RĒZEKNES TEHNOLOĢIJU AKADĒMIJA
Inženieru fakultāte

REZEKNE ACADEMY OF TECHNOLOGIES
Faculty of Engineering

VIDE. TEHNOLOĢIJA. RESURSI

XII starptautiskās zinātniski praktiskās konferences materiāli
2019.gada 20.-22.jūnijs

2.SĒJUMS

ENVIRONMENT. TECHNOLOGY. RESOURCES

Proceedings of the 12th International Scientific and Practical Conference
June 20 – 22, 2019

Volume II

Rēzekne
2019

VIDE. TEHNOLOĢIJA. RESURSI. 12. starptautiskās zinātniski praktiskās konferences materiāli. 2019.gada 20.-22. jūnijs. *1.sējums.* Rēzekne: Rēzeknes Tehnoloģiju akadēmija, 2019. 263 lpp.

ENVIRONMENT. TECHNOLOGY. RESOURCES. Proceedings of the 12th International Scientific and Practical Conference on June 20-22, 2019. *Volume I,* Rezekne: Rezekne Academy of Technologies, 2019. pp. 263.

Rekomendējusi publicēšanai Rēzeknes Tehnoloģiju akadēmijas Zinātnes padome
2019. gada 23. maijā.
Recommended for publication by the Scientific Council of Rezekne Academy of Technologies on May, 23 rd, 2019.

ERDF co-funded project "Funding of international projects in research and innovation at Rezekne Academy of Technologies" No. 1.1.1.5/18/I/012

12.starptautiskās zinātniski praktiskās konferences ''Vide. Tehnoloģija. Resursi'' materiālu trijos sējumos ir pārstāvēti jaunākie pēdējo divu gadu pētījumi vides inženierzinātnē, vides un dabas aizsardzībā, ilgtspējīgā lauksaimniecībā, enerģētikā, materiālzinātnē, mehānikā, metālapstrādē, lāzeru tehnoloģijās, matemātiskajā modelēšanā, elektrotehnikā, vides ekonomikā un vadībā, informācijas tehnoloģijās un sociotehnisko sistēmu modelēšanā, vides izglītībā un ilgtspējīgas attīstības procesos. Krājumā pārstāvēto pētījumu joma ir daudzpusīga un starpdisciplināra, balstīta uz starptautisko zinātnieku kolektīvu sasniegumu rezultātiem. Konferences materiālos iekļauti 173 zinātniskie raksti, kuru autori pārstāv 14 valstis.

Proceedings of the 12th International Scientific and Practical Conference ''Environment. Technology. Resources'' include studies of the last two years in fields of engineering, environmental and nature protection, sustainable agriculture, energy, material science, mechanics, metalworking, laser technologies, mathematical modelling, electrical engineering, environmental economics and management, information technologies and sociotechnical systems modelling, environmental education and sustainable development. The research area presented in the proceedings is comprehensive and cross disciplinary-based, on advances of international researchers. The proceedings comprise 173 scientific papers from 14 countries.

Print ISSN 1691-5402
Online ISSN 2256-070X

This journal is licenced under
Creative Commons Attribution 4.0 International (CC BY 4.0) License
The author of the paper takes responsibility for the content of the paper.

Published by Rezekne Academy of Technologies, 2019

Scientific Committee Chairman

Dr.sc.ing. Edmunds Teirumnieks, Rezekne Academy of Technologies, Latvia

Scientific Committee Co-Chairmen

Dr. Walter Leal, Hamburg University of Applied Sciences, Germany
Dr.sc.ing. Andris Martinovs, Rezekne Academy of Technologies, Latvia
Dr.sc.ing. Artis Teilāns, Rezekne Academy of Technologies, Latvia

Scientific Committee

PhD Andres Annuk, Estonian University of Life Sciences, Estonia
Dr. Dragen Antic, University of Nis, Serbia
Dr.oec. Dzintra Atstāja, BA School of Business and Finance, Latvia
DSc. Hristo Ivanov Beloev, Angel Kanchev University of Ruse, Bulgaria
PhD Plamen Bogdanov, Vasil Levski National Military University, Bulgaria
Dr. Xiao Chen, Hefei University, China
Dr.sc.ing. Jurijs Cizovs, Riga Technical University, Latvia
PhD Lubomir Dimitrov, Tehnical University of Sofia, Bulgaria
Dr.-Ing. Horst Exner, Mittweida University of Applied Sciences, Germany
PhD Vladimir Golenkov, Belarusian State University of Informatics and Radioelectronics, Belarus
Dr. Ants Kallaste, Tallinn University of Technology, Estonia
PhD Tsanko Karadzhov, Technical University of Gabrovo, Bulgaria
Dr. Vasily Karasev, Institute of Problems of Mechanical Engineering of Russian Academy of Sciences, Russian Federation
PhD Veiko Karu, Tallinn University of Technology, Estonia
Dipl.-Kfm. Lars Kempt, Chemnitz University of Technology, Germany
Dr. Linas Kliucininkas, Kaunas University of Technology, Lithuania
PhD Egons Lavendelis, Riga Technical University, Latvia
PhD Gilberto Marzano, President of Ecoistituto Udine, Italy
PhD Cristian-Emil Moldoveanu, Military Technical Academy ‘’Ferdinand I’’, Romania
Dr.sc.ing. Vladislav Morozov, Geophysical Centre of the Russian Academy of Sciences, Russian Federation
Dr.geogr. Oļģerts Nikodemus, University of Latvia, Latvia
Dr. Vlastimir Nikolic, University of Nis, Serbia
Dr. Nenad Pavlovič, University of Nis, Serbia
Dr.habil.sc.ing. Igor Plokhov, Pskov State University, Russian Federation
Dr. Svetlana Polukoshko, Ventspils University College, Latvia
PhD Gabriel Raducanu, Air Force Academy Henri Coanda, Romania
Vilnis Rantins, Chairman of the Management Board of the Association of Mechanical Engineering and Metalworking Industries, Latvia
Dr. Anton Rassolkin, Tallinn University of Technology, Estonia
Dr.sc.ing. Andrejs Romanovs, IEEE Latvia Section, Riga Technical University, Latvia
Dr.habil.phys. Vladimir Solovyev, Pskov State University, Russian Federation
PhD Chrysostomos Stylios, Technological Educational Institute of Epirus, Greece
PhD Arita Takahashi, Riga Technical University, Latvia
Dr.sc.ing. Victor Tatarinov, Geophysical Centre of the Russian Academy of Sciences, Russian Federation
Dr.-Ing. Josef Timmerberg, Jade University of Applied Sciences, Germany
Dr. Toomas Vaimann, Tallinn University of Technology, Estonia
Dr.habil.sc.ing Janis Viba, Riga Technical University, Latvia

PhD Constantin-Iulian Vizitiu, Militar Technical Academy ‘’Ferdinand I’’, Romania

Organising Committee Chairman

Dr.habil.geol. Gotfrīds Noviks, Rezekne Academy of Technologies, Latvia

Organizing Commitee

Dr.sc.ing. Egils Ginters, Riga Tehnical University, Latvia
Dr.sc.ing. Pēteris Grabusts, Rezekne Academy of Technologies, Latvia
DSc. Dimitar Grekov, Agricultural University, Bulgaria
PhD Aleksandrs Gorbunovs, Riga Technical University, Latvia
PhD Krzysztof Krakowski, War studies University, Poland
Dr.sc.ing. Ēriks Kronbergs, University of Latvia, Latvia
Dr.sc.ing. Sergejs Kodors, Rezekne Academy of Technologies, Latvia
PhD Egons Lavendelis, Riga Technical University, Latvia
Dr.sc.ing. Lyubomir Lazov, Rezekne Academy of Technologies, Latvia
PhD Līga Mazure, Rezekne Academy of Technologies, Latvia
PhD Magdalena Mitkova, Prof. Assen Zlatarov University Burgas, Bulgaria
Dr.phys. Pāvels Narica, Rezekne Academy of Technologies, Latvia
PhD Valyo Nikolov, Technical University of Sofia, Plovdiv Branch, Bulgaria
Dr.sc.ing. Imants Nulle, Latvia University of Life Sciences and Technologies, Latvia
Dr.geol. Janis Prols, SIA “Geo Consultants”, Latvia
PhD Leonardo Quattrocchi, Confindustria, LUISS Business School, Italy
Dr.geol. Valdis Segliņš, University of Latvia, Latvia
Dr.biol. Artūrs Škute, Daugavpils University, Latvia
Dr.sc.ing. Victor Tatarinov, Geophysical Centre of the Russian Academy of Sciences, Russian Federation

