernel functions into consideration, one effective research approach is to optimize parameters of a kernel function employed during the training process. For example, it is possible to optimize the σ (sigma) parameter of a chosen Gaussian (RBF) function to achieve 'better classifying' Non-linear SVM model. As associated with this approach, the research way followed in this study has been to optimize this parameter with an AI based optimization in every diagnosis problem case. In the literature, there are lots of different AI based optimization algorithms – techniques (i.e. Particle Swarm Optimization [31-33], Cuckoo Search [34, 35], Artificial Bee Colony [36, 37], Firefly Algorithm [38, 39] and many others) that can be employed for the same purpose. In the study, a recent algorithm: Vortex Optimization Algorithm (VOA) has been chosen for this task.

B. Vortex Optimization Algorithm (VOA)

VOA is a recent AI based optimization technique, which was designed and developed by Kose and Arslan [6, 40]. Briefly, it is inspired from the dynamics of vortices in the nature and employ some Swarm Intelligence based mechanisms to ensure an intelligent, meta-heuristic optimization process. With its particle elimination step(s), it also considers evolutionary solution processes through an optimization problem. Since its first introduction to the literature, it has received some modifications and also been employed in many multidisciplinary problems. Figure 2 illustrates particle flow in the simple form of the VOA [40].

By referring the readers to the sources: [6, 40] for more detailed information about it, algorithmic structure of the VOA can be expressed in a shorter form as follows:

- **Step 1 (Initialization Phase):** Determine the initial parameters (N: number of particles, initial vorticity (v) values of the particles, max. and min. limits (min. limit is the negative of the max.) for v value (max_v and min_v), e for the elimination rate, any other parameters for i.e. in-system optimization. Furthermore, adjust the related values for the problem considered.
- **Step 2:** Place the particles randomly in the solution space and calculate fitness values over the objective function(s). Update the v value of the best particle with a random value and accept it as a 'vortex'.
- **Step 3 (Loop Phase):** Perform loop steps below according to the stopping criteria(s):
- **Step 3.1:** Accept the particles having fitness values equal to or better than the average fitness of all particles as 'vortices'. Accept other ones as 'normal' particles.
- **Step 3.2:** Update v value of each particle by dividing the best v with its v and adding its randomly improved value to the current v.
- **Step 3.3:** Improve v values of 'vortex' particles except from the best ones with random values.
- **Step 3.4:** Move each particle to the best one by using position differences and v strengths.
- **Step 3.5:** Calculate fitness values over the objective function(s). Accept the best particle as 'vortex'.
- **Step 3.6 (Elimination):** If total number of 'normal' particles is equal to / under e, eliminate normal particles and create-place new ones as the same number.
- **Step 3.7 (In-system optimization):** If the optimization process so far does not seem stable enough, perform in-system optimization (in especially bigger problems).
- **Step 4:** Solution is the optimum value(s) after the whole loop process.

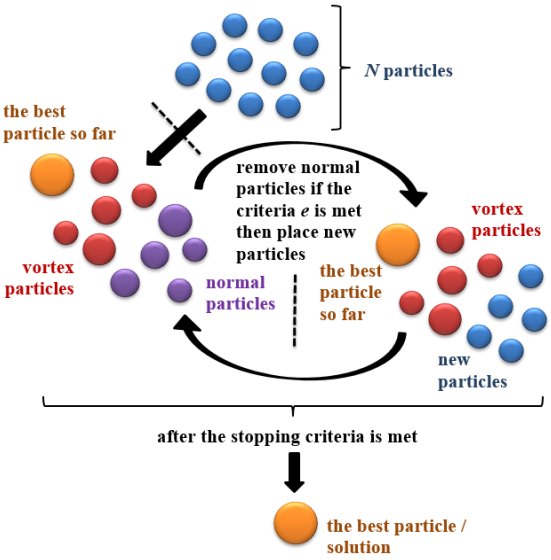


Figure 2. Particle flow in the simple form of the VOA [40].

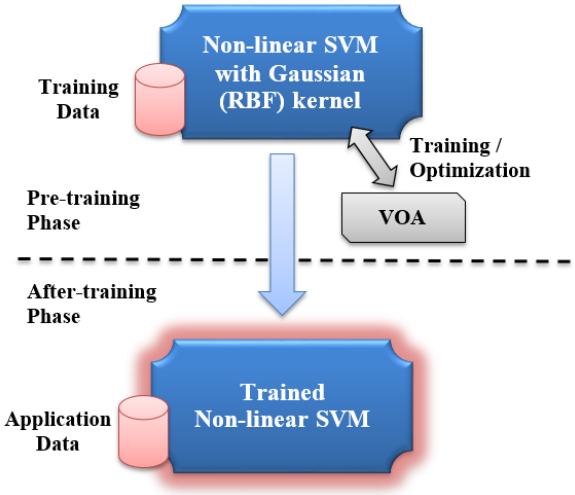


Figure 3. A brief schema explaining the hybrid model – system structure of the SVM-VOA.

III. SVM-VOA FOR DIAGNOSIS

Research performed along this study consists of developing a hybrid model based on the SVM as a popular AI – Machine Learning – classifying technique and determining an alternative AI based optimization technique to optimize its kernel function parameter as leading to especially diagnosis problems. At this point, integration of SVM and VOA has resulted to an alternative system to be considered along the related diagnosis applications. This section includes some brief explanations on model – system structure and the applied training – diagnosis processes.

A. *Model – System Structure*

The developed SVM-VOA can be examined under two different phases as 'pre-training' and 'after-training'. Model – system approach formed in the study can be examined in a clearer when the pre-training phase is examined.

Structure of the model in the pre-training phase includes a Non-linear SVM employing Gaussian (RBF) kernel function. The most difficult task is related to the training process performed on this pre-training phase. VOA takes an important role on the pre-training and according to the objective diagnosis problem, it is run for training the SVM by optimizing the Gaussian (RBF) kernel function. After it is seen that the SVM is trained enough to achieve good diagnosis results over the problem, then the SVM structure can be employed in the after-training phase.

After-training phase is just the application life-cycle of the SVM. Of course it is important to see if there is newer aspects of the considered diagnosis problem when running the training SVM. This can be achieved by running a self-evaluation task if desired. But generally it is not too necessary if all aspects of the diagnosis problem are covered enough within the training data used for training the SVM.

Figure 3 presents a brief schema explaining the hybrid model – system structure of the SVM-VOA.

B. *Training the SVM and the Diagnosis*

Training process performed for the SVM is generally based on finding the optimum σ (sigma) parameter of the Gaussian (RNF) kernel function, as it is mentioned before. At this point, particles of the VOA have been used to determine the optimum value through some iterative steps on employing the SVM on the objective diagnosis problem. A typical training and diagnosis process on SVM-VOA based diagnosis system can be explained as below:

- Each VOA particle is associated with the sigma (σ) value used in the Gaussian (RBF) kernel function.
- By taking a specific iteration number, the VOA is run for determining alternative sigma (σ) values. Within each new iteration, alternative sigma (σ) values of the particles are used in the kernel function (and so in the SVM) and correctness accuracy values (Equation 11) are used to determine which particle (value) is good for classifying training data of the objective diagnosis problem.

$$\frac{100 * CC}{TS} \qquad (11)$$

- (In the Equation 11, CC is for the number of total correctly classified symptoms and TS is for the total number of symptom sets considered.)
- With the end of VOA iteration steps, the sigma (σ) value enabling the SVM to achieve the best correctness accuracy is used under the Gaussian (RBF) kernel function of the SVM.
- After the training process, the trained SVM is now can be applied in 'real' applications to get decision support for newly encountered symptoms regarding to the objective problem.

In order to see if the formed hybrid model of SVM-VOA (diagnosis system) and the followed aspects in training –

application process are effective enough, the system has been used in some diagnosis applications.

IV. DIAGNOSIS APPLICATIONS

The developed diagnosis system has been applied in two different diagnosis problems in order to understand more about its effectiveness – efficiency and evaluate the success. Details related to the applications have been provided within next sub-titles.

A. Medical Diagnosis

Medical diagnosis application on the system has been done over 'headache disease diagnosis'. The International Headache Society (HIS) classifies headache diseases in 13 different classes while 4 of them are known as primer headache diseases and the remaining 9 ones are as secondary headache diseases [41]. Especially diseases in the classification 'secondary' are related to usually eye diseases, brain tumors, celebral hemorrhage, and even problems related to the neck [42, 43]. Because of that it is a vital approach to determine if an individual is associated with one of the headache classifications or healthy according to the related symptoms taken into consideration.

In the study, symptoms reported – considered in [42] are used as data for the SVM and at this point, a total of 439 symptom sets gathered from Usak State Hospital and Afyonkarahisar State Hospital in Turkey are used to train and test the diagnosis system. Table 1 provides information on the symptoms evaluated over just two states (yes or no) easily.

TABLE I. HEADACHE DIAGNOSIS SYMPTOMS [42].

No	Symptom
1	Is there any vomitus – vomiting?
2	Is there any nausea?
3	Is there any photophobia?
4	Is there any phonophobia?
5	Is there any aura symptom?
6	Can the headache be categorized as a throb?
7	Is the attack time is between 4-72 hours?
8	Does the headache intensity improve after any move?
9	Does the headache take place on multi points?
10	Is the headache level moderate – medium?

221 of the related symptom sets were used for training process while the remaining ones (218) are used within tests. In detail, 152 symptom sets in the training data are associated with the individuals with illness and the remaining 69 symptoms sets are from healthy – normal individuals.

Table 2 provides correctness accuracy (rates on correctly classified test symptom sets) findings, found σ (sigma) parameters for the Gaussian (RNF) kernel function, total applied VOA particle numbers, and total training iteration numbers on four different applications.

TABLE II. FINDINGS OVER MEDICAL DIAGNOSIS TESTS WITH SVM-VOA ON FOUR DIFFERENT APPLICATIONS.

App. No	Correct. Accur.	σ (sigma)	VOA particle no.	Total train. Iteration
1	72,86%	9,8941	25	3000
2	76,15%	8,9016	50	5000
3	78,58%	0,9601	50	7500
4	86,42%	0,8851	75	10000

The obtained findings show that the diagnosis system is able to provide effective enough decision support on determining headache diseases.

B. Fault Diagnosis

In the fault diagnosis application, a 'Dissolved Gas Analysis (DGA)' based oil-filled transformer fault diagnosis problem has been considered. In detail, data set called as 'IEC TC 10' and from [44] was used through the application (More details about the problem are in [44]).

Table 3 and Table 4 provides brief information on the symptom factors and fault types in the IEC TC 10, respectively.

TABLE III. IEC TC 10 – SYMPTOM FACTORS [44].

No	Symptom Factor
1	H2
2	CH4
3	C2H2
4	C2H4
5	C2H6
6	CO
7	CO2

TABLE IV. IEC TC 10 – FAULT TYPES [44].

No	Faulty	Cause
1	Partial Discharges (PD)	"cold plasma (corona) type with possible X-wax formation, and of the sparking type inducing small carbonized punctures in paper"
2	Discharges of Low Energy (D1)	"larger punctures in paper, tracking, or carbon particles in oil"
3	Discharges of High Energy (D2)	"extensive carbonization, metal fusion, and possible tripping of the equipment"

4	Thermal Faults Below 300 ºC (T1)	"paper has turned brownish"
5	Thermal Faults Above 300 ºC (T2)	"paper has carbonized"
6	Thermal Faults Above 700 ºC (T3)	"oil carbonization, metal coloration, or fusion"

The IEC TC 10 is formed by a total of 167 data samples and 120 of these samples were used within the training processes of SVM. As similar to the four different applications done for the medical diagnosis (explained under the previous sub-title), the fault diagnosis problem on the IEC TC 10 has also been tried to be solved within four applications, over 47 data samples used for testing. Table 5 presents correctness accuracy (rates on correctly classified sets) findings, found σ (sigma) parameters for the Gaussian (RNF) kernel function, applied VOA particle numbers, and also total training iteration numbers on four different applications done.

TABLE V. FINDINGS OVER FAULT DIAGNOSIS TESTS WITH SVM-VOA ON FOUR DIFFERENT APPLICATIONS.