Reviewers

PhD Nikolaj Angelov, Technical University of Gabrovo, Bulgaria
PhD Andres Annuk, Estonian University of Life Sciences, Estonia
PhD Iluta Arbidane, Rezekne Academy of Technologies, Latvia
Dr.oec. Dzintra Atstāja, BA School of Business and Finance, Latvia
Dr.agr. Skaidrīte Būmane, Institute of Food safety, Animal Health and Environment “BIOR”, Latvia
Dr.sc.ing. Jurijs Cizovs, Riga Technical University, Latvia
Dr.sc.ing. Edgars Čubars, Rezekne Academy of Technologies, Latvia
PhD Peteris Daugulis, Daugavpils University, Latvia
PhD Sandra Ezmale, Rezekne Academy of Technologies, Latvia
Dr.sc.ing. Pēteris Grabusts, Rezekne Academy of Technologies, Latvia
PhD Aleksandrs Gorbunovs, Riga Technical University, Latvia
Dr.oec. Elita Jermalajeva, Preili Municipality, Latvia
Dr.sc.ing. Aivars Kaķītis, Latvia University of Life Sciences and Technologies, Latvia
Dr.paed. Ilmārs Kangro, Rezekne Academy of Technology, Latvia
Dr. Vasily Karasev, Institute of Problems of Mechanical Engineering of Russian Academy of Sciences, Russian Federation
Dr.paed. Janis Kapenieks, Riga Technical University, Latvia
Dr.geol. Andris Karpovičs, Rezekne Academy of Technologies, Latvia
PhD Tsanko Karadzhov, Technical University of Gabrovo, Bulgaria
PhD Aivars Kaupuzs, Rezekne Academy of Technologies, Latvia
Dr.sc.ing. Sergejs Kodors, Rezekne Academy of Technologies, Latvia

PhD Egons Lavendelis, Riga Technical University, Latvia
Dr.sc.ing. Lyubomir Lazov, Rezekne Academy of Technologies, Latvia
PhD Angel Lengerov, Technical University of Plovdiv, Serbia
Dr.oec. Lienīte Litavniece, Rezekne Tecnologies Academy, Latvia
Dr.sc.ing. Jānis Locs, Riga Tehnical University, Latvia
Dr.habil.sc.ing. Zigurds Markovics, Riga Technical University, Latvia
Dr.sc.ing. Andris Martinovs, Rezekne Academy of Technologies, Latvia
PhD Gilberto Marzano, President of Ecoistituto Udine, Italy
Dr.habil.sc.ing. Viktors Mironovs, Riga Tehnical University, Latvia
Dr.phys. Pāvels Narica, Rezekne Academy of Technologies, Latvia
Dr.habil.geol. Gotfrīds Noviks, Rezekne Academy of Technologies, Latvia
Dr.sc.ing. Imants Nulle, Latvia University of Life Sciences and Technologies, Latvia
Dr.chem. Sergejs Osipovs, Daugavpils University, Latvia
Dr.sc.ing. Romāns Petričenko, Riga Technical University, Latvia
PhD Desislava Ivanova Petrova, Technical University of Gabrovo, Bulgaria
Dr.habil.sc.ing. Igor Plokhov, Pskov State University, Russian Federation
Dr.Sc.Ing Svetlana Polukoshko, Ventspils University College, Engineering Research Institute "VSRC", Latvia
Mg.sc.ing. Aivars Ratkevičs, Latvia University of Life Sciences and Technologies, Latvia
Dr.habil.math. Andrejs Reinfelds, University of Latvia
Dr.sc.ing. Andrejs Romanovs, IEEE Latvia Section, Riga Technical University, Latvia
Dr.geol. Valdis Segliņš, University of Latvia, Latvia
PhD Irēna Silineviča, Rezekne Academy of Technologies, Latvia
Dr.sc.ing. Andris Skromulis, Rezekne Academy of Technologies, Latvia
Dr.habil.phys. Vladimir Solovyev, Pskov State University, Russian Federation
Dr.geol. Juris Soms, Daugavpils University, Latvia
Dr.sc.ing. Sandis Spolītis, Riga Technical University, Latvia
Dr.agr. Veneranda Stramkale, Latvia University of Life Sciences and Technologies, Institute of Agricultural Resources and Economics, Latvia
PhD Aina Strode, Rezekne Academy of Technologies, Latvia
PhD Gunars Strods, Rezekne Academy of Technologies, Latvia
Dr.biol. Artūrs Škute, Daugavpils University, Latvia
PhD Arita Takahashi, Riga Technical University, Latvia
Dr.sc.ing. Artis Teilāns, Rezekne Academy of Technologies, Latvia
Dr.sc.ing. Edmunds Teirumnieks, Rezekne Academy of Technologies, Latvia
Dr.-Ing. Josef Timmerberg, Jade University of Applied Sciences, Germany
Dr.sc.ing. Mārtiņš Vilnītis, Riga Technical University, Latvia
Dr.sc. ing. Imants Zarembo, Rezekne Academy of Technologies, Latvia
PhD Erika Zubule, Rezekne Academy of Technologies, Latvia
PhD Anda Zvaigzne, Rezekne Academy of Technologies, Latvia

Technical Program Committee Chairman

Mg.sc.chem. Ērika Teirumnieka, Rezekne Academy of Technologies, Latvia

Technical Program Committee

Mg.sc.env. Sintija Augule, Rezekne Academy of Technologies, Latvia
Mg.sc.chem. Inese Bernāne, Rezekne Academy of Technologies, Latvia
Mg.sc.comp. Gundega Bēriņa, Rezekne Academy of Technologies, Latvia

II

Information Technologies and Engineering Education

II SĒJUMS
SATURS
CONTENTS

Information Technologies

Engineering Education

Information Technologies

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 17-20

Input Data Collection for the Fit-Gap Analysis Method: a Literature Review

Ieva Ancveire
Information Technology Institute
Riga Technical University
Riga, Latvia
ieva.ancveire@edu.rtu.lv

***Abstract*—Development of Enterprise Resource Planning (ERP) systems has become an independent industry for the improvement of information systems. It can be stated that ERP systems are designed to support the operation of a company. The fundamental objective of the system is to create a business value that aims at reducing the time and costs of the business while increasing the profit of the company. Although the adaptation and deployment of the ERP system, in general, are complex and protracted processes that require a lot of resources, the obtained results sometimes differ from the expected results of users. Mostly for that reason, there are incomplete internal enterprise business processes and software requirements analysis and development. Therefore, there is a need to determine the compliance of the main enterprise requirements and business processes with the ERP system. To choose the most appropriate ERP system, it is necessary to identify all possible methods of input data for the fit-gap analysis method. Thus, the main aim of the present study is to identify possible input data for the fit-gap analysis method, which can be used for the selection of the most appropriate ERP system.**

***Keywords*—Fit Gap analysis method, ERP systems, Input data.**

I. INTRODUCTION

The development of Enterprise Resource Planning (ERP) systems nowadays [1] has become an independent industry of information system development. It can be stated that ERP systems are designed to support the operation of the company [2]. The fundamental objective of the system is to create a business value [3] that aims at reducing the time and costs of the business while increasing the profit of the company [4]. There are several reasons as to why the ERP systems are underused, for example: (1) insufficient analysis and development of the enterprise internal business processes and software requirements, (2) need for new servers, (3) obtaining additional or new servers, (4) updates and other required specifications [5]. However, adjusting and introducing the ERP system, in general, are complex and time-consuming processes [6] [7] requiring many resources [8], and the obtained results tend to be different than the results expected by the users [4] [9].

Therefore, it is important to analyse compliance of the ERP system with the enterprise main requirements and business processes. Fit-gap analysis method is one of the ways [10] how to successfully choose the right ERP system [11]. Overall, there are many studies addressing selection and introduction of the ERP system [12]; however, studies on the use of the fit-gap analysis method in the process of adjusting the ERP system are not that common. These include scientific publications analysing the main input and output data types of the said method for ERP system projects.

Therefore, it is possible that there are unresolved challenges related to the identification of the main sources of information in the decision-making process of the fit-gap analysis [13], and not only for selecting the most suitable ERP system solution.