App. No	Correct. Accur.	σ (sigma)	VOA particle no.	Total train. Iteration
1	68,94%	9,9048	25	3000
2	70,27%	7,0091	50	5000
3	74,11%	3,7088	50	7500
4	80,94%	1,0872	75	10000

Considering the findings with the fault diagnosis application, it is possible to indicate again that the diagnosis system seems effective enough to give decision support on determining faults in the problem of DGA based oil-filled transformer.

V. CONCLUSIONS AND FUTURE WORK

In this study, a diagnosis system formed by a hybrid model of SVM-VOA has been introduced. By using the system, it is possible to deal with diagnosis problems including especially medical and fault oriented ones. In detail the diagnosis mechanism has been done over SVM structures trained thanks to the VOA technique. At this point, kernel function for the SVM is chosen as the widely-used Gaussian (RBF) function and the σ (sigma) parameter of this function was optimized by VOA, in order to obtain better classifying results on the problems. The realized, two different diagnosis applications have shown that the formed SVM-VOA based system is applicable enough at performing diagnosis processes and giving accurate outputs as a typical decision support approach. In detail, it can be used as a fault diagnosis decision support system and an automated medical decision support system, which are both remarkable application types of Artificial Intelligence and Machine Learning [45-47].

A. *Future Work*

Based on the introduced SVM-VOA hybrid model, the authors think about more future studies regarding to the diagnosis system. In this context, infrastructure of the system will be evaluated from different, additional perspectives in order to see if there is anything to do in order to improve success more and more. Future works will include also focus on using different kernel functions to make the diagnosis system having wider scope on alternative SVM background. Finally, more alternative types of diagnosis problems will also be considered as continuation of evaluations regarding to success of the system on diagnosis.

REFERENCES

[1] M. R. Sabhnani, P. P. Rao and A. V. Panchal, "AI software in everyday life: A survey". Proceedings of MAICS Conference, 2001. Online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.202.5326&rep=rep1&type=pdf (Retrieved 4th May 2017).

[2] D. Faggella, "Everyday examples of artificial intelligence and machine learning", TechEmergence.com, 2017, Online: https://www.techemergence.com/everyday-examples-of-artificial-intelligence-and-machine-learning/ (Retrieved 4th May 2017).

[3] J. Twentyman, "Artificial intelligence moves from sci-fi to daily life", Financial Times, 2016, Online: https://www.ft.com/content/0213693e-81a7-11e6-8e50-8ec15fb462f4 (Retrieved 4th May 2017).

[4] P. Vasant (Ed.), "Handbook of Research on Modern Optimization Algorithms and Applications in Engineering and Economics". IGI Global, 2016.

[5] B. Pătruţ and C. Tomozei, "Agent technology in monitoring systems". International Journal of Computers Communications & Control, 5(5), 852-861, 2010.

[6] U. Kose, "Development of Artificial Intelligence Based Optimization Algorithms". (Turkish). Doctoral Thesis. Selcuk University, Dept. of Computer Engineering, Turkey, 2017.

[7] P. R. Cohen and E. A. Feigenbaum (Ed.), "The Handbook of Artificial Intelligence". Butterworth-Heinemann, 2014.

[8] P. Hamet and J. Tremblay, "Artificial intelligence in medicine. Metabolism". 69 (April 2017), S36-S40, 2017. DOI=https://doi.org/10.1016/j.metabol.2017.01.011

[9] H. Jaeger, "Artificial intelligence: Deep neural reasoning", Nature, 538(7626), 467-468, 2016, DOI=10.1038/nature19477

[10] R. S. Michalski J. G. Carbonell and T. M. Mitchell (Ed.), "Machine Learning: An Artificial Intelligence Approach". Springer Science & Business Media, 2013.

[11] A. I. Strong, "Applications of Artificial Intelligence & Associated Technologies". Science (ETEBMS-2016), 64-67, 2016.

[12] Oxford Living Dictionaries, "Diagnosis – Definition for the concept of diagnosis, in English", 2017, Online: https://en.oxforddictionaries.com/definition/diagnosis (Retrieved 5th May 2017).

[13] A. N. Ramesh, C. Kambhampati, J. R. T. Monson and P. J. Drew, "Artificial intelligence in medicine". Annals of The Royal College of Surgeons of England, 86(5), 334, 2004, DOI=10.1308/147870804290

[14] J. Korbicz, J. M. Kościelny, Z. Kowalczuk and W. Cholewa (Eds.), "Fault Diagnosis: Models, Artificial Intelligence, Applications". Springer Science & Business Media, 2012.

[15] E. Turban and L. E. Frenzel, "Expert Systems and Applied Artificial Intelligence". Prentice Hall Professional Technical Reference, 1992.

[16] J. P. Ignizio, "Introduction to Expert Systems". McGraw-Hill, 1991.

[17] A. Abraham, E. Corchado and J. M. Corchado, "Editorial: Hybrid learning machines". Neurocomputing, 72(13-15), 2729-2730, 2009. DOI=https://doi.org/10.1016/j.neucom.2009.02.017

[18] E. Corchado, M. Graña and M. Woniak, "Editorial: Editorial: New trends and applications on hybrid artificial intelligence systems".

Neurocomputing, 75(1), 61-63, 2012. DOI=https://doi.org/10.1016/j.neucom.2011.06.021

[19] S. Wermter and R. Sun, "Hybrid Neural Systems (No. 1778)". Springer Science & Business Media, 2000.

[20] A. Bahrammirzaee, "A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems". Neural Computing and Applications, 19(8), 1165-1195, 2010. DOI=10.1007/s00521-010-0362-z

[21] E. Corchado, A. Abraham and A. de Carvalho, "Editorial: Hybrid intelligent algorithms and applications". Information Sciences: An International Journal, 180(14), 2633-2634, 2010. DOI=https://doi.org/10.1016/j.ins.2010.02.019

[22] E. Comak, A. Arslan and I. Turkoglu, "A decision support system based on support machines for diagnosis of the heart valve diseases". Computers in Biology and Medicine, 37, 21-27, 2007.

[23] V. Vapnik, S. Golowich and A. Smola, "Support vector method for function approximation, regression estimation, and signal processing". Advances in Neural Information Processing Systems, 9, 281–287, 1997.

[24] U. Kose, G. E. Guraksin and O. Deperlioglu, "Cognitive development optimization algorithm based support vector machines for determining diabetes". BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 7(1), 80-90, 2016.

[25] OpenCV, "Introduction to Support Vector Machines". OpenCV – Tutorials – Docs.–OpenCV.org / Machine Learning", 2017, Online: http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html (Retrieved 7th May 2017).

[26] H. Tamura and K. Tanno, "Midpoint validation method for support vector machines with margin adjustment technique". International Journal of Innovative Computing, Information and Control, 5[11(A)], 4025-4032, 2009.

[27] T. Santhanam and M. S. Padmavathi, "Application of k-means and genetic algorithms for dimension reduction by integrating SVM for diabetes diagnosis". Procedia Computer Science, 47, 76 – 83, 2015.

[28] I. Javed, M. N. Ayyaz and W. Mehmood, "Efficient training data reduction for SVM based handwritten digits recognition". Electrical Engineering, 2007. ICEE'07. International Conference on (pp. 1-4). IEEE, 2007.

[29] C. Cortes and V. Vapnik, "Support-vector networks". Machine Learning, 20(3), 273-297, 1995, DOI= 10.1007/BF00994018

[30] L. J. Cao, K. S. Chua, W. K. Chong, H. P. Lee and Q. M. Gu, "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine". Neurocomputing, 55(1), 321-336, 2003, DOI=https://doi.org/10.1016/S0925-2312(03)00433-8

[31] M. Dorigo, M. A. M. de Oca and A. Engelbrecht, "Particle swarm optimization". Scholarpedia, 3(11), 1486, 2008, Online: http://scholarpedia.org/article/Particle_swarm_optimization (Retrieved 8th May 2017).

[32] J. Kennedy, "Particle Swarm Optimization". Encyclopedia of Machine Learning (pp. 760-766), 2011, Springer US.

[33] M. Clerc, "Particle Swarm Optimization (Vol. 93)". John Wiley & Sons, 2010.

[34] X. S. Yang and S. Deb, "Cuckoo Search via Lévy Flights. Nature & Biologically Inspired Computing", NaBIC 2009. World Congress on (pp. 210-214). IEEE, 2009.

[35] A. H. Gandomi, X. S. Yang and A. H. Alavi, "Cuckoo search algorithm: A metaheuristic approach to solve structural optimization problems". Engineering with Computers, 29(1), 17-35, 2013.

[36] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm". Journal of Global Optimization, 39(3), 459-471, 2007, DOI=10.1007/s10898-007-9149-x

[37] D. Karaboga and B. Akay, "A comparative study of artificial bee colony algorithm". Applied Mathematics and Computation, 214(1), 108-132, 2009, DOI=https://doi.org/10.1016/j.amc.2009.03.090

[38] X. S. Yang, "Firefly algorithm, stochastic test functions and design optimisation". International Journal of Bio-Inspired Computation, 2(2), 78-84, 2010.

[39] X. S. Yang, "Firefly Algorithm, Levy Flights and Global Optimization". Research and Development in Intelligent Systems XXVI (pp. 209-218), 2010, Springer London.

[40] U. Kose and A. Arslan, "On the idea of a new artificial intelligence based optimization algorithm inspired from the nature of vortex". BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 5(1-4), 60-66, 2015.

[41] J. Olesen, "The IHS Members' Handbook". USA, 2000.

[42] M. Dogan, A. Babanli and S. Telceken, "Classification of Headache Diseases with Artificial Neural Networks Methods". (Turkish). Symposium on Electric – Electronic and Computer Engineering (ELECO 2004), 424-427, 2004.

[43] G. Tezel and U. Kose, "Headache Disease Diagnosis by Using the Clonal Selection Algorithm". International Advanced Technologies Symposium (IATS 2011). 144-148, 2011.

[44] M. Duval, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases". IEEE Electrical Insulation Magazine, 17(2), 31-41, 2001.

[45] B. S. Yang, D. S. Lim and A. C. C. Tan, "VIBEX: An expert system for vibration fault diagnosis of rotating machinery using decision tree and decision table". Expert Systems with Applications, 28(4), 735-742, 2005, DOI=https://doi.org/10.1016/j.eswa.2004.12.030

[46] R. C. M. Yam, P. W. Tse, L. Li, and P. Tu, "Intelligent predictive decision support system for condition-based maintenance". The International Journal of Advanced Manufacturing Technology, 17(5), 383-391, 2001.

[47] N. Pombo, N. Garcia, K. Bousson and V. Felizardo, "Machine Learning Approaches to Automated Medical Decision Support Systems". Artificial Intelligence: Concepts, Methodologies, Tools, and Applications (pp. 1653-1673), 2017, IGI Global.

Cogeneration (combined heat and power production) in Europe

Sigrid Kusch
ScEnSers Independent Expertise
Ulm, Germany
sigrid.kusch@scensers.org

Abstract—**Implementation of cogeneration, which is also termed as combined heat and power production (CHP), achieves high energy efficiency degrees due to the simultaneous supply of electricity and heat. This paper explores the current situation of cogeneration throughout Europe and beyond. The technology today is state of the art. Despite clear environmental benefits, adoption of cogeneration remains significantly behind existing potentials (70 to 80% of the potential not used). In Europe, the share of electricity produced in cogeneration mode is higher than in other regions, and cogeneration has a rather long tradition. Nevertheless, the share of electricity produced in cogeneration mode in the European Union remains below 12%, and no significant increase of that share has occurred throughout the last decade. In absolute figures, cogeneration in the EU declined by more than 10% in the last 10 years. This work includes a discussion of factors that merit more attention, both with view to interpretation of cogeneration data and with view to more widespread adoption of the technology.**

Keywords- electricity; heat; cogeneration; combined heat and power production (CHP); energy efficiency; decarbonisation

I. INTRODUCTION

Cogeneration, or combined heat and power production (CHP), delivers both power (usually electricity) and heat simultaneously. In contrast to power-only schemes, integration of heat valorises energy that would otherwise be considered waste heat of power generation. This make better use of the energy carrier and consequently contributes to a decarbonisation (reduction of carbon dioxide emissions) of the energy sector. Decarbonisation of the energy sector has highest priority on the pathway to limit global warming to below 2°C: around two thirds of global anthropogenic greenhouse gas emissions are energy-related [1]. Increasing energy efficiency is one of the most effective means to reduce emission of the greenhouse gas carbon dioxide (CO2) [2].

In conventional power supply, which is predominantly based on thermal energy processes, the largest share of the energy content of energy carriers is lost in the form of waste heat, and only around one third of the energy content is delivered as electricity to the consumer [3][4]. In thermal energy processes, which consist of a multistep energy conversion chain (release of chemical energy contained in the energy carrier through combustion to reach thermal energy, generation of mechanical energy via heat engines and finally conversion into electricity), occurrence of waste heat as a by-product is unavoidable due to the underlying thermodynamic principles. Advanced technologies today can achieve up to 45% electrical efficiency [3]. If waste heat is valorised to generate additional electricity, the electrical efficiency might be higher than 50%, but in any case, a significant discrepancy exists between energy indeed made available as electricity and the energy content of used energy carriers.

At the same time, a high demand for energy in the form of heat exists worldwide. Heat is often given less attention then electricity or transport fuels, however, globally and in Europe, demand for heat exceeds the demand for electricity or for energy in the form of fuels for transport [3][4]. Heat demand, both worldwide and in Europe, is currently met primarily through the conversion of fossil fuels via specific heat supply equipment such as boilers [5].

A key characteristics of cogeneration schemes is that heat is valorised in addition to the recuperation of the main target product electricity. Heat valorisation can contribute to ensuring comfortable temperatures in buildings, to supply hot water, or to cover industrial heat demands. This replaces other sources of heat supply, and therefore, achieves net savings of energy resources and a decline in corresponding CO2 emissions [6].

Cogeneration can reach up to 90% energy conversion efficiency, and most common are schemes that convert 75 to 80% of fuel input into useful energy [3]. Since not all heat is indeed valorised in practice, the realised average efficiency of cogeneration units operated worldwide was assessed with 58% [7]. This is below the theoretical potential, but still significantly higher than efficiency degrees in electricity-only schemes. This translates into significant environmental benefits of cogeneration schemes, along with economic benefits, although precise quantification of the benefits is not a simple task and remains challenging [8][9].

II. CURRENT STATUS OF IMPLEMENTATION OF COGENERATION

Globally, around 10% of electricity is produced in cogeneration mode [7]. This share has not changed significantly over the past decade, even if absolute cogeneration has increased moderately (absolute consumption of electricity has also increased) [7]. Cogeneration has a tradition of more than 130 years: the first central power plant in the USA (New York City) started operation in 1882 in cogeneration mode (heat was supplied to nearby buildings).

Although the untapped potential is still huge (see section IV), many thousand CHP schemes are already in operation worldwide. Installations span a wide range of sizes (electrical

capacities), and they cover a diversity of fields of applications. They also include different technologies (types of engines, heat recuperation equipment) [8][10].

Installations range from large-scale CHP schemes (more than one megawatt power output, usually using gas turbines or steam turbines, generally custom-build according to the individual needs of the one specific site) to packaged small-scale units (few kilowatts power output up to 1 MW) that are available on the market with predefined specification and sizes. Large-scale CHP units are operated at central power stations or at sites with high energy requirement such as large industrial sites in the chemical industry or the oil-refining sector, while small-scale CHP units are common at smaller industrial sites or smaller commercial applications, in community settings, at hotels, hospitals, or in domestic settings [10].

Most CHP units are operated with fossil fuel, but operation on solid biomass or biogas is increasingly common. At biogas plants, installation of a CHP unit is standard, since part of the heat is usually used to maintain a favourable temperature in the anaerobic digester. In Europe, biogas plants without temperature control are very rare exceptions limited to specific single installations [11][12], therefore, the many thousand biogas plants in Europe (more than 10,000 in Germany alone [13]) are examples of manifold CHP installations in practice. Nevertheless, the best-known example of a cogeneration scheme remains the setting where a large, centralised power plant (usually operated on coal or natural gas, but wood biomass is also relatively common) is equipped with heat valorisation via a district heating network.

Implementation of cogeneration shows high variations among different countries and regions. Although the first cogeneration plant in the 19th century was located in the USA (see above), adoption of the technology remained at rather low level in the USA throughout the 20th century and until today. Rising electricity demand was answered by building very large power plants, predominantly operated on coal; therefore, plants could not suitably be placed in city areas [14]. Considering that heat cannot be transported over long distances, the large distance to potential heat consumers is one explanation why cogeneration is not widespread in the USA [14]. Another reason is a regularly framework that did not incentivise energy efficiency and/or cogeneration throughout most of the 20th century. The US Public Utilities Regulatory Policies Act of 1978 (public utilities were obliged to use electricity from renewable sources and from cogeneration) boosted cogeneration implementation, and the share of cogeneration in electricity supply rose to around 8% [15][16]. Nevertheless, cogeneration has remained less common compared to Europe.

In particular, Scandinavian and continental European countries were (and remain) highly active in implementing cogeneration schemes, which can partially be explained by higher fuel costs compared to other regions such as North America [16]. Another element is the dense population in European cities, with many people living in apartments rather than single houses, which facilitates heat distribution [14]. With the CHP Directive, the European Union (EU) formally incorporated cogeneration in 2004 into its energy policy and has established a framework to foster implementation of cogeneration schemes.

In the EU, the average share of electricity produced in cogeneration mode is slightly higher than worldwide, amounting to around 11% in 2015 (Figure 1) according to official Eurostat data [17][18]. However, the data reveal some fluctuation rather than a constant increase, and overall an increase did not occur when comparing the last available data (year 2015) to the situation 10 years earlier (2005). In this period (2005 to 2015), Eurostat data [20] indicate that EU electricity consumption fell from 1,131,937 GWh to 996,363 GWh (decline of around 12%). This means that in absolute figures, cogeneration in the EU also declined by more than 10%. This is in contrast to the global situation, where electricity consumption is rising, and with the rise in electricity consumption cogeneration in absolute terms also rises, even if the share of cogeneration shows a plateauing performance at some 10% worldwide [7].

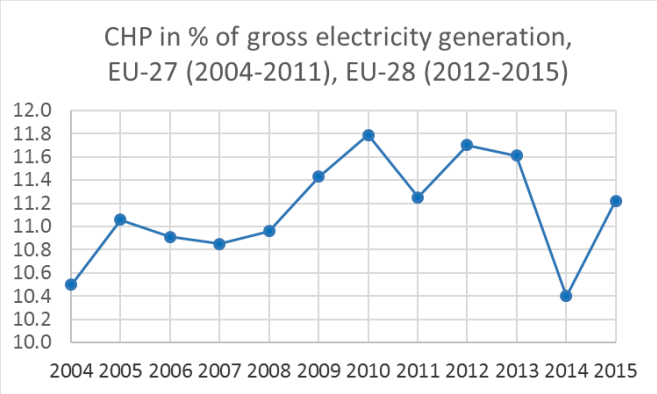


Figure 1. Share of electricity produced in cogeneration mode in the European Union (data source: Eurostat [17][18])

The fluctuations observed in Figure 1 can be assumed to be the result of highly different situations in the individual Member States of the EU. As an example, in Slovakia nearly 80% of electricity was produced in cogeneration mode in 2015, while in other countries, including Greece and France, the share was below 5% [17]. The situation in the EU Member States is shown with Figure 2 (countries with a cogeneration share below 10% in 2015 are not shown in Figure 2, which includes Romania: 8.45%, Sweden: 8.43%, and Spain: 8.08%, all other countries: below 8%). The five EU countries with the highest shares of cogeneration electricity, all with a share of more than 30% in 2015, are Slovakia (78.49%), Latvia (44.75%), Denmark (39.96), Finland (31.69%) and Lithuania (31.26%).

It is worth keeping in mind that the EU is a region with a particularly high diversity of countries, including size of countries and number of inhabitants, and consequently energy demands. Although Germany generates only 12% of its electricity in cogeneration mode, in absolute figures, it has the highest installed cogeneration capacity in the EU.

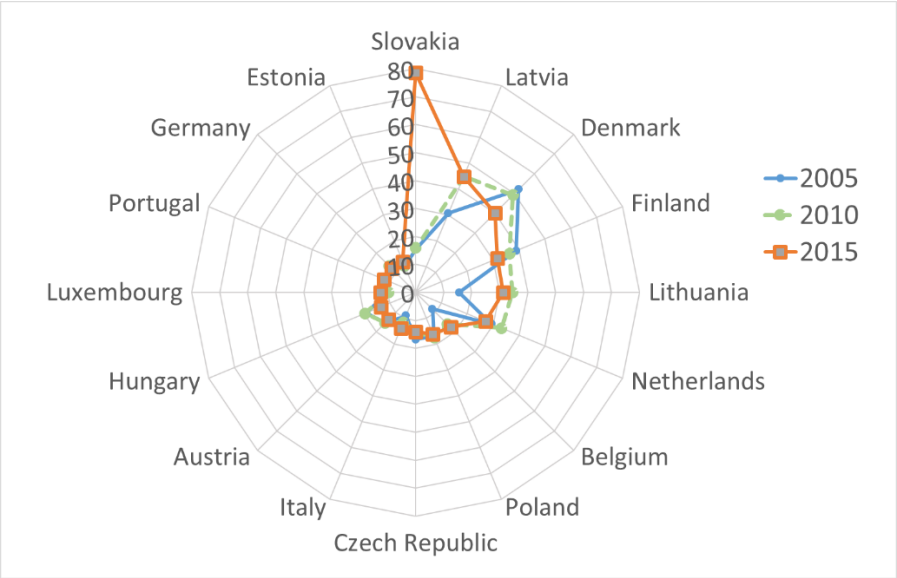


Figure 2. Share of electricity produced in cogeneration mode in EU countries in 2005, 2010 and 2015, in per cent of gross electricity generation (only countries with a share of more than 10 per cent cogeneration shown) (data source: Eurostat [17])

III. Interpretation of Cogeneration Data: Key Factors to be Considered

When interpreting cogeneration data, some factors need to be taken into consideration. Low shares of electricity produced in cogeneration mode can be the result of a generally low adoption of the technology in a country or region, however, such low shares might also primarily reflect specific elements of the prevailing power system. In France, for example, more than 75% of electricity comes from exploitation of nuclear energy [5]. Nuclear power plants are usually not equipped with heat valorisation schemes, therefore, the overall share of electricity produced in cogeneration mode is very low in the country, even if most thermal power plants are operated as cogeneration facilities. Similarly, an increasing share of electricity produced from renewables, which is given for example in Denmark, Italy and Austria, can drastically reduce the share of electricity produced in cogeneration mode: technologies such as solar energy or wind energy do not typically result in waste heat to be valorised (note that this is not the case for bioenergy: increased usage of biomass usually does generate waste heat, since energetic biomass valorisation is based on thermal energy conversion processes).

Therefore, it is important to be aware that cogeneration data usually refer to the total electricity produced in a country and not only to electricity generated in thermal power plants. Countries with high shares of electricity from renewable sources and/ or nuclear plants will usually score low in cogeneration statistics, even if they perform well in implementing cogeneration. One example is Finland, where hydropower is available at large scale and has been exploited since many decades: a major share of electricity generated in Finland comes from hydropower. At the same time, more than 80% of existing thermal power plants in Finland make use of cogeneration [3][8].

IV. Unused Cogeneration Potential

When considering that not all electricity supply solutions result into occurrence of excess heat, it becomes evident that the share of electricity produced in cogeneration mode is not usually a comprehensive indicator for the remaining cogeneration potential [8]. A lack of official statistical data on cogeneration implemented at thermal power plants, i.e. a lack of cogeneration statistics based on electricity generated in thermal processes only, hinders a precise assessment to what extent the cogeneration potential is indeed being used throughout different countries and regions.

Nevertheless, when considering that more than 60% of all electricity worldwide comes from fossil energy via thermal conversion processes [2][5], and still more than 40% in the EU [5], the currently low shares of energy generated in cogeneration schemes reveal a huge unused potential for more widespread adoption of the cogeneration technology, worldwide and in Europe. Therefore, it can be assessed that below 30% of the cogeneration potential in the EU is indeed being used, and below 20% of the potential worldwide, while the remaining 70 to 80% of the potential remains untapped.