Analysing the collected scientific publications, it is important to identify all possible types of input data in the fit-gap analysis method and to determine whether the obtained output data comply with the expected results. Therefore, the following research questions are defined:

RQ1: What input data are identified by publications on the fit-gap method?

RQ2: Can these data be used in the fit-gap method for the purposes of ERP system development and deployment projects?

RQ3: How can these data be categorized?

RQ4: What output data should be received for the input data?

RQ5: What kind of improvements are available for the pre-defined fit-gap methods?

This publication contributes by identifying and collecting possible input data of the fit-gap analysis method used in the selection of the most suitable ERP system.

The rest of the paper is organized as follows: Section 2 details the research design process, describing the process of selection and analysis of scientific publications. The description of scientific publications and answers to research questions **RQ1**, **RQ2** and **RQ3** are given in

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4160
© 2019 Ieva Ancveire.
Published by Rezekne Academy of Technologies.
This is an open access article under the Creative Commons Attribution 4.0 International License.

Section 3. Section 4 addresses study questions where the input data are contrasted with the output data of fit-gap method, as well as identifies the possible improvements of the fit-gap analysis method. Finally, Section 5 presents conclusions related to literature review and suggests directions for further research.

II. RESEARCH DESIGN

The literature analysis of the publication is aimed at collecting scientific publications about the input data of fit-gap method in the ERP system development projects, devoting special attention to the need of identifying the ways for improvement of the already identified fit-gap methods. The process of literature analysis is based on a systematic literature analysis seeking for answers to the research questions. Therefore, scientific literature available in the scientific databases has been analysed within the framework of the present study.

The literature analysis (see Fig. 1) has been performed in 4 stages to identify and collect scientific publications addressing the input data of fit-gap method in the ERP system development projects.

First, it is necessary to define key words for literature selection. For the purposes of the present literature analysis, five key words have been identified. The following key words have been used in the selection process of scientific literature: "fit gap method erp data", "gap analysis method inputs", "gap analysis method inputs erp", "Fit-Gap Analysis inputs", "Gap Analysis inputs erp", "gap analysis erp artefacts", "gap analysis artefacts".

Scientific databases have been used to select scientific publications covering the time period starting with 2000. Once abstracts have been analyzed, Stage 3 involved selecting twenty scientific publications for closer review of the scientific literature source, including the previously conducted research into the use of the fit-gap analysis method in ERP projects.

Thereafter, the selected scientific publications have been assessed in full. To carry out an in-depth investigation of the issue addressed by the publication, special attention has been devoted to references in the main texts of the selected scientific publications. Out of all initially selected scientific publications, only three address the use of the fit-gap analysis method in the ERP system selection and deployment projects, allowing the author to answer the research questions. Since the previously conducted [14] study on the fit-gap analysis methods was also identified among the selected publications, the input data of the method mentioned in the sources of the said publication would also be studied when gathering the available input data type. To answer the research questions, overall 13 scientific publications will be analyzed in more detail.

The input data types in the identified scientific publications can be categorized in two groups: (1) input data, the source of which is the customer, and (2) input data, the source of which is the ERP system and its functional description.

After analysing the main texts of the scientific publications, it has been concluded that in order to answer research question **RQ2**, it is not necessary to carry out additional literature analysis because while identifying scientific publications during the process of abstract analysis, only publications addressing fit-gap analysis input data and their use in the ERP system development processes have been selected. The following sections offer a more detailed analysis of the scientific publications.

III. The Method Input Data

For the purposes of the fit-gap analysis method and ERP systems, the available literature does not provide a definition for the concept of "data", but it follows from the literature that "data" can be defined as quantitative facts or information on events [6], which can be processed and used to compare the ERP system with the requirements of the users.

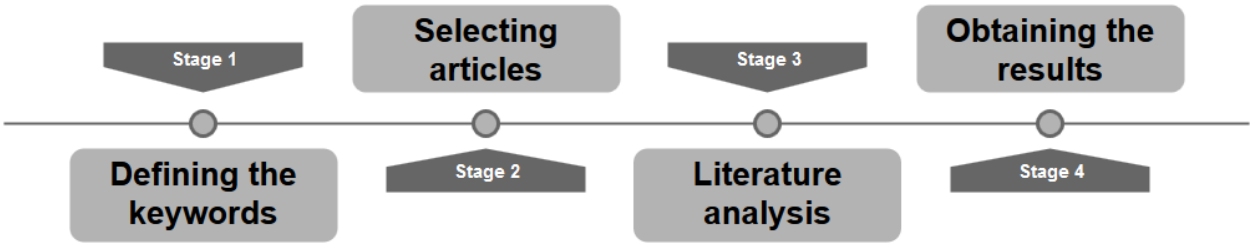


Fig. 1. Literature analysis process

This section collects and investigates not only scientific publications found in the process but also all pre-studied and analysed scientific literature [14], identifying input data types for the fit-gap analysis method. The section aims at answering research questions **RQ1**, **RQ2** and **RQ3**.

As previously mentioned, the input data of the analysis method can be divided into two categories: input data, the source of which is the customer and input data, the source of which is the ERP system and its functionality. Bearing in mind both categories, the identified scientific publications have also been categorized and analysed in detail (see Table 1). Overall, the author of the present study has found 13 scientific publications identifying business, functional and non-functional requirements and business processes as input data for the fit-gap method.

It has been identified in 13 scientific publications that the customer requirements are one of the input data types. In four out of 13 publications, the ERP systems and infrastructures have been specified as a type of input data.

However, three out of 13 publications argue that national legislation shall be considered when comparing customer requirements and ERP systems or selecting the ERP system.

TABLE I. CATEGORIZATION OF INPUT DATA

Input data type	Publication references
National legislation	[8], [15], [16]
Customer requirements	
Business requirements	[2], [4], [11], [15] – [21]
Enterprise business processes	[2], [8], [11], [15], [16], [18], [19], [21], [22]
Functional and non-functional requirements	[2], [4], [8], [16], [17], [19]
Not detailed	[23]
ERP system and infrastructure	
ERP systems and functionality, and business processes	[2], [15], [16], [19], [21], [22]
ERP system infrastructure requirements	[11], [18]
ERP system prototype	[20]
Not detailed	-

The analysis of scientific publications demonstrates that the most common input data types are business requirements, enterprise business processes and ERP system functionality and business processes. The customer's business requirements and processes can be regarded as the primary input data [8] divided into various enterprise events, enterprise operations, enterprise objects, metrics and data [17], [18], [21], activity diagrams, sketches, data vocabularies, and business process details [19]. The main ingredients of the business processes are client interviews and client enterprise business process observations [8], detailed descriptions of utilization cases (problem outline, assumptions, preconditions, prescriptions and statistics) [2].

After collecting the publications, it has been concluded that the best way how to improve the input data quality is to develop the business requirements in a working group composed not only of developers but also of customers [2]. It is also suggested to compile and describe the business processes in documents in a structured manner [2] by using pre-prepared templates [22]. This would allow incorporating the functional and non-functional requirements in the descriptions of business processes, as well as grouping and prioritizing the requirements according to a selection approach.

However, documentation is secondary data [8]. To prepare documentation, business processes are reviewed and documented to implement these business processes as part of a new solution. The requirement documents are composed of the business, functional and non-functional requirements [4]. The two main types of input data documentation arise therefrom: (1) specification of requirements of priority data software where all software requirements are categorized in the order of priority, (2) specification of software requirements including therein – for the purposes of the project development team – application requirements (describing customer needs with respect to software), process requirements (describing functions and tasks to comply with application requirements) and design requirements (describing requirements ensuring compliance with process) [11], [15], [16].

The category where input data are sourced from the ERP system and its functionality can be divided into three sub-categories. In six out of 13 scientific publications, authors argue that the ERP system functionality and ERP system business processes can be used as input data for the investigated comparison method. In two out of 13 scientific publications, authors argue that the ERP system infrastructure requirements can be used as input data, but only one publication addresses the ERP system prototype.

After an in-depth analysis of the scientific publications, it can be concluded that the national legislation can also be regarded as one of the input data categories. The national legislation does not address requirements defined by customers or requirements describing the ERP system, but instead deal with the legislation in the customer's country.