Pathways to decarbonise the energy system give high priority to more widespread implementation of renewable energy schemes. As outlined above, such a transition results into less potential to integrate cogeneration. The goals to increase renewable energy and the goal to increase cogeneration unlock little synergy, but they do not exclude each other. On longer perspectives, cogeneration might be a technology of transition [19], and clearly the prioritised promotion of renewable energy schemes is highly justified. Nevertheless, fostering implementation of cogeneration is equally justified as long as the energy systems remain strongly based on thermal processes.

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

V. A Focus on Cogeneration in a Business Environment

Assessments how to increase uptake of cogeneration often focus on large, centralised power plants in combination with district heating [3]. Such settings are closely interlinked with urban planning and fall into the sphere that can be designed or at least significantly influenced by the public sector. While this area is relatively well studied, less attention is usually given to implementation in an industrial setting or in other decentralised schemes [8].

Implementation of combined heat and power production in an industrial setting clearly requires a different approach compared to cogeneration as a task for urban planning [8]. Elements to ensure successful implementation of CHP projects in a business environment include technical issues, efficient heat valorisation strategies, and successful management of economic challenges [8][21]. A CHP project in a business environment can be approached as part of an integrated environmental strategy of a company [8]. One obstacle is that funding programmes or other promotional programmes are often designed to address district heating or similar initiatives, while the industrial setting is usually less in focus, resulting in a situation where industrial initiatives or other decentralised projects fall between the regulations and might not meet eligibility criteria.

The central driver for CHP implantation in the industrial sector will be the generation of economic benefits. Therefore, programmes that aim at fostering cogeneration uptake should ensure that applications of all sizes (electrical capacities) in the industrial sector are included among the eligible projects. Another key element to encourage uptake of cogeneration by companies is access to knowledge and guidance (both on financial and technical issues), which is particularly beneficial for small and medium-sized enterprises (SME) who often lack human and financial resources to engage in cleaner production measures [8][22][23].

VI. Summary and Conclusions

Cogeneration is a powerful means to increase energy efficiency. However, around 70% of the potential to implement cogeneration in the EU remains untapped, and worldwide around 80% of the potential is not used. Most countries worldwide, and many countries in the EU, generate below 10% of their electricity in cogeneration mode. However, there are huge variations among countries. Five EU countries have shares of cogeneration electricity above 30%, with a maximum value reaching nearly 80%. Programmes fostering uptake of cogeneration often focus on district heating to valorise excess heat from centralised power plants. While that is one suitable application of cogeneration, its potential implementation in industrial settings or other decentralised schemes merits increased attention.

References

[1] International Energy Agency (IEA), Energy and climate change – world energy outlook special briefing for COP21, IEA publication, Paris, 2015.

[2] International Energy Agency (IEA), World Energy Outlook 2015, IEA publication, Paris, 2015.

[3] International Energy Agency (IEA), Co-generation and renewables, IEA publication, Paris, 2011.

[4] International Energy Agency (IEA), Heating without global warming, IEA publication, Paris, 2014.

[5] European Commission (EC), EU energy in figures – statistical pocketbook 2016, Brussels, 2016.

[6] U. Çakir, K. Çomakli, and F. Yüksel, "The role of cogeneration systems in sustainability of energy," Energy Conversion and Management 63, 2012, pp. 196-202.

[7] International Energy Agency (IEA), Tracking clean energy progress 2015 – Energy technology perspectives 2015 (excerpt IEA input to the Clean Energy Ministerial), IEA publication, Paris, 2015.

[8] S. Kusch, "An overview concerning combined heat and power production: a smart way to improve energy efficiency," CSE City Safety Energy 2-2015, 2016, pp. 132-141.

[9] M. Bianchi, L. Branchini, A. De Pascale, and A. Peretto, "Application of environmental performance assessment of CHP systems with local and global approaches," Applied Energy 130, 2014, pp. 774-782.

[10] Carbon Trust, Introducing combined heat and power (Technology guide), Carbon Trust, London, 2010.

[11] S. Kusch, and H. Oechsner, "Vergärung landwirtschaftlicher Substrate in Feststofffermentern," in Fachagentur Nachwachsende Rohstoffe (FNR) (Ed.), Trockenfermentation - Evaluierung des Forschungs- und Entwicklungsbedarfs. Gülzower Fachgespräche, vol. 23, 2004, pp. 105-113. (in German)

[12] S. Kusch, Methanisierung stapelbarer Biomassen in diskontinuierlich betriebenen Feststofffermentationsanlagen, PhD thesis, University of Hohenheim, Herbert Utz Verlag, Munich, 2007. (in German)

[13] Agentur für Erneuerbare Energien, website https://www.unendlich-viel-energie.de/, accessed 26 September 2017.

[14] B. Bridgeman, Combined heat and power (cogeneration). In: N. Cohen, and P. Robbins (Eds.), Green Cities: An A-to-Z Guide, Thousand Oaks, SAGE Publications, 2011, pp. 96-100.

[15] K. Ehrhardt-Martinez, and V. McKinney, Combined heat and power. In D. Mulvaney and P. Robbins (Eds.), Green Energy: An A-to-Z Guide. Thousand Oaks, SAGE Publications, 2011, pp.88-90.

[16] A.J. Waskey, Cogenerators. In P. Robbins (Ed.), SAGE Encyclopedia of Environment and Society. Thousand Oaks, SAGE Publications, 2007, pp. 298-299.

[17] Eurostat (2017a), Combined heat and power generation (official statistical dataset of the European Union), retrieved via http://ec.europa.eu/eurostat on 26 September 2017.

[18] Eurostat (2015), Combined heat and power generation (official statistical dataset of the European Union), retrieved via http://ec.europa.eu/eurostat on 27 October 2015.

[19] K.A. Kelly, M.C. McManus, and G.P. Hammond, "An energy and carbon life cycle assessment of industrial CHP (combined heat and power) in the context of a low carbon UK," Energy 77, 2014, pp. 812-821.

[20] Eurostat (2017b), Electricity consumption by industry, transport activities and households/services (GWH) (official statistical dataset of the European Union), retrieved via http://ec.europa.eu/eurostat on 26 September 2017.

[21] M. Köttner, S. Kusch, A. Kaiser, D. Dörrie, and D. Collins, Economic modelling of anaerobic digestion/ biogas installations in a range of rural scenarios in Cornwall. Cornwall Agri-Food Council, 2008.

[22] Umweltbundesamt (UBA) (German Federal Environment Agency), Cleaner production Germany (website), http://www.cleaner-production.de/, last accessed on 26 September 2017.

[23] United Nations Environment Programme (UNEP), Resource efficient & cleaner production (website), http://www.unep.org/recp/, last accessed on 26 September 2017.

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Using mobile devices while driving in Croatia – preliminary analysis

Dragan Peraković, Ivan Grgurević, Željko Šarić, Ivan Forenbacher, Siniša Husnjak, Ivan Jović, Ivan Cvitić, Gordana Kordić, Rosana Elizabeta Sente, Petra Zorić
University of Zagreb, Faculty of Transport and Traffic Sciences
Department of Information and Communication Traffic
Zagreb, Republic of Croatia, EU
nemobiteluuvoznji@fpz.hr

Abstract—**Research has shown that using a mobile device while driving increases the risk of traffic accidents by up to five times. From 2010 to 2014, Croatia has doubled the number of offenses caused using mobile device while driving. Current methods of prevention have succeeded in reducing the number of offenses (reduction of 6% in 2016), which justifies the investment in ways to improve prevention. Except the financial penalties, activities in the form of prevention and awareness raising about the negative impact of the use of mobile devices on driver behavior and the safety of road traffic in Croatia are insufficiently pronounced. The aim of this research is to determine the intensity of the use of mobile devices and to find to what extent would certain technical, technological and regulatory measures trigger a reduction in use.**

Keywords- mobile phones, drivers, safety, distraction

I. INTRODUCTION

Depending on the life situations and the type of job, people are accustomed to frequent use of mobile phone. The use of mobile device while driving allows the driver to immediately access the information that is essential to him. However, this ability comes with few negative effects for the driver as its one of the ways of driver distraction. According to currently available literature, there are several types of distraction (talking to in-car companion, radio device, navigational device, etc.), but the use of mobile device while driving has the most dangerous effect. This is because actions such as telephony, writing messages, or reviewing websites require visual, sound and cognitive attention of drivers, and hence the safety of road traffic gets disrupted.

Details from police reports of traffic accidents are included in 'Bulletin about road safety' published by the Ministry of the Interior of the Republic of Croatia – MUP. This Bulletin is the only official national statistical database about the number of traffic accidents and the most frequent causes of accidents, publicly available for the education of traffic participants about the risk factors in traffic. Although the previously mentioned available basis contains the most frequent causes of accidents, including the category of the roads, there are some accident causes that are not easily identified. This is primarily the case of the use of all types of mobile devices while driving (mobile phones, smart watches, embedded information and entertainment systems, etc.).

The most pronounced risk factors related to the problem of mobile phone influence on driving are insufficient awareness and the fact that the use of mobile phone while driving is an increasing trend that is considered as common. This research is presented and financed by the project of The National Road Safety Programme of the Republic of Croatia 2011 – 2020, under the title 'The impact of mobile device usage on drivers' behavior'. The following chapters present developed research methodology and the preliminary research results gathered through questionnaires form Q1 and Q2 stages of the project. The aim of the project is the introduction of preventive actions for raising the collective awareness, informing and drivers education about the negative effects of the use of mobile devices while driving. During the project, the objective is to study driver's behavior in real-life situations and in car-test polygon environment. The research results will provide the ground for forming the national informative Web 2.0 portal as the media intended for educating on the dangerous effects of using mobile device while driving.

II. PREVIOUS RESEARCH

Drivers who use a mobile device while driving are distracted for 5 seconds in average, i.e. do not follow the road in front of them, and by driving at a speed of 50 km/h exceed 60 meters of road completely unaware of the traffic. Due to this reason, 3,328 drivers were killed in the United States – US in the year of 2012, while the 48% of US drivers stated that they regularly use their mobile devices while driving [1]. That is why it is extremely important to warn all traffic system participants on the sources of threats and to form informative and educational actions for reducing the threat and strengthen the safety.

Research results from Canada have shown 52% of respondents continue to use mobile phone while driving, regardless the awareness of the driver about the possible road safety threats. About one fifth (1/5) respondents answer their phone immediately no matter of the present situation in traffic [2]. Respondents who use their phone while driving and the ones who don't, both, agreed that the using of mobile phone for writing and reading messages while driving is very dangerous, and that the use of mobile phone while driving is as dangerous as driving drunk [3].

In 2013 distracted driving was the cause of 2,910 deaths in car accidents in the US, from which 14% was caused because the use of mobile phone while driving, i.e. 455 lives were lost due to the mobile phone usage behind the wheel. Recordings

from police reports say that drivers where having phone conversations or used their mobile device for listening to music or some other possibility when the accident occurred. According to the data of this research, drivers no younger than 20 and no older than 29 use mobile phone while driving in 38% of cases [4].

One of the chapters within the research contains drivers as the subject of research. Through the research of mentioned safety phenomenon, the authors strive to find out if negative potentially socially dangerous deriver behavior is the characteristic of the mass. The conclusion of the research brings out the fact that more than half of drivers are being distracted while driving because of the use of mobile phone [5]. The authors state that young and inexperienced drivers, belonging to the 'web population', make the most endangered group. These drivers intensively use sophisticated applications, read and write messages, take photographs and make videos while driving and therefore are labeled as the 'mobile killers'.

Because of these distractions, drivers monitor the ongoing traffic situation four times less while driving [6]. Similar results are given within simulated traffic environment where young drivers were using social network applications on their mobile device while driving [7].

The research on psychological factors of mobile phone addiction results with conclusions of positive correlation on mobile phone usage while driving. According to stated, drivers between 17 and 25 years of age are more likely to use mobile phone while driving than older ones.

Also, it is more likely that driver would use his phone while waiting on traffic light at road intersection than while driving at speed of 100 km/h. The research found no difference in mobile phone usage regarding drivers who are not in a hurry or are turning late. Driving conditions have had more impact on the intent of using mobile devices while driving than the motivational factors of respondents [8].