IV. FINDING THE METHOD OUTPUT DATA

Based on research question **RQ4**, the identified scientific publications have attempted to determine whether there is a relation between the input and output data. At first, all possible output data have been identified in each scientific publication (see Table 2).

The most common output data for the fit-gap method is the list of high-level ERP system functionality compliances/non-compliances against enterprise business processes, requirements and data. This is reflected in eight out of 13 scientific publications. Having gathered information from the publications, such output values contain descriptions and reviews of compliances and non-compliances between [22]: (1) main architectural elements, (2) solutions [19]; (3) enterprise, sectoral, integration and national legislation requirements [4], [15], (4) enterprise requirements [19] and (5) prevention of identified non-compliances, e.g., activities to be carried out [11] and assessment of work capacity thereof [4] and adjustment expenses [8], as well as suitability of the ERP system [23] and success factors of the ERP system [18].

TABLE II. POSSIBLE OUTPUT DATA OF THE METHOD

Output data types	Reference
List of high-level ERP system functionality compliances/non-compliances against enterprise business processes, requirements and data	[4], [8], [11], [15], [18], [19], [22], [23]
List of the most suitable ERP systems or the most suitable ERP system	[2]
Documentation on adjustment possibilities of the ERP functionality or *enterprise* business processes	[8], [17], [20], [21]
Suitability of an enterprise for introducing the ERP system	[16], [18]
Not mentioned	-

However, in four out of 13 publications the possible output data are documentation on the ERP system functionality or adjustment possibilities of the enterprise business processes. Statistically, the fit-gap methods offer less chances to obtain the list of the most suitable ERP systems or the most suitable ERP system or to determine suitability of the enterprise for the ERP system deployment. This is addressed only in three out of 13 publications.

V. CONCLUSIONS

Although the development of the Enterprise Resource Planning (ERP) systems has become an individual industry of information system development, one can still encounter cases when the ERP systems are insufficiently utilized for the purposes of enterprise business processes.

The present paper has attempted to identify and collect the possible fit-gap analysis input data used for the selection of the most suitable ERP system. The paper has also aimed at finding out whether the obtained output data comply with the expected results and whether it is possible to identify improvements of the existing methods based on the available information.

Within the framework of the study, 23 scientific publications have been investigated within the literature analysis finding answers to the three research questions: **RQ1**, **RQ2** and **RQ3**. Having gathered information about the questions under consideration, it can be stated that the most common type of input data is enterprise business requirements and processes, as well as ERP system functionality and business processes. However, for it to be possible to correctly use the said input data, it is necessary to develop a unified system for defining and describing input data and ensuring correct data input.

The author of the paper has strived to find a relation between input and output data (**RQ4**) and attempted to answer research question **RQ5**. Having analysed the identified publications, it has been concluded that the gathered information does not allow coming up with a clear answer to research questions **RQ4** and **RQ5**. Therefore, to answer research questions **RQ4** and **RQ5**, it is necessary to carry out an additional literature review, which is the next challenge in the study of the fit-gap analysis methods.

References

[1] N. Kountouridou, P. Antoniou, and I. Stamelos, "A Comprehensive Approach for Implementing an Open Source ERP in a Greek Industry," in Proceedings of the 20th Pan-Hellenic Conference on Informatics - PCI '16, 2016, pp. 1–5.

[2] M. Nöbauer, N. Seyff, N. Maiden, and K. Zachos, "S 3 C : Using Service Discovery to Support Requirements Elicitation in the ERP Domain," pp. 18–32, 2011.

[3] T. Chandrakumar and S. Parthasarathy, "An approach to estimate the size of ERP package using package points," Comput. Stand. Interfaces, vol. 47, pp. 100–107, 2016.

[4] D. Pajk and A. Kovacic, "Fit Gap Analysis – The Role of Business Process Reference Models," Econ. Bus. Rev., vol. 15, no. 4, pp. 319–338, 2013.

[5] B. K. Muslmani, S. Kazakzeh, E. Ayoubi, and S. Aljawarneh, "Reducing Integration Complexity of Cloud-based ERP Systems," in Proceedings of the First International Conference on Data Science, E-learning and Information Systems, 2018, p. 37:1--37:6.

[6] A. Skoogh, T. Perera, and B. Johansson, "Input data management in simulation - Industrial practices and future trends," Simul. Model. Pract. Theory, vol. 29, pp. 181–192, 2012.

[7] L. Shaul and D. Tauber, "Critical success factors in enterprise resource planning systems," ACM Comput. Surv., vol. 45, no. 4, pp. 1–39, 2013.

[8] E. Hustad, M. Haddara, and B. Kalvenes, "ERP and organizational misfits: An ERP customization journey," Procedia Comput. Sci., vol. 100, no. 1877, pp. 429–439, 2016.

[9] Panorama Consulting Solutions, "2015 Erp Report," Panor. Consult. Solut., pp. 1–19, 2015.

[10] J. Nugra, M. Wiyarta, and Y. Kurniawan, "The Evaluation of Inventory Management Module of Enterprise Resource Planning System (A Case Study Approach)," in Proceedings of the 2018 International Conference on Internet and e-Business - ICIEB '18, 2018, pp. 79–83.

[11] G. Juntao and Z. Li, "Detecting gaps between ERP software and organizational needs: A semantic similarity based approach," in Proceedings - 1st IEEE International Workshop on Semantic Computing and Systems, WSCS 2008, 2008, pp. 21–26.

[12] M. Haddara, "ERP systems selection in multinational enterprises: A practical guide," Int. J. Inf. Syst. Proj. Manag., vol. 6, no. 1, pp. 43–57, 2018.

[13] J. Grabis, "Optimization of Gaps resolution strategy in implementation of ERP systems," in Proceedings of 21st International Conference on Enterprise Information Systems (ICEIS 2019), 2019, p. To appear.

[14] I. Ancveire, "Fit Gap Analysis Methods for ERP Systems," 2018 IEEE 12th Int. Symp. Appl. Comput. Intell. Informatics, pp. 161–166, 2018.

[15] T. Shiang-yen, "A Framework for Classifying Misfits Between Enterprise Resource Planning (ERP) Systems," Asian Acad. Manag. J., vol. 16, no. 2, pp. 53–75, 2011.

[16] S. Parthasarathy and M. Daneva, "An approach to estimation of degree of customization for ERP projects using prioritized requirements," J. Syst. Softw., vol. 117, pp. 471–487, 2016.

[17] P. A. Millet, "Toward a model-driven, alignment-oriented ERP methodology," Comput. Ind., vol. 64, no. 4, pp. 402–411, 2013.

[18] H. Sun, W. Ni, and R. Lam, "A step-by-step performance assessment and improvement method for ERP implementation: Action case studies in Chinese companies," Comput. Ind., vol. 68, pp. 40–52, 2015.

[19] J. H. Wu, S. S. Shin, and M. S. H. Heng, "A methodology for ERP misfit analysis," Inf. Manag., vol. 44, no. 8, pp. 666–680, 2007.

[20] M. Nöbauer et al., "Managing variability of ERP ecosystems: Research issues and solution ideas from microsoft dynamics AX," ACM Int. Conf. Proceeding Ser. (VaMoS '12), pp. 21–26, 2012.

[21] R. Agarwal, A. R. Raha, and B. Ghosh, "Our experience and learning in ERP implementation," ACM SIGSOFT Softw. Eng. Notes, vol. 25, no. 2, pp. 31–34, 2000.

[22] T. R. Gulledge, "Erp Gap-Fit Analysis From a Business Process Orientation," Int. J. Serv. Stand., vol. 2, no. 4, pp. 339–348, 2006.

[23] I. Syamsuddin, "Novel gap analysis framework for cloud health information systems," J. Theor. Appl. Inf. Technol., vol. 87, no. 3, pp. 415–421, 2016.