The research made in United Kingdom - UK has shown that the combination of youth and inexperience poses a great risk to safety of young drivers. Young drivers need a high level of concentration on the current situation in traffic which makes them more prone to distraction risks. Also, young drivers use mobile devices while driving more often. Even 19% of young drivers are messaging at least once in a month while driving while older drivers are included in such activities in 11% of cases [9]. Unlike the UK, research from the USA shows even more disturbing numbers. Even 80% of young drivers in USA use a mobile device while driving to make calls, and 72% for writing messages [10].

In 2016, 32,757 traffic accidents occurred in Croatia [11]. However, the number of traffic accidents caused by using mobile phone while driving and the factors of phone usage affecting the traffic accidents remain unknown. The primary issue is a hardly possible subsequent identification of mobile phone usage as a main cause of the traffic accident. Participants of traffic accidents usually don't mention the use of mobile phone while driving and the associated distraction. Mentioned circumstances are the reason why Croatian police registers often don't contain accurate and/or detailed data on the accidents causes by mobile phone usage. Drivers involved in passenger and transit traffic, including the truck and bus drivers, are also exposed to high safety risks due to the use of mobile device while driving. There is a space for improvement of the safety of for mentioned drivers, but in the Republic of Croatia it has been done very little. Due to insufficiently specified insight into the complexity of this problem and the lack of awareness of traffic participants, there is a need for an easily accessible and up-to-date information source where users form awareness of the negative impact of using mobile devices when driving [12], [13]. Proposed novelty should help in decreasing the number of people killed in traffic accidents for 50% by 2010 [14].

The fines are one of the methods used for raising the awareness about the negative impact of mobile phone usage while driving. Considering the fines are not most efficient methods, as stated by, it is necessary to act proactively in the terms of education [15], [16].

III. METHODOLOGY OF RESEARCH

The most prominent risk factors related to the problem of the impact of the use of mobile devices on driver behavior during driving covered by this research are:

- Today's employers such as carriers, taxi services, driver-renters and ordinary drivers imply the use of a motor vehicle as a mobile office, and assuming the significant use of mobile device while driving,
- The fact that trend of mobile devices owning also reflects in driving,
- Insufficient level of awareness regarding the problem,
- Irresponsibility of the wider community that is insufficiently informed about the problem,
- Lack of a national information portal for broader education, but also narrowly defined groups of users such as professional drivers, ordinary drivers as well as the employers,
- Inadequate engagement on prevention despite prescribed financial penalties (i.e. fines) and the fact that research has shown that this does not necessarily have a greater positive effect.

Considering previously mentioned this project's focus is directed to the research, dissemination of research results, education and the transfer of knowledge.

Research on the frequency of use of mobile devices while driving was carried out by the method of counting. Measurements are made at a few relevant locations in the city of Zagreb, Zagreb County and the city of Opatija.

Counting was made at city road points of higher traffic density, in different days of week and at different hours. The measurement was made on two different ways. One part of measuring is performed by project collaborators wearing project's official T-shirts who were therefore recognized as the project participants. Another part of measuring is performed by project collaborators dressed as common passengers (Figure 1).

Figure 1. Conducting of counting on two different ways

During the first way of measuring, drivers have easily spotted that something is going on and therefore have adjusted their behavior to unexpected circumstances. In the second way of measuring, drivers were completely unaware of conducted measuring and therefore haven't changed their behavior, i.e. drivers have used their phones more freely and therefore more of them was spotted.

The activities of presenting the results of the research, informing and drivers' education are conducted in: Autoclub Siget, Rijeka and Zadar, within the International Scientific Conference ZIRP-LST 2017, in the Ministry of the Sea, Transport and Infrastructure in Zagreb and on the seminar Technical analysis of traffic accidents 2017 at the Faculty of Traffic and Transport Sciences in Zagreb. Distribution of relevant information, i.e. informing of drivers is conducted on the city roads of Zagreb as well, in cooperation with several companies supporting the research labeled as the 'friends of the project'. Promotional materials, such as flyers containing educational information and facts, are frequently distributed.

The increasing number of tourists on state roads is defined as risk factor for traffic safety within the Country. Therefore, a special campaign has been prepared within the project to inform the tourists on dangers of mobile phone use while driving on social media and at toll station Lučko at the time of highest traffic in seasonal jams (Figure 2).

Figure 2. Informing the tourists at the Lučko toll station

Social networks are counting increase in the number of users daily and present a low-cost form of advertising and content sharing. Regarding to mentioned fact, this project is promoted through four most popular social networks in Croatia: Facebook, Instagram, Twitter and YouTube. In addition, there is an informative website of project formed to provide relevant information about the project, such as the overview of past and planned activities.

By promoting through Facebook social network, during the first 6 months, there were a very high number of 1,091,891 accesses to the project's Facebook page 'NE mobitelu u vožnji'. Out of the total number of accesses, the page has been assessed by 355,246 unique users. Many campaign posts were made, of which 18 were financially supported promotions. The results of the most successful campaign posts, ranked from the latest to the oldest, are shown in Table 1, and the image of A and B posts can be seen in Figure 3.

TABLE I. STATISTICS OF PROMOTING THROUGH FACEBOOK SOCIAL NETWORK

CAMPAIGN POST	NUMBER OF VIEWS	NUMBER OF REACTIONS, COMMENTS AND SHARES	NUMBER OF CLICKS
A	60,410	1,790	1,643
B	49,052	1,605	997
C	34,385	994	467
D	95,727	2,152	3,317
E	103,112	2,280	4,354
F	212,347	5,725	7,285

Figure 3. Image of A and B posts published in social networks (Facebook)

Four different surveys have been conducted on four different stakeholders, i.e. users:

- Domestic population - web questionnaire and field survey using the tablets on relevant locations,
- Taxi drivers – field survey using tablets and the help of interviewers in the rooms of taxi companies and associations (Figure 4),
- Rent-a-car users – field survey using tablets and the help of interviewers,
- Foreigners – web questionnaire adjusted to tourists that are having vacation in Croatia.

Figure 4. Conducting the survey among taxi drivers

By conducting the survey, the research gathered personal attitudes and experiences of the respondents on the use of mobile devices while driving.

IV. RESULTS

This paper contains preliminary results of counting the frequency of use of mobile devices and preliminary results of questionnaire on domestic population conducted within the first six months of project duration.

A. Results of counting the frequency of use of mobile devices

Since results obtained from questionnaires depend on the level of respondents' sincerity, the method of measuring the frequency of use of mobile device has given different results from ones obtained through questionnaires.

The Figure 5 shows locations of conducting measuring in Zagreb. Locations have been marked with letters from A to N, each containing the obtained percentage of drivers who were using their phone while driving.

According to results obtained by this method of research, drivers were using their phones while driving in 8.9% of cases. The usage frequency ranges from 5.6% (location G) to 13.3% (location D), in dependence of the part of the town. At the far north of the town (location B) and at the far east of town (location N) there is an equal number of noticed mobile phone users while driving (5.9%).

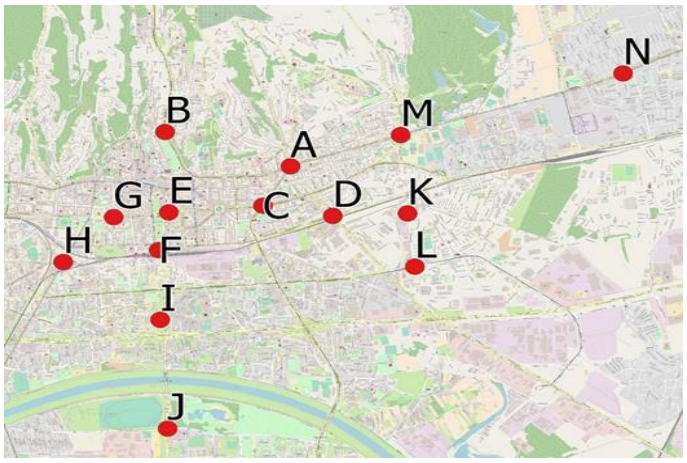


Figure 5. Locations of counting the frequency of use of mobile devices in Zagreb

Additionally, locations A and F show the frequency of usage of 9%, while locations L and M show 10.5%. Followed by location H with 10.2% of usage, the location I with 9.2%, the location C with 8.7% and the location K with 8.1% of usage. Also, there is undeniable percentage of usage in amount of 7.5% at the measurement point E (6.9%). The average rate of mobile phone usage is 8.6%.

B. Results of questionnaires on domestic population

The survey conducted on the entire state population through web questionnaire and a field survey on relevant locations in a six months duration has included 795 respondents. The most of respondents were between 18 and 25 years old (20%), and 13% of respondents were from age 25 to 29. Besides, 16% of respondents were 35 to 39 years old, 10% were 40 to 44 years old, and 6% were 55 to 59 years old. Among respondents, 70% were men and 30% were women.

Out of the entire set of respondents, 97% own driver's license and are drivers on daily basis. Only 3% of respondents do not use mobile phone for any kind of related activities while driving, and other 97% of respondents do.

The most of respondents who have confirmed to use mobile phone while driving use their phone when waiting at traffic lights (71%), while even 67 % of respondents uses mobile phone during the entire driving time.

Given the fact that traffic jams are common in city environment, 60% of respondents shorten their waiting time with their mobile phone, and 29% of respondents use their mobile phone when waiting at toll stations.

The Figure 6 shows the frequency of mobile phone usage while driving regarding individual mobile phone services. It is visible that 15% of mobile phone usage while driving is due to making phone calls, i.e. mobile conversation.

Also, interesting fact is that 89% of respondents claim to never play games on the mobile phone while driving.

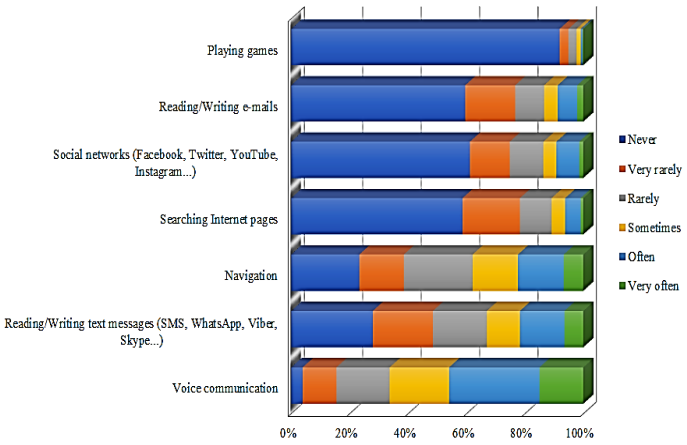


Figure 6. The frequency of mobile phone service usage while driving

Voice communication is still the main reason of using a mobile phone while driving (86%), and the navigation and texting (SMS, WhatsApp, Viber, Skype) are equally present (44%). Social networking is the reason why 15% of respondents uses their mobile phone, and followed by e-mailing with 13%. Playing of mobile games while driving isn't overly popular with respondents (1%).

The user's opinion on the impact of activities while driving shown in Figure 7 were also investigated, regarding making calls, using navigation or social networks. Respondents believe that gaming on a mobile device during driving significantly affects their reaction time (49%). In an equal percentage of 24% of respondents, searching the web, social networking and e-mailing is considered not to affect the time of drivers' reaction.

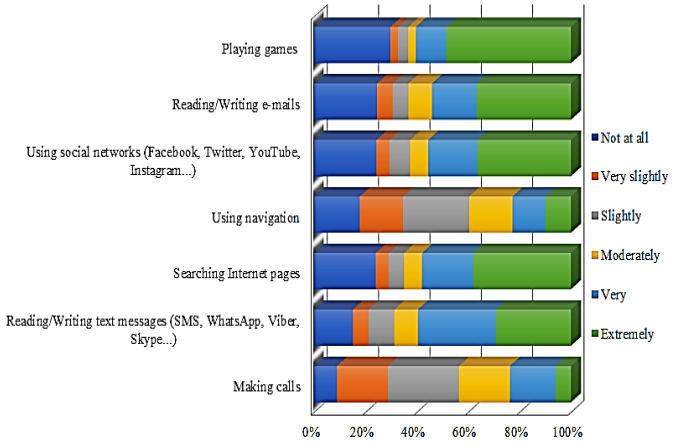


Figure 7. Considered effect of individual mobile phone activity while driving

Respondents are highly aware (91%) that their mobile phone usage while driving effects their own safety and the safety of other traffic participants, but the number of those who claim not to even think about it (4%) shouldn't be ignored. Since they are aware of the possible safety dangers, most respondents usually adjust the driving speed when using mobile phone behind the wheel (59%). Many respondents, up to 55%, believe that they can react quickly enough to the new situation and that their use of mobile devices will not endanger the traffic. As many as 34% of respondents say they use handsfree because they are aware of the dangers, and 4% of users justify their behavior claiming the others use their phone while driving as well, i.e. claim to follow the example of other drivers.