Models and Algorithms for Constructing a Formalized Description of Production Technologies

Dmitry Andreev
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
dandreev60@mail.ru

Sergey Lyokhin
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
slyokhin@gmail.com

Lilia Motaylenko
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
lvs@mail.ru

Sergey Verteshev
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
verteshev@mail.ru

***Abstract*—The problem of technological knowledge formalization is being considered. Discussed features of systems in the scope of ontological engineering of processes, applied to solving the problems of the structural representation of technologies associated with the organization of their analysis. Proposed a model of the concept of technological action, which identifies technological actions as holistic conceptual formations. Constructed the model of formalized description of technologies, which provides unambiguous transition from verbal descriptions of technologies to their ontological representations. Described algorithms for constructing a formalized description of technologies, which implement the logical procedures for automated and automatic construction of ontological representations of technologies. Concluded in the development of a new method of formalized description of technologies, the fundamental principles of which constitute the distinctive features of the proposed models and constructed algorithms.**

***Keywords*—algorithmic procedure, formalization, relation, technological action.**

I. Introduction

In the modern world technological knowledge with full confidence can be considered as the most valuable information. This is because of the increased intensity of a struggle for possession of information resources, among which the knowledge of production technologies (hereinafter – technologies) occupies one of the leading positions [1] – [4]. The presence of a huge number of technologies that differ in the most diverse groups of their components makes it necessary to analyze and compare them. Solution of scientific and practical problems within the framework of this problem receives a new impulse of development in the presence of effective tools for constructing a formalized description of technologies.

At present, the fixation of knowledge about technologies occurs through various methods of their description, which are quite heterogeneous from the standpoint of formalization aspects [5]. At the same time, the possibility of computer processing of this knowledge in the interests of analysis of technologies is significantly limited by the means of implementing these methods. In this regard, a trend towards a more efficient use of technological knowledge is needed to overcome this situation, which can be ensured by their successful operation within a progressive methodological basis, the role of which in recent decades has been actively claimed by the concept of a system-ontological approach [6] – [9].

The study of the features of systems in the field of ontological engineering of processes shows the impossibility of their application to solving the problems of formalized representation of technological knowledge associated with the organization of analysis of technologies. Problem arises in weak elaboration of formal mechanisms of the applied level of description of technologies, in the part of the formation of their decomposition structures (DST) that is expressed in the absence of representation of the activity element as a holistic conceptual formation and in the replacement of the axiomatics of ontological models with graphical notations of the corresponding tools [10], [11].

Thus, it seems to be very important to develop a method of formalized description of technologies by means of a system-ontological approach, the basic principles of which allows operating technological knowledge by computer tools in automated and automatic modes and solving a whole range of tasks of analysis of technologies.

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4063
© 2019 Dmitry Andreev, Sergey Lyokhin, Sergey Verteshev, Lilia Motaylenko.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

II. THE MODEL OF THE CONCEPT OF TECHNOLOGICAL ACTION

From the semantic point of view, the ontology should store information about the sets of concepts, whose purpose is to conduct a detailed formalization of certain field of knowledge [12]. In accordance with the generally accepted practice of ontological modeling, conceptual elements of the field of knowledge are required to act as such concepts. The features of this research indicate the need to consider such concept as the concept of technological action, which is a single holistic formation of a formalized description of technology [13].

The model of the concept of technological action is the basic representation of each of the technological actions located in the nodes of the DST, is proposed to be represented by a tuple [14]:

$$TD_u =< TP_u, Y_u, X_u \,|\, W_u, H_u, Z_u > \ , \qquad (1)$$

where TD_u – a concept of technological action, TP_u – a kernel of the concept, Y_u – a set of resulting components, X_u – a set of source components, W_u – a set of invariant components, H_u – a set of cost characteristics, Z_u – a set of own characteristics.

When designing the DST, the following initial degrees of content formation of the concepts are determined:

Definition 1. The concept is fully formed if

$$\forall\, TD_u\ (TP_u \neq \varnothing, Y_u \neq \varnothing, X_u \neq \varnothing, W_u \neq \varnothing, H_u \neq \varnothing, Z_u \neq \varnothing \vee Z_u = \varnothing) \qquad (2)$$

Definition 2. The concept is preliminary formed if

$$\forall\, TD_u{}'\ (TP_u{}' \neq \varnothing, Y_u{}' \neq \varnothing, X_u{}' = \varnothing, W_u{}' = \varnothing, H_u{}' = \varnothing, Z_u{}' = \varnothing) \qquad (3)$$

The designation u represents the quantity of positions in the index of each of the technological actions located in the nodes of the DST. The index is an ordered, well-defined sequence of natural numbers, the order of formation of which is that the set of values of each subsequent position is put in accordance with a specific element of the set of values of the previous position [15]. Thus, the unambiguous identification of the location of the concepts of technological actions in the nodes of the DST is formalized. For example, for a certain concept of λ-level of the DST, the index u is the sequence of natural numbers $(\alpha_i, \beta_j, ..., \eta_l, \lambda_m)$. If this concept is preliminary formed, then it can be detailed to the presentation in the form of a set of other concepts, the index of each of which will be supplemented by an additional position. Such formation is also called a "unified decomposition construction" (UDC), at the apex of which is the only holistic concept, and at the basis is a set of private concepts. "Fig. 1" shows an example of one of the possible variants of the graphical view of the UDC.

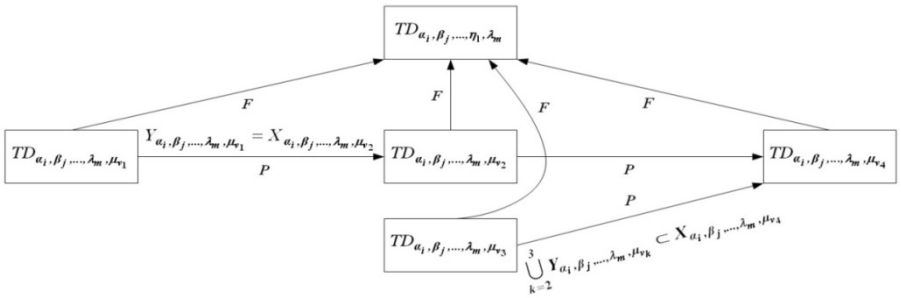


Fig. 1. The example of the UDC

III. THE MODEL OF THE FORMALIZED DESCRIPTION OF TECHNOLOGIES

In information technologies and computer science, ontology is most often considered as an explicit specification of conceptualization, which implies the formalization of a certain field of knowledge not only through the presence of a set of concepts, but also the description of relations between them [16].

In order to describe relations between the fully formed private concepts of the same level of decomposition, the relation P of "immediate precedence" is introduced [17].

Definition 3. Fully formed concepts are in the relation of immediate precedence if

$$\forall\, TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \forall\, TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\ (TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ P\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}): \ (4)$$
$$:(Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r} \subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_s})$$

When designing a DST, each of the levels of decomposition will have so-called "initial" and "finite", and some of them will also have "complementary" fully formed concepts.

Definition 4. A fully formed concept of a certain level of decomposition is initial if

$$\forall\, TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \exists\, TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\ \bar{\exists}\, TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_q}\ ((Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r} \subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}) \wedge \wedge (Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_q} \subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_r})) \qquad (5)$$

Definition 5. A fully formed concept of a certain level of decomposition is finite if

$$\forall\, TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \exists\, TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_q}\ \bar{\exists}\, TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\ ((Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_q} \subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}) \wedge \wedge (Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r} \subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_s})) \qquad (6)$$

Definition 6. A fully formed concept is complementary to another fully formed concept related to the same level of decomposition if

$$\begin{aligned}&\forall\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_t}\ \ \exists\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\\&\exists\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \ \exists\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_q}\\&\bar{\exists}\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_u}\ ((Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_t}\ \subset\\&\subset X_{\alpha_i,\beta_j,...,\lambda_m,\mu_s})\wedge(Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \subseteq\\&\subseteq(X_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\ \setminus X^*_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}))\wedge\\&\wedge(Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_q}\ \subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_r})\wedge\\&\wedge(Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_u}\subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_t}))\end{aligned}\tag{7}$$

In order to describe relations between each of the fully formed private concepts and the already fully formed holistic concept, the relation F of "part-whole" is introduced [18].

Definition 7. Fully formed concepts are in the relation of "part-whole" if (8).

The ordered totality of this kind of UDC forms DST.