Many respondents (72%) believe that legal sanctions regarding the use of mobile devices while driving should be stricter. Higher fines would greatly encourage respondents not to use mobile devices while driving (20%). 28% of respondents find that they would be much encouraged by a stricter law and dense police control. The education about the negative effect of using mobile devices while diving is found relevant by 15% of respondents.

V. Discussion and Conclusion

The use of mobile device while driving is one of the many types of driver distraction and is considered a new era disease. Given the fact that activities such as telephony, writing messages or browsing websites require visual, sound and cognitive attention of drivers, the road traffic safety is compromised. The official publicly available national statistical database has identified main causes of traffic accidents. However, there are causes that are not easy to record, such percentage of the use of all types of mobile devices while driving.

Many worldwide conducted studies show that a large majority of drivers regularly use mobile devices while driving. Drivers are aware of the dangers that may occur, but continue with the usage of mobile phone in traffic. Young and inexperienced drivers belonging to the online generation are considered the riskiest driver group. This lies in the fact that they use the most sophisticated applications, read and write messages, take photos and make videos while driving, also it takes more time for them to react then for experienced drivers.

In addition to the financial penalties used by many states, drivers need to be educated about the dangers and consequences of using mobile devices while driving.

Presented questionnaire results show that 97% of drivers in Croatia use their mobile phones while driving. The results of counting have given results of 8,9% of drivers who use their phone behind the wheel. Respondents most often use their phone for talking behind the wheel, while texting is considered to noticeable affect the time of reaction. Stricter law, higher fines, more dense police control and higher level of education are considered to encourage reduction in the use of mobile devices while driving.

The obtained results indicate the necessity of further research on the issue of using mobile devices while driving to increase driver awareness and the traffic preventive.

Acknowledgment

This research is conducted within the entitled 'The impact of mobile device usage on drivers' behavior' lead by Department of Information and Communication Sciences of the Faculty of Transport and Traffic Sciences in collaboration with Croatian chamber of transport engineers. Project is

finances from The National Road Safety Programme of the Republic of Croatia 2011 – 2020 of the Ministry of the Interior of the Republic of Croatia – MUP.

REFERENCES

[1] National Highway Traffic Safety Administration, "National Telephone Survey on Distracted Driving Attitudes and Behaviours - 2012, "National Center for Statistics and Analysis, Washington, USA, April 2013.

[2] A. S. Nurullah, J. Thomas, F. Vakilian, "The prevalence of cell phone use while driving in a Canadian province," *Elsevier*, no. Transportation Researh Part F 19, pp. 52-62, 2013.

[3] P. C. Burns, A. Parkes, S. Burton, R. K. Smith, D. Burch, How dangerous is driving with a mobile phone? Benchmarking the impairment to alcohol, Wokingham, United Kingdom, Transport Research Laboratory, 2002.

[4] National Highway Traffic Safety Administration, "Distracted Driving 2013," National Center for Statistics and Analysis, Washington, USA, April 2015.

[5] V. Obradović, Risky Delinquent Criminal Behavior: Empirical Studies (cro. Rizično delinkventno kriminalno ponašanje: empirijske studije), Zagreb, Politička kultura, 2014.

[6] S. G. Hosking, K. L. Young, M. A. Regan, The Effects of Text Messaging on Young Novice Driver Performance, Mealbourne, Australia, Accident Recearch Centre, 2006.

[7] D. Basacik, N. Reed, R. Robbins, "Smartphone use while driving: a simulator study", Wokingham, United Kingdom, EU, Transport Research Laboratory, 2012.

[8] S. P. Walsh, K. M. White, B. Watson, M. K. Hyde, "ATSB Research and Analysis report Road Safety Research Grant Report: Psychosocial factors influencing mobile phone use while driving," Australian Government, Canberra, December 2007.

[9] M. Haque, S. Washington, "A parametric duration model of the reaction times of drivers distracted by mobile phone conversations", *Accident Analysis and Prevention*, vol. 62, pp. 42-63, January 2014.

[10] J. Ehsani, A. Brooks, K. Li, J. Perlus, A. Pradha, B. G. Simmons-Morton, "Novice Teenage Driver Cell Phone Use Prevelance," in *International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Maryland, USA, 2014.

[11] Ministry of the Interior of the Republic of Croatia, Department of Strategic Planning, Analysis and Development, "Bulletin about road safety 2016.", Zagreb, Republic of Croatia, EU, Ministry of the Interior of the Republic of Croatia, 2017.

[12] D. Peraković, I. Jovović, R. Magušić, "An Influence Study of the Use of Modern Communication Equipment on Traffic Safety," in *15th International Conference on Transport Science ICTS 2012*, Portoroz, Slovenia, 2015.

[13] R. Magušić, V. Obradović, "The occurrence of distraction by using cell phone while driving (cro. Distrakcije do kojih dovodi korištenje mobitela tijekom vožnje)," in *Proceedings of 22nd Ramiro and Zoran Bujas Days Conference*, Zagreb, Croatia, 2015.

[14] Government of Republic of Croatia, "National Road Safety Programme of the Republic of Croatia 2011 – 2020", Zagreb, Republic of Croatia, EU, Narodne novine, 2011.

[15] B. Sampaio, "On the Identification of the Effect of Prohibiting Hand-held Cell Phone Use while Driving," *Transportation Research*, vol. 44, no. 9, pp. 766-770, November 2010.

[16] World Health Organization (WHO), "Mobile phone use: A growing problem of driver distraction," World Health Organization Press, Geneva, 2011.

Methodology to define the use of liquid natural gas in the short sea shipping in Europe using data mining techniques

Beatriz Molina Serrano
Civil Engineering. Transport
Universidad Politécnica de Madrid
Madrid, Spain
beatriz.molinas@alumnos.upm.es

Nicoletta González-Cancelas
Civil Engineering. Transport
Universidad Politécnica de Madrid
Madrid, Spain
nicoleta.gcancelas@upm.es

Francisco Soler-Flores
Civil Engineering. Transport
Universidad Politécnica de Madrid
Madrid, Spain
f.soler@upm.es

Abstract—**Present research presents a methodology which has been developed to determine the relationship between the variables that define the use of Liquefied Natural Gas in Short Sea Shipping in Europe, through the use of data mining techniques. The research takes place in the European space, which includes data from 30 countries (28 members of the European Union plus Norway and Iceland). Bayesian network is constructed with 35 indicators selected, which are classified into 5 different categories: international trade and transport, economy and finance, population and social condition, environment and energy, and institutional and political. Results show that variables RCU (capacity of LNG regasification terminals under construction) and MSFT INW (modal distribution of cargo transport by inland waters) are the two root nodes of the network. In addition, the variables of transport and international trade category and economy and finance category become the most important in the decision to implement LNG as fuel in the maritime stretch of motorways of the sea. However, environment and energy variables and population and condition ones are the most dependent on the network.**

Keywords- liquid natural gas, short sea shipping, data mining

I. INTRODUCTION

Congestion in some sectors of the European road network grew in the 1990s, as a result of the growth of international trade and the entry of new members into the European Union (EU). It was as important as even it threatened economic competitiveness because it included additional costs into the productivity of the companies. Consequently, land transport congestion and its externalities, such as pollution and injuried accidents, show the need to create a continental strategy. So, it was proposed to carry the most of goods by sea. Main advantages of this alternative are energy efficiency, little damage to the environment, safety and reduction in infrastructure saturation [1].

Promoting maritime transport was expected to optimize modes of transport, decongest roads, reduce carbon dioxide (CO_2) and sulfur oxides (SO_x) emissions that affect the environment and human health, create new routes between European or neighboring ports and provide safe transport and quality. Likewise, it was hoped to guarantee the connection between the sea, the waterways and the railroad, through intermodality. In addition, it was hopped to develop the motorways of the sea [2].

However, shipping sector generated a lot of pollution on European coasts, so it was necessary to establish an energy policy proposing solutions on supply the effects on growth and on European environment [3]. In this way, maritime transport becomes a complement to the goods transport by road, managing to increase the performance and capacity of the system in general. The RECORDIT [5] and REALISE [6 and 7] projects demonstrated that intermodal transport reduces total costs (internal and external ones), making it a better economic alternative when compare it to unimodal transport [8 and 9]. So, European Commission promotes Short Sea Shipping (SSS) as an alternative to unimodal transport.

Nevertheless, implementation of the SSS brought with it pollution in coastal zones, due to carbon dioxide (CO_2) and sulfur oxides (SO_x) emissions, which was generated by fuel consumption [10]. For this reason, Annex VI of the International Maritime Organization (IMO) MARPOL convention laid down rules for the prevention of air pollution, which allowed for 0.1% sulfur content by 2015. Fuel in SECAs and internationally is expected to be 0.5% by 2020 [11].

Psaraftis and Kontovas [12] affirm that there are several ways to reduce greenhouse gases such as having low consumption engines, find optimum propulsion, use alternative fuels or purifiers in the exhaust, etc. Other interesting contributions have been made on the environmental impact of SSS and policy recommendations on climate change, including

the developments of Corbett, Fischbeck and Pandis [13], Endresen [14], Corbett and Koehler [15], Corbett, Wang and Winebrake [16] and Devanney [17], Devanney and Beach [18]. However, there is an absence of developments on the financial impact of regulation in the maritime transport sector [19].

Studies and research about fuel used in SSS in Europe, that could meet the environmental requirements for sulfur content, are basically concentrated in three: Heavy Fuel Oil (HFO), Marine Gas Oil (MGO) and Liquefied Natural Gas (LNG). The first one is the most used nowadays despite having a sulfur content of at least 1% of the mass. Although the latter two ones comply with the regulations. The latter one is the best alternative not only to comply with the strict regulations for sulfur oxides (SO_x) emissions, but also to meet those related to Nitrogen Oxides (NO_x) and carbon dioxide (CO_2) [4].

Regarding greenhouse gas emissions, LNG produces slightly less than one generated by diesel fuels and is only 1% less efficient. However, its implementation requires significant investment in medium and small-scale. LNG need to be carried to bring gas to intermediate terminals or storage tanks, and its strict environmental regulation challenges traditional forms of maritime governance [20, 21, 22 and 23]

Comparing the SSS fuels mentioned above, it is possible to conclude that LNG is the best alternative due to the availability of infrastructure within the geographical limits, at current market-adjusted price, reduction of Europe's dependence on oil, and mainly reduction of environmental impacts [20 and 24]. So, briefly, the use of LNG generates a positive impact on environment and strengthen the motorways of the sea, reinforcing existing routes and fostering the creation of new ones [23].

II. METHODOLODY

The methodological proposal followed is divided into three stages as it is shown in Figure 1.

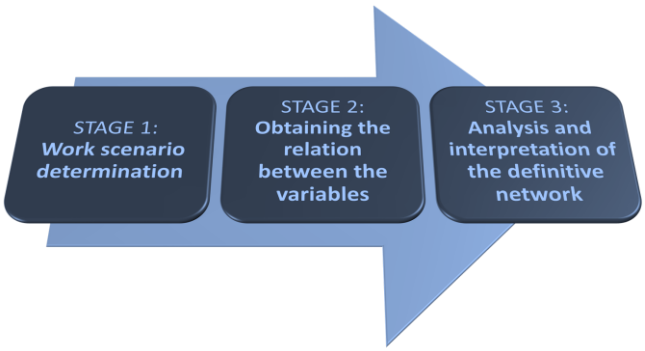


Figure 1. Methodology scheme. Source: Own production

Stage 1: Work scenario determination

The following is a brief explanation of each of the stages of the methodology.