Then the model of the formalized description of technologies is determined by the tuple [19]:

$$Ont_{Tech} =< TD, P, F >,\tag{8}$$

where $TD = \bigcup\limits_u TD_u$ – the set of concepts of technological actions, P – the intra-level relation of immediate precedence, F – the inter-level relation of "part-whole".

$$\begin{aligned}&\forall\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \forall\ TD_{\alpha_i,\beta_j,...,\eta_l,\lambda_m}\\&(TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ F\ TD_{\alpha_i,\beta_j,...,\eta_l,\lambda_m}):\\&:(Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\subset Y_{\alpha_i,\beta_j,...,\eta_l,\lambda_m})\wedge\\&\wedge(((W_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}=\bigcup_{k=1}^{n} w^k_{\alpha_i,\beta_j,...,\lambda_m,\mu_r})\subset\\&\subset(W_{\alpha_i,\beta_j,...,\eta_l,\lambda_m}=\bigcup_{k=1}^{n} w^k_{\alpha_i,\beta_j,...,\eta_l,\lambda_m}))\Big|\\&\Big|\exists w^{k_a}_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \exists w^{k_a}_{\alpha_i,\beta_j,...,\eta_l,\lambda_m},\\&(w^{k_a}_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\subset w^{k_a}_{\alpha_i,\beta_j,...,\eta_l,\lambda_m}))\wedge\\&\wedge((X_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\subset X_{\alpha_i,\beta_j,...,\eta_l,\lambda_m})\vee\\&\vee((X_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\setminus X^*_{\alpha_i,\beta_j,...,\lambda_m,\mu_r})\subset\\&\subset X_{\alpha_i,\beta_j,...,\eta_l,\lambda_m})\vee\\&\vee(\neg(X_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\subset X_{\alpha_i,\beta_j,...,\eta_l,\lambda_m})))\end{aligned}\tag{9}$$

The proposed model of the form (9) allows making a set-theoretic analysis of the considered descriptions of technologies by means of the following set of axiomatic properties [20]:

1. The set of resulting components of a fully formed concept is unique:

$$\begin{aligned}&\forall\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \ \forall\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\\&(Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \mathrm{I}\ Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}=\varnothing)\end{aligned}\tag{10}$$

2. The sets of source and resulting components of a fully formed concept do not have common elements:

$$\begin{aligned}&\forall\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\\&(X_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \mathrm{I}\ Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}=\varnothing)\end{aligned}\tag{11}$$

3. On the sign of the decomposition of the kernel of a preliminary formed holistic concept:

$$\begin{aligned}&\exists\ TD'_{\alpha_i,\beta_j,...,\eta_l,\lambda_m}\ \ \exists\ \bigcup_{\mu} TD'_{\alpha_i,\beta_j,...,\lambda_m,\mu}\\&((Y'_{\alpha_i,\beta_j,...,\eta_l,\lambda_m}=\bigcup_{\mu} Y'_{\alpha_i,\beta_j,...,\lambda_m,\mu}\ \Big|\\&\Big|\ \mu=\overline{1,n},\ n\geq 2)\Rightarrow\\&\Rightarrow(TP'_{\alpha_i,\beta_j,...,\eta_l,\lambda_m}=\bigcup_{\mu} TP'_{\alpha_i,\beta_j,...,\lambda_m,\mu}\Big|\\&\Big|\ \mu=\overline{1,n},\ n\geq 2))\end{aligned}\tag{12}$$

4. On the definition of a complement to a set of source components of a fully formed concept:

$$\begin{aligned}&\forall\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\ \ \forall\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\\&((Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\subset X_{\alpha_i,\beta_j,...,\lambda_m,\mu_s})\Rightarrow\\&\Rightarrow(\bar{\exists}\bigcup_{k=1}^{n} TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_{v_k}}\ \ \vee\\&\vee\ \exists\bigcup_{k=1}^{n} TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_{v_k}},\\&\bigcup_{k=1}^{n} Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_{v_k}}\ \subseteq\\&\subseteq(X_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\setminus X^*_{\alpha_i,\beta_j,...,\lambda_m,\mu_s})))\end{aligned}\tag{13}$$

5. The set of resulting components of a fully formed concept cannot be a subset of several sets of source components of other fully formed concepts:

$$\begin{aligned}&\forall TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\\&\exists\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_s}\ \bar{\exists}\ TD_{\alpha_i,\beta_j,...,\lambda_m,\mu_q}\\&((Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_s})\wedge\\&\wedge(Y_{\alpha_i,\beta_j,...,\lambda_m,\mu_r}\subseteq X_{\alpha_i,\beta_j,...,\lambda_m,\mu_q}))\end{aligned}\tag{14}$$

6. On the definition of a fully formed holistic concept on the basis an aggregated set of source components:

$$\exists\, TD'_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}\ \ \exists\underset{\mu}{\mathrm{Y}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}$$
$$((\underset{\mu}{\mathrm{Y}} X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu} = X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}) \mid$$
$$\mid(\forall TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}\ (\bar{\exists}\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q}$$
$$((Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q} \subseteq X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}) \Rightarrow$$
$$\Rightarrow (X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} := X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \cup$$
$$\cup X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r})) \vee \exists\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q}$$
$$(((Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q} \subset X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}) \Rightarrow$$
$$\Rightarrow (X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} := X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \cup$$
$$\cup (X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r} \setminus X^*_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}))) \vee$$
$$\vee ((Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q} = X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}) \Rightarrow$$
$$\Rightarrow (X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} := X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \cup$$
$$\cup \varnothing))))) \Rightarrow (TD'_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \rightarrow$$
$$\rightarrow TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m})^{\mathrm{X}}) \qquad (15)$$

7. On the definition of a fully formed holistic concept on the basis an aggregated set of invariant components:

$$\exists\, TD'_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}\ \ \exists\underset{\mu}{\mathrm{Y}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}$$
$$((\underset{\mu}{\mathrm{Y}} W_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu} = W_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}) \mid$$
$$\mid(\forall TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}\ \exists\, w^{k_a}_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$$
$$((w^{k_a}_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r} = w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}) \Rightarrow$$
$$\Rightarrow (w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} := w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \cup \varnothing)) \vee$$
$$\vee ((w^{k_a}_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r} \cap w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \neq \varnothing) \Rightarrow$$
$$\Rightarrow (w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} := w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \cup$$
$$\cup\ (w^{k_a}_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r} \setminus w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}))) \vee$$
$$\vee ((w^{k_a}_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r} \cap w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} = \varnothing) \Rightarrow$$
$$\Rightarrow (w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} := w^{k_a}_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \cup$$
$$\cup\ w^{k_a}_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}))) \Rightarrow (TD'_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \rightarrow$$
$$\rightarrow TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m})^{\mathrm{W}}) \qquad (16)$$

8. On the definition of a fully formed holistic concept on the basis the additivity of the set of cost characteristics:

$$\exists\, TD'_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}\ \ \exists\underset{\mu}{\mathrm{Y}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}$$
$$((\underset{\mu}{\mathrm{Y}} H_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu} = H_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \mid \forall\, z \in [1,k],$$
$$\left(\sum_{\mu=1}^{n} \{h^z_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}\} = \{h^z_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}\}\right) \Rightarrow \qquad (17)$$
$$\Rightarrow (TD'_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m} \rightarrow TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m})^H)$$

IV. ALGORITHMIC PROCEDURES FOR CONSTRUCTING ONTOLOGICAL REPRESENTATIONS OF TECHNOLOGIES

The process of constructing a formalized description of technologies is the process of implementing formulated and proved theoretical propositions, which leads to the possibility of the formation of the DST by means of a staged combined design of the UDC, starting with the root representations of the technologies [21].

The implementation of main stages of UDC construction goes in accordance with the execution of the following algorithmic procedures (dotted line in "Fig. 2"):

1. The algorithm for the automated construction of decomposition of a preliminary formed concept (stage of downward design);

2. The algorithm for the automatic determination of relations between fully formed private concepts (stage of intra-level design);

3. The algorithm for the automatic determination of the complete formation of a holistic concept (stage of upward design).

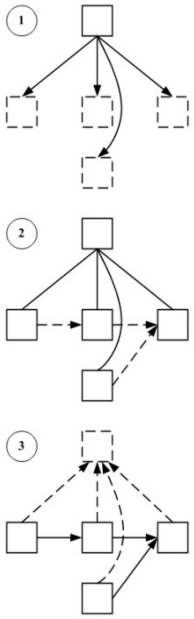


Fig. 2. Staged algorithmic procedures for constructing UDC

The stage of downward design is associated with the transition from a preliminary formed holistic concept to its representation as a set of private concepts of various initial degree of content formation. The methodological basis of this stage is the position of the axiomatic property 3.

The stage of intra-level design is responsible for establishing the relation of immediate precedence

among the received fully formed private concepts. Methodologically this stage is provided by the positions of theoretical propositions 1-3.

Proposition 1. If the set $Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ of the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is equal to the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$ of the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$, then the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is the only immediately preceding concept for the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$: :

$$\begin{aligned}&\forall\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}\ \forall\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}\\&((Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r} = X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}) \Rightarrow\\&\Rightarrow (\exists!\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r},\\&(TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}\ P\ TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s})))\end{aligned} \tag{18}$$

Proof. By definition 3 it follows that the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is in the relation P with the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$.

Uniqueness is proved by the method of the opposite. Suppose that there is some fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q}$, which is also in the relation P with the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$. However, as a result of comparison of definition 3 for the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q}$ with the first part of the proof of this proposition there is a contradiction with the position of the axiomatic property 1. Therefore, the assumption made is incorrect and the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is the only immediately preceding concept for the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$.

Proposition 2. If the set $Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ of the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is an own subset of the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$ of the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$ and there is no set of such fully formed concepts $\sum_{k=1}^{n} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_k}}$, that the totality of sets $\sum_{k=1}^{n} Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_k}}$ is a subset of the complement $\setminus X^*_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$ of the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$ of the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$, then the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is the only immediately preceding concept for the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$:

$$\begin{aligned}&\forall\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}\ \forall\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}\\&(((Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r} \subset X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}) \wedge\\&\wedge (\bar{\exists} \sum_{k=1}^{n} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_k}},\\&(\sum_{k=1}^{n} Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_k}} \subseteq (X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s} \setminus\\&\setminus X^*_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s})))) \Rightarrow (\exists!\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}\\&(TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}\ P\ TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s})))\end{aligned} \tag{19}$$

Proof. By definition 3 it follows that the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is in the relation P with the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$.

The proof of uniqueness is constructed by correlating the position of the axiomatic property 4 and the conditional part of this proposition for the fully formed concepts $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ and $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$. As a result of the analysis, it turns out that the subset of the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$ is exclusively the set $Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$, which by the axiomatic property 1 is unique. Thus, the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is the only immediately preceding concept for the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$.

The situation when such set of fully formed $\sum_{k=1}^{n} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_k}}$ exists, the absence of which is stated in the position of the proposition 2, is considered in the proposition 3. In that case, the investigative part of the position of the proposition 3 contains a conclusion about the multiplicity of immediate precedence to the fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$. This proposition is proved by a similar scheme of the given proof of the proposition 2.

The final stage of UDC construction is the stage of upward design, which is the establishing relation of "part-whole" between each of the fully formed private concepts and the already fully formed holistic concept. The implementation of the corresponding algorithm is methodologically determined by the position of the proposition 4.

Proposition 4. If all the concepts $\mathop{Y}_{\mu} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}$ are fully formed, and for any concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ there is no such concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q}$, that the set $Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q}$ of the concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q}$ is not a subset of the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ of the concept $TD_{\alpha_i,\beta_i,\ldots,\lambda_m,\mu_r}$, and there is also no such concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$, that the set $Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ of the concept $D_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}$ is not a subset of the set

$X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$ of the concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}$, then each of the concepts $\underset{\mu}{\mathrm{Y}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}$ is a private concept for the fully formed holistic concept $TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$:

$$\begin{aligned}
&\exists\, TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}\ \exists \underset{\mu}{\mathrm{Y}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}\\
&((\forall\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}\\
&\bar{\exists}\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q}\ \bar{\exists}\, TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}\\
&((\neg(Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_q} \subseteq X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r}))\wedge\\
&\wedge(\neg(Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_r} \subseteq X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_s}))))\Rightarrow\\
&\Rightarrow(\forall \underset{\mu}{\mathrm{Y}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}\, F\, TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}))
\end{aligned} \tag{20}$$

Proof. There is a sufficient number of possible cases in the proof of this proposition, which is associated to the total quantity and mutual arrangement of the fully formed concepts $\underset{k=1}{\overset{n}{\mathrm{Y}}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_k}}$. In this regard, the proof will be conducted on the example of "Fig. 1", when $n=4$.

The determination of the complete formation of the holistic concept $TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$ is based on the conditional part of this proposition and the positions of the axiomatic properties 6-8, and there are some special features. They are associated with the application of the position of the axiomatic property 6, according to which it turns out that: a) the set $X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$ of the holistic concept $TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$ is supplemented by the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_1}}$ of the fully formed private concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_1}}$, which by definition 4 is initial;

b) since the fully formed private concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_1}}$ is the only immediately preceding concept for the fully formed private concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_2}}$, at that $Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_1}} = X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_2}}$, the set $X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$ from the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_2}}$ is not replenished by anything; c) the set $X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$ of the holistic concept $TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$ is supplemented by the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_3}}$ of the fully formed private concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_3}}$, which by definition 6 is complementary; d) since the union of the sets $Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_2}}$ and $Y_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_3}}$ is an own subset of the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_4}}$, the set $X_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$ of the holistic concept $TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$ is supplemented by the complement $\setminus X^*_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_4}}$ of the set $X_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_4}}$ of the fully formed private concept $TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_4}}$.

Thus, by definition 7 each of the fully formed concepts $\underset{k=1}{\overset{4}{\mathrm{Y}}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu_{v_k}}$ is in the relation of "part-whole" with the already fully formed concept $TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$, while the requirements of the axiomatic property 2 are also met for the fully formed holistic concept $TD_{\alpha_i,\beta_j,\ldots,\eta_l,\lambda_m}$.

For other quantities and mutual arrangements of the fully formed concepts $\underset{\mu}{\mathrm{Y}} TD_{\alpha_i,\beta_j,\ldots,\lambda_m,\mu}$, the proof of this proposition will be constructed on a similar scheme.

V. CONCLUSION

By the results of consideration of the proposed models' features and the constructed algorithms it can be concluded that as a result the method of formalized description of technologies was developed, differing from the existing methods by presence of the possibility of the formation of the DST by means of the staged combined design of the UDC, with the purpose of structuring of knowledge representation about technologies with various degree of detail.

The distinctive feature of the proposed model of the concept of technological action is the concentration of all the semantics of technological actions, located in the nodes of the DST, exclusively within the sets that are part of the structure of these conceptual formations.

The specificity of the constructed model of the formalized description of technologies is in organizing the process of the formation of the DST, with the purpose of reception of analytical tools for ontological representations of technologies by determining and establishing all the entered relations, based on the constructive features of the embedded model of the concept of technological action. The axiomatic component of this model determines the basic laws of UDC construction, from which the DST are already formed.

The distinctive features of the developed algorithms include: constructing ontological hierarchies of concepts basing on predetermined signs of decomposition, automatic establishment of relations between concepts of the same level of decomposition and automatic receiving full information about the concepts located in the root nodes of the DST, based on the implemented principle of level-by-level aggregation of knowledge.

References

[1] R. Koller and N. Kastrup, Prinziplösungen zur konstruktion technischer produkte. Berlin: Springer, 1994.

[2] R. J. Van Wyk, „Technology: a fundamental structure," Knowledge, Technology, and Policy, vol. 15 (3), pp. 14–35, 2002. V. V. Sadovsky, M. V. Samoilov, and N. P. Kokhno, Production technologies. Minsk: BSEU, 2008. (in Russian)

[3] D. A. Makhotin, "Technological knowledge in modern education," Bulletin of Moscow state regional University, Pedagogics, No. 2, pp. 116-121, 2010. (in Russian)

[4] D. A. Andreev, "Methods of formalized description of technologies: an attempt to overview," Proceedings of Pskov Polytechnic Institute, Mechanical engineering. Electrical engineering, No. 15.3, pp. 291-297, 2011. (in Russian)

[5] S. Staab and R. Studer, Handbook on ontologies. Berlin: Springer, 2004.

[6] A. V. Palagin and N. G. Petrenko, "System-ontological analysis of the subject area," Control systems and machines, No. 4, pp. 3-14, 2009. (in Russian)

[7] I. V. Antonov and M. V. Voronov, "Methods of data analysis in the automation tasks of domain ontology," Distance and virtual training, No. 8, pp. 19–35, 2011. (in Russian)

[8] G. B. Evgenev, "The ontological methodology for creating intelligent systems in mechanical engineering," Proceedings of higher educational institutions, Mechanical engineering, vol. 6 (651), pp. 39–47, 2014. (in Russian)

[9] V. B. Artemenko and A. A. Ivlev, "The ontology of military technologies based on conceptual maps," Investigated in Russia, vol. 14, pp. 285–294, 2011. (in Russian)

[10] R. A. Calvo and J. Villalon, „Concept maps as cognitive visualizations of writing assignments," Educational technology and society, vol. 14 (3), pp. 16–27, 2011.

[11] R. Carnap, „Empiricism, Semantics, and Ontology," Revue Internationale de Philosophie, vol. 4, pp. 20–40, 1950.

[12] V. F. Turchin, „The cybernetic ontology of actions," Kybernetes, vol. 22 (2), pp. 10–30, 1993.

[13] D. A. Andreev, "Model of conceptual elements of activity in the formalized descriptions of production technologies," in proceedings of the XXXV International scientific and practical conference on prospects of development of information technologies, Novosibirsk, 2017, pp. 6-13. (in Russian)

[14] D. A. Andreev, "Decomposition representation of knowledge as the basis of ontological engineering of technological processes," in proceedings of the second international scientific and technical conference on computer science and technology, Belgorod, 2011, pp. 140-144. (in Russian)

[15] T. R. Gruber, „A translation approach to portable ontology specifications," Knowledge acquisition, vol. 5 (2), pp. 199–220, 1993.

[16] S. V. Maltseva, "Application of ontological models for solving identification problems and domain monitoring," Business Informatics, No. 3, pp. 18-24, 2008. (in Russian)

[17] B. V. Dobrov and N. V. Lukashevich, "Transitive non-taxonomic relations in ontological modeling," in proceedings of the symposium on ontological modeling, Zvenigorod, 2008, pp. 229-259. (in Russian)

[18] D. A. Andreev and M. V. Voronov, "Model of unified constructions of technology description in their ontological representations," in proceedings of the XXVI international scientific conference on mathematical methods in engineering and technologies, Nizhny Novgorod, 2013, vol. 8, pp. 138-140. (in Russian)

[19] D. A. Andreev, "Axiomatics of the model of formalized description of arbitrary technology," in proceedings of the XXVIII international scientific conference on mathematical methods in engineering and technologies, Saratov, 2015, vol. 4, pp. 77-80. (in Russian)

[20] D. A. Andreev and M. V. Voronov, "Procedural mechanism for constructing ontological representations of the decomposition structures of technologies," in proceedings of the XV all-Russian joint conference on the Internet and Modern Society, Saint-Petersburg, 2012, pp. 37-40. (in Russian)

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II,28-33

Development of Software for Design Ontological Representations of Production Technologies

Dmitry Andreev
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
dandreev60@mail.ru

Sergey Lyokhin
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
slyokhin@gmail.com

Victor Nikolaev
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
nvv60@mail.ru

Olga Poletaeva
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
oapoletaeva@mail.ru

***Abstract*—The features of the progressive methodological basis for the construction of formalized descriptions of technologies for their subsequent analysis are discussed. The existing possibilities of ontological knowledge engineering applied to the problem of the structural representation of technologies using computer tools are noted. Presented functionality of the developed software that allows automating staged algorithmic procedures for constructing unified decomposition structures of formalized descriptions of technologies. Analysed following development tools: Java Standard Edition programming platform, "Eclipse" IDE, Java programming language, PostgreSQL database management system, "Swing" library for creating a graphical interface and "JGraphX" library for graphs visualization. Database structure of the developed software is described: shown database schema, database tables are defined and the links between them are indicated. The architecture of the developed software is presented: shown data flow scheme and the purpose of each of the modules is described. The main advantages of the developed software "OntoTechnology" is designated, which shows the practical significance of the results.**

Keywords — concept, link, module, table.

I. Introduction

In recent years, methodological conceptions based on ontologies received development in the field of formalization of technological knowledge based on their structural representation. This fact looks logical, since it is ontologies that provide the possibility of comprehensive and detailed formalization of a certain field of knowledge by means of conceptual schemes, which are the systems of concepts interconnected according to certain rules [1] – [3].

The ontology as a generalized scheme of knowledge representation is based on various methods of knowledge conceptualization and methodological considerations of the development of tools for their analysis [4], [5]. Conceptualization is one of the most important processes of human cognitive activity, it implies processing the incoming information and leads to the formation of concepts, conceptual structures and the entire conceptual system in the human mind. Its purpose is constructing an abstract model that determines the structure of the simulated field of knowledge, the properties of its components and causal relations connecting them [6]. Thus, such conceptual model first and foremost consists of certain cognitive structures of special knowledge (mental essences, notions, concepts) and the relations between them.

Technologies need obligatory ontologization before their application, thus, if the knowledge about these technologies goes beyond anthropocentric representations, ontologization becomes the only way of mastering the essence of technologies [7].

The activity of many modern specialists in the field of ontological knowledge engineering varies from the declarative approach while defining possible models of concepts ontologies [8] – [11] and to the implementation of procedural mechanisms for the automated construction of subject fields ontologies [12]. Most of the existing software ontological design tools are aimed at constructing an ontological hierarchy of objects [13], which at best can reflect only the essential aspect of the components of material nature used in the implementation of technologies. In this regard, the usability of such systems, regarding their possible application to display the structural representation of technologies, is significantly limited, since the "coverage

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4064
© 2019 Dmitry Andreev, Sergey Lyokhin, Sergey Verteshev, Lilia Motaylenko.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

area" of technological knowledge is, among other things, a set of actions aimed at transforming objects in conditions of specific production.

This article describes the development of software that provides effective operation of technological knowledge through the design of appropriate ontological representations for the subsequent analysis of production technologies.

II. FUNCTIONALITY AND THE SOFTWARE DEVELOPMENT TOOLS

For the practical application of the developed method [14], the software was designed to automate staged algorithmic procedures for constructing unified decomposition structures (UDC) of formalized descriptions of technologies.

The developed software has the following functionality:

- Creating a new project, loading a previously saved project, deleting a project that contains a formalized description of the technology;
- Adding, editing, deleting concepts of technological actions, located in the nodes of the decomposition structures of technology (DST);
- Determining the initial degree of content formation of each of the concepts of technological actions;
- Managing the number of private concepts within the UDC based on the formed signs of decomposition;
- Automatic relations establishment between fully formed private concepts within the UDC;
- Automatic determining the complete formation of a holistic concept within the UDC;
- Displaying the results of the design using the library for visualization graphs JGraphX;
- Importing and exporting projects in XML format (eXtensibleMarkupLanguage).

Java Standard Edition (SE) programming platform, "Eclipse" IDE, Java programming language, PostgreSQL database management system (DBMS), "Swing" library for creating a graphical interface and "JGraphX" library for graphs visualization were chosen to develop the software. Currently all these development tools are among the most advanced ones.

Java SE platform is the standard platform Java version 2, designed to create and execute applets and applications intended for individual use and for use in scales of small and medium enterprises [15]. It should be noted that this platform is mainly intended to develop and run desktop applications that do not require preliminary installation on a work computer to start working with them. The advantages of the platform include following features: ability to run applications under the control of most modern operating systems, high reliability and security, portability and high performance.

"Eclipse" IDE is positioned as a free integrated development environment for modular cross-platform applications [16]. It is a fully-fledged Java IDE (Integrated Development Environment) used by a huge community of software developers and it is the corporate standard for application development in many organizations. The main advantage of this design environment is the ability to connect a variety of extensions (modules, plug-ins, etc.) that extend the functionality of the environment for specific practical needs of the developer (for working with databases, application servers, etc.).

"Java" language is distinguished by effective support for the object-oriented programming paradigm [17]. Java programs are translated into a byte code executed by the Java virtual machine (a program that processes byte code and sends instructions to the hardware as an interpreter). The advantage of this method of program execution is the complete independence of the byte code from the operating system and hardware, which allows running Java-applications on any device that has a corresponding virtual machine. Another important feature of the Java language is the flexible security system because the execution of the program is fully controlled by the virtual machine. Any operations that exceed the established permissions of the program (an attempt of unauthorized access to data, connections to another computer, etc.) cause an immediate interruption. The main features of the language are automatic memory management, advanced exception handling, a rich set of I/O filtering tools, a wide range of standard collections, tools for creating multi-threaded applications built into the language and unified access to databases.

PostgreSQL is a free object-relational DBMS [18]. The strengths of PostgreSQL are: support for databases of virtually unlimited size, powerful and reliable transaction and replication mechanisms, an extensible system of embedded programming