For the determination of the work scenario, the previous state-of-the-art review is taken as a basis. From it, it was possible to establish the categories of indicators that make up the database and the sources from which they were consulted.

The geographical delimitation for the present study is 28 members of the European Union plus Norway and Iceland, because their location and because they are part of the SECA of the North Sea.

Determination and Obtaining of Study Variables

The indicators considered were selected according to their direct or indirect relationship with LNG and the motorways of the sea; also included some that have no apparent relationship with the topics to be analyzed. In total, it was decided to include 35 indicators, drawn from Eurostat, from the European Gas Infrastructure Association (GIE) and the European LNG terminal infrastructure 2015: Status and Outlook document.

Variable classification

With the objective of observing the relationship between the variables that affect the motorways of the sea and LNG, variables were classified into 5 types of indicators according to their nature: international trade and transport, economy and finance, population and social condition, environment and energy, and institutional and political. The set of variables and their category are presented in Table I.

TABLE I. LIST OF VARIABLES

Category	Indicator	Variable	Unit
Transportation and international trade	Maritime Transport	MTG	Thousands of tones
	Modal distribution of freight transport by rail	MSFT RAIL	Percentage
	Modal distribution of freight transport by road	MSFT ROAD	Percentage
	Modal distribution of freight transport by continental waters	MSFT INW	Percentage
	TEUs in SSS	SSS VTEU	1000 TEU's
Economy and Finance	Gross domestic product	GDP	Thousands of euros
	Exports from member states to third countries	EGS	Thousands of euros
	Imports from third countries to member states	IGS	Thousands of euros
	Debt, Public surplus	GOVDEF	% GDP
	Production in industry	PROIND	Index, 2010=100%
	Production in construction	PROCON	Index, 2010=100%
	Participation in imports by EU member	SHAIMPT	% (EU=100%)
	Participation in exports by EU member	SHAEXPO	% (EU=100%)
Environmental and Energy	SOx emissions from transport and communication	AETSO	Tones
	Emissions of greenhouse gases from transport	GHGTR	Thousands of tones
	Environmental taxes on transportation	ETXTRANS	Thousands of euros
	Total generation of waste	GENWAS	Tones
	Prices of natural gas for industrial use	GPICONS	Euro/Kilowatt-hour
	Waste electrical and	WELEC	Kilograms per

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Category	Indicator	Variable	Unit
	electronic equipment produced by households		capita
	Productivity of resources	REPRO	Euro/Kilogram
	Emissions of greenhouse gases from agriculture	GHGAGR	Thousands of tones
Institutional and Politic	Sulfur Emission Control Areas	SECA	Number
	Geographical location of the coasts	SFRONT	Number
	Capacity of LNG regasification terminals in operation	RCO	billons m³ per year
	Capacity of LNG regasification terminals under construction	RCU	billons m³ per year
	Capacity of LNG regasification terminals in plans	RCP	billons m³ per year
	Form of government	FDG	Number
	Models of port management in Europe	EPG	Number
Population and social condition	Total expenditure on research and development	RDEXPN	Euro/habitant
	Population	POBL	Habitants
	Life expectancy	LIFEEXP	Age
	Population at risk of poverty or social exclusion	RISKPOV	Population percentage
	Unemployment rate	UNER	Active population percentage
	Students enrolled in tertiary education	STENTER	Population
	Expenditure on social protection	EXPSPBE	Euro/habitant

Generation of the database

Database was generated with 35 indicators: 28 belong to Eurostat, 3 indicators were consulted in the "European LNG terminal infrastructure 2015: Status and Outlook of the Gas Infrastructure Association in Europe (GIE)" and the last 4 ones were built (own production).

Stage 2: Obtaining the relation between the variables

This section shows the process to create and carry out the Bayesian network interpretation.

Generation of a Bayesian network from database

Bayesian networks represent, in a compact way, a multivariate probability distribution. Technically, it is an acyclic directed graph where each of its nodes refers to a random variable. Relationships between variables are encoded in the graph structure itself according to the criterion of separation. Each node is associated with a probability distribution conditioned to the parents of that node. Equation (1) shows a network with n variables X1, X2, ...Xn [25].

$$p(x_1, ..., x_n) = \prod_{i=1}^{n} p(x_1 \mid x_{pa(i)}) \quad (1)$$

Probabilistic networks allow the automation of the probabilistic modeling process by using the expressivity of the graphs. The models resulted from them have both the results of graph theory: relations between the set of variables, and probability (quantification of these relations).

Rodríguez and Dolado [25] assert that knowledge engineering in Bayesian Networks is the process of construction, validation and use of Bayesian networks. This is done by means of the definition of graph structure and the parameters of the network. So, it is necessary to follow six steps: to know the domain to model, to select the variables of interest, to define a type for each variable, to set the network topology, to build probability tables for each node, and finally to evaluate and verify the Bayesian network.

The first step is to obtain additional information from the available data through an experiment in the subject or an additional process [26]. Secondly, it involves carefully selecting the set of variables. So, it enriches the analysis without adding unnecessary complexity to the study. In the third step, each variable must be typed in Boolean, label or numeric, and if necessary, its discretization must be defined. In the topology of the network, it is necessary to establish the relations between the variables, being preferential the causal relations. This causality implies a correlation, but not the reverse. The fifth step is the construction of the probability tables for each node and the last one consists on analyzing the sensitivity, where variation of the values introduced can be observed.

The use of Bayesian networks for the graphical representation of the relations between variables facilitates the understanding of the causal relations through the graph [27]. Other advantages are: it allows simultaneous representation of the quantitative dimension (probability tables) and qualitative one (graph) of a problem, it can be worked with incomplete data [27 and 28], it is possible to perform a sensitivity analysis [25 and 29] and finally, it can make predictions with incomplete information, among others.

However, use of this kind of networks presents some difficulties in their construction, such as the computational cost of exploring a previously unknown network, the quality and scope of the initial parameters that can distort the results, among others Niedermayer [30].

Discretization of variables

After selecting variables, it is necessary to discretize them because in Bayesian networks it is common to use discrete or nominal variables. The development of the application allows to implement algorithms in search of that objective like, for example, the search 'glotona' (hill-climbing).

Construction of the model

Structural learning is emphasized in finding the relations of dependence between the variables. In such a way, it is possible to determine the topology of the Bayesian network. Depending on the type of structure, application of different methods of structural learning is possible: multiconnected networks, learning trees, learning poly-trees, learning methods based on

measures and search, and even, methods based on dependency relationships.

Inference, classification and validation

Bayesian networks simulate different conditions of uncertainty when it is not possible to know the truth of a hypothesis in a range of variation. Regarding, Lozano [31] states that "all Bayesian models have in common the assignment of probability as a measure of belief of a hypothesis, so that inference is a process of readjustment of belief measures when new ones are known axioms". For Bayesian inference it is common to use observations and evidence to find the truth of an assumption, so it is necessary to observe the evidence and estimate a value relating to the degree of belief raised in the hypothesis. In this way, when a greater number of data is available, results obtained are better.

After the construction of the models, it must be proceed to study the inference capacity. The inference system generated by a Bayesian network modifies the probability tables by finding new evidences regarding the state of some nodes. These new probabilities are propagated to the other nodes. This is known as probabilistic inference and allows calculate the probability of some of the variables from the evidence in other ones [25].

It is possible to use Bayesian networks for the construction of classifiers [32]. A cross-validation study is used, estimating the model with a random sample set, conformed by 70% or 80% of the data. After that, model is tested with the remaining 30% or 20%. The degree of adjustment of the model to the data set shows evidence of its degree of validity [33]. In addition, it is possible to construct a confusion matrix or classification table. It makes possible to compare the observations with the predictions cast by the model.

Stage 3: Analysis and interpretation of the definitive network

At this stage way in which the different variables are related is analyzed. In addition, those variables that are taken down from the network are identified. Therefore does not have any relation with any of the variables under consideration.

III. RESULTS

According to methodology used, Bayesian network is constructed from the database formed by the 35 indicators, classified in 5 different categories. Values of these indicators corresponds to the countries that make up the European Union, plus Norway and Iceland (Figure 2).

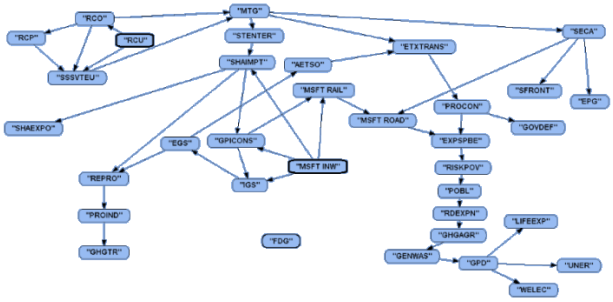


Figure 2. Relation between variables. Source: Own production

RCU and MSFT INW are decisive variables in the network. Each of them appear in the network as a variable "node" from which arcs come out. So, in each one divergent connections are generated. The first one, RCU, is a parent node that projects all its arcs to 2 children. This means that RCO (capacity of LNG regasification terminals in operation) and SSSVTEU (TEUs in short sea shipping) are a consequence of RCU (capacity of LNG regasification terminals under construction) as it is shown in Figure 3.

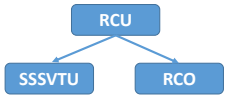


Figure 3. Divergent RCU Variable Connection. Source: Own production

The same happens with MSFT INW variable in which the arcs diverge towards their 4 children. So, IGS (imports from member states to third countries), GPICONS (prices of natural gas for industrial use), MSFT RAIL (modal distribution of freight by rail) and SHAIMPT (participation in imports by EU member) are a consequence of MSFT INV (Figure 4).

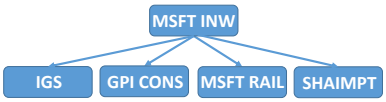


Figure 4. Divergent MSFT INW Variable Connection. Source: Own production

Analyzing relationship between SHAIMPT (participation in imports by EU member) it is observed that there are not only divergent relationship but also convergent ones. This variable has two parents, STENTER (students enrolled in tertiary education) and MSFT INW (modal distribution of cargo transport by inland waters), which converge in SHAIMPT. In other words, there are two alternative causes of it. In this type of connections, parent variables are independent of each other. However, if one of them has the evidence about the child variable, parent variables become dependent (Figure 5).

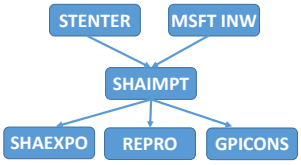


Figure 5. Variable SHAIMPT (Participation in imports by EU member). Source: Own production

Regarding divergent relations, SHAIMPT is considered as a decision variable, because it appears in the network as a variable "node". 3 arcs diverge from it towards their children. So, SHAEXPO (participation in exports by EU member), REPRO (resource productivity) and GPICONS (industrial natural gas prices) are an indirect effect of STENDER and MSFT INW.

In Figure 6, another decision variable can be seen. SSSVTEU (TEU's in SSS) has three parents which converge on it: RCP (capacity of LNG regasification terminals in plans), RCO (capacity of LNG Regasification Terminals in Operation) and RCU (capacity of LNG regasification terminals under

construction). All of these parents belong to institutional and political category. Additionally, SSSVTEU diverges towards MTG (maritime transport). So, MTG is a consequence of SSSVTEU. At the same time, that it means that RCU is related to the rest of the network variables, as it is shown in Figure 3.

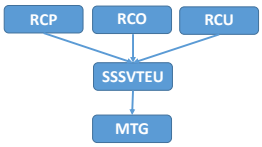


Figure 6. Variable SSSVTEU (TEU's in SSS). Source: Own production

SECA (sulfur emission control areas) is considered as another decision variable. MTG (maritime transport) is its single parent. Relating to SSSVTEU (TEU's in short sea shipping), SECA depends on MTG and it depends on SSSVTEU at the same time. Furthermore, this variable differs in 3 variables: MSFT ROAD (modal distribution of road freight), SFRONT (geographical location of the coasts) and EPG (port management models in Europe). All of them depends indirectly on RCU, as it is shown in Figure 2.

On the other hand, network shows some relations between variables called serial connections. In Figure 9 it can be seen that EXPSPBE (social protection expenditure) depends on MSFT ROAD (modal distribution of road freight transport) and PROCON (production in construction). Latter, it depends on ETXTRANS (Environmental taxes on Transportation) and so on. In term of cause-effect, it would be said that ETEXTRANS is cause of PROCON and, at the same time, PROCON is cause of EXPSPBE. In this case, given the dependence relationship between these variables, if some information about ETXTRANS is known, certainty could be modify related to the state of EXPSPBE. Consequently, if something about the state of EXPSPBE is known, the belief about the state of ETXTRANS can be altered (Figure 7).

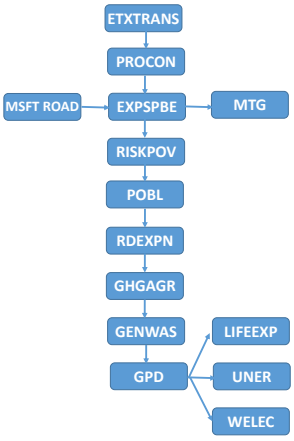


Figure 7. Variable EXPSPBE (Social Protection Expenses), Serial Connection. Source: Own production

Furthermore, RISKPOV (population at risk of poverty or social exclusion) depends on EXPSPBE (expenditure on social protection) and POBL (population), which depends, at the same time, on RISKPOV and so on. It finalizes at the end of the series. On the other hand, as it is shown in Figure 8, another shorter serial connection is observed. REPRO (productivity of resources) has 2 parents on whom it depends on: SHAIMPT (participation in imports by EU member) and EGS (exports from member states to third countries). It has a son, PROIND (production in the industry), which depends on REPRO at the same time. Finally, it is observed that GHGTR (emissions of greenhouse gases from the transport) depends on PROIND.

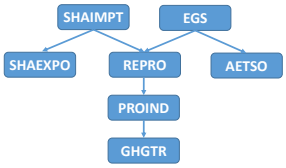


Figure 8. REPRO (Resource Productivity) Variable. Serial Condition. Source: Own production

Moreover, SECA depends on MTG (maritime transport), which depends on variables related to the capacity of the European gas infrastructure that is in operation and under construction, as well as, TEU'S in the SSS SSSVTEU.

In addition, a hierarchy was established between variables. Importance degree of each one was defined taking into account the number of arcs connecting each indicator with the most important nodes or indicators (RCU and MSFT INW). So, it was possible to find that transport and international trade category and economy and finance category are the most important ones, because their indicators are the closest to the root nodes of the network and they are the least dependent. Consequently, both categories are significant in the decision to implement LNG as fuel in the maritime stretch of the motorways of the sea. However, environment and energy category and population and social status category are the most dependent ones. They are far from the root nodes of the network.

Finally, it is possible to infer that all of the variables depend on RCU (capacity of LNG regasification terminals under construction) and 75% of them depend on MSFT INW (modal distribution of water cargo transportation continental). Both variables are the most important ones because they are independent and they do not depend on any other variable.

IV. CONCLUSIONS

Intensification of maritime use for freight transport has shown the need to more environmentally friendly fuels. LNG tries to be an alternative to this. Consequently, this study includes a methodology to determine the relationship between the variables that define the use of Liquefied Natural Gas in Short Sea Shipping in Europe. This methodology has been successfully developed using data mining techniques.

A Bayesian network has been constructed with the 35 indicators classified in 5 categories. It makes possible to establish that RCU (capacity of LNG regasification terminals under construction) and MSFT INW (modal distribution of water cargo transportation continental) variables are the two root nodes of the network. It means that they are the most important variables in the decision to implement LNG as fuel in the maritime stretch of the motorways of the sea.

Moreover, in the analysis by category, transport and international trade variables and economy and finance ones are

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

the most relevant in the decision-making process when implementing LNG in the maritime stretch of the motorways of the sea. However, environment and energy variables and population and social status ones are the most dependent on the entire network, so, they are the most affected ones in the network.

REFERENCES

[1] Cánovas, B. (2015). Short Sea Shipping, una estrategia Europea. Documento de Opinión, No. 70. Instituto Español de Estudios Estratégicos.

[2] Comisión Europea (2001). Libro Blanco, la política europea de transportes de cara al 2010: la hora de la verdad (pp. 11- 53). COM (2001) 370 final. Bruselas: Commission of the European Communities.

[3] Comisión Europea (2006). Libro Verde, estrategia europea para una energía sostenible, competitiva y segura (pp. 3- 19). COM (2006) 105 final. Bruselas: Commission of the European Communities.

[4] Bengtsson, S., Andersson, K., & Fridell, E. (2011). A comparative life cycle assessment of marine fuels: liquefied natural gas and three other fossil fuels. Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment, 225(2), 97-110.

[5] RECORDIT (2003). Actions to Promote Intermodal Transport: Final Report No. WP9. Available in: http://www.recordit.org/

[6] REALISE (2002). Regional Action for Logistical Integration of Shipping across Europe. The Alliance of Maritime Regional Interests in Europe (AMRIE), GTC2-2000-33032. Available in: http://www.realise-sss.org

[7] REALISE (2005). The Alliance of Maritime Regional Interests in Europe, (AMRIE), Final Report No. GTC2-2000-33032. Available in: http://www.realise-sss.org

[8] Lloyd, M. (2003). Regional action for logistical integration of shipping across Europe (realise). Work package 4. Task 4.1 multi-modal pricing and costing analyses inception report. Available in: http://www.realise-sss.org/

[9] Vassallo, W. (2004). Regional action for logistical integration of shipping across Europe (Realise). Available in: http://www.realise-sss.org/

[10] Psaraftis, H.N. and Kontovas, C.A. (2009). CO 2 Emission Statistics for the World Commercial Fleet. WMU Journal of Maritime Affairs, Vol. 8, No. 1, pp. 1-25.

[11] Bloor, M., Baker, S., Sampson, H. and Dahlgren, K. (2011). Effectiveness of international regulation of pollution controls: the case of the governance of ship emissions. Seafarers International Research Centre. Cardiff University.

[12] Psaraftis, H.N. and Kontovas, C.A. (2010). Balancing the Economic and Environmental Performance of Maritime Transportation. Transportation Research Part D: Transport and Environment, Vol. 15, No. 8, pp. 458-462.

[13] Corbett, J.J. , Fischbeck, P.S. and Pandis, S.N. (1999). Global Nitrogen and Sulfur Inventories for Oceangoing Ships. Journal of Geophysical Research, Vol. 104, No. D3, pp. 3457-3470.

[14] Endresen, Ø., Dalsøren, S., Eide, M., Isaksen, I. S., & Sørgård, E. (2008). The environmental impacts of increased international maritime shipping, past trends and future perspectives. In Proceedings of the Global Forum on Transport and Environment in a Globalising World, Guadalajara, Mexico, November (pp. 10-12).

[15] Corbett, J.J. and Korhler, H.W. (2003). Updated Emissions from Ocean Shipping. Journal of Geophysical Research, Vol. 108, No. D20, pp. 4650.

[16] Corbett, J.J., Wang, H. and Winebrake, J.J. (2009). The Effectiveness and Costs of Speed Reductions on Emissions from International Shipping. Transportation Research Part D: Transport and Environment, Vol. 14, No. 8, pp. 593-598.

[17] Devanney, J. (2010). CO2 Emissions from Shipping: The Case for Taking our Time. Available in: http://www.c4tx.org/ctx/pub/take_time.Pdf

[18] Devanney, J. and Beach, S. (2010). Detailed Studies of the Impact of Eedi on Vlcc Design and co2 Emissions. Center for Tankship Excellence. Available in: http://www.c4tx.org/ctx/pub/eedi_vlcc.pdf

[19] Almazán, J.L., Palomino, M.., González, N. Soler, F. y Iribarren, E. (2013). Estimation of Spanish bunkering at EU level of secas. Research Conference In Technical Disciplines, Section Transport and Logistic, pp. 136-141.

[20] Verbeek, R., Kadijk, G., Mensch, V.P., Wulffers, C., Beemt, B.V. and Fraga, F. (2011). Environmental and Economic aspects of using LNG as a fuel for shipping in the Netherlands. TNO report.

[21] Gullberg, M. and Gahnström, J. (2011). North European LNG Infrastructure Project: A feasibility study for an LNG filling station infrastructure and test of recommendations. Draft Feasibility Report. Trans-European Transport Network.

[22] Gritsenko, D., & Yliskylä-Peuralahti, J. (2013). Governing shipping externalities: Baltic ports in the process of SOx emission reduction. Maritime Studies, Volumen 12, No.1, p. 10.

[23] Acedo, A., Almazán, J. (2015). Pasado, presente y futuro de las Autopistas del Mar en Europa. Revista de Obras Públicas, No.3565, pp. 31- 38.

[24] Bengtsson, S.K., Fridell, E. and Andersson, K. (2014). Fuels for short sea shipping: A comparative assessment with focus on environmental impact. Journal of Engineering for the Maritime Environment, Vol. 228, No., pp. 44–54

[25] Rodríguez, D. & Dolado, J. (2007). Redes Bayesianas en la Ingeniería del Software.

[26] Moral, S. (2014). Una Introducción a las Redes Bayesianas.

[27] Puga, J.L. García, J.L., De la Fuente Sánchez, L. & de la Fuente Solana, E.I. (2007). Las redes bayesianas como herramientas de modelado en psicología. Anales de psicología, Volumen 23, No. 2, pp. 307-316.

[28] Heckerman, D. (2008). A tutorial on learning with Bayesian networks. Innovations in Bayesian networks, (pp. 33-82). Springer Berlin Heidelberg.

[29] Laskey, K. B. (2014, May). Sensitivity analysis for probability assessments in Bayesian networks. In Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence (p. 136).

[30] Niedermayer, D. (2008). An introduction to Bayesian networks and their contemporary applications. In Innovations in Bayesian Networks (pp. 117-130). Springer Berlin Heidelberg.

[31] Lozano, M.R. (2011). El papel de las redes bayesianas en la toma de decisiones, La simulación al servicio de la academia No. 2.

[32] Fernández, E. (2004). Análisis de Clasificadores Bayesianos, Trabajo Final de Especialidad en Ingeniería de Sistemas Expertos. Escuela de Postgrado. Instituto Tecnológico de Buenos Aires.

[33] Puga, J.L. (2012). Cómo construir y validar Redes bayesianas con netica. REMA Revista electrónica de metodología aplicada, Volumen 17, No.1, pp. 1-17.

eISSN: 2453-6571, cdISSN: 1339-5076
ISBN: 978-80-554-1375-4

Published by: EDIS - Publishing Institution of the University of Zilina
Univerzitna 1
01026 Zilina
Slovak Republic

Editors: Ing. Michal Mokrys; Ing. Stefan Badura, Assoc. Prof. Dragan Peraković

ISBN: 978-80-554-1375-4

cdISSN: 1339-5076

eISSN: 2453-6571

DOI: 10.18638/rcitd.2017.5.1

Pages: 67

Printed in: 65 copies

Publication year: 2017

- All published papers undergone single blind peer review.
- All published papers are in English language only. Each paper had assigned 2 reviewers and each paper went through two-tier approval process.
- Papers are published as delivered by authors without content modification. All accepted papers have been formally checked by the conference Technical Committee and regularly reviewed in single blind peer reviewing process by the conference Scientific Committee (Reviewers Committee).
- The publication contents are the sole responsibility of the publication author and it does not reflect the Publishing Society, Zilina, Slovakia.

Open Access Online archive is available at: **http://www.rcitd.com/archive**
(proceedings will be available online one month after the publication release).

In case of any questions, notes or complaints, please contact us at: **info(at)rcitd.com**.

Paper Citation Example: In (Eds.) S. Brown, S. Larsen, K. Marrongelle, and M. Oehrtman, Proceedings of The 5th International Virtual Research Conference in Technical Disciplines (RCITD-2017), Vol. 5, pg #-#. Zilina, Slovakia.

RCITD 2017

Proceedings in
Research Conference In Technical Disciplines 2017
The 5th International Virtual Research Conference In Technical Disciplines

13. - 17. November 2017
Slovak Republic

ISBN **978-80-554-1375-4**
cdISSN **1339-5076**
eISSN **2453-6571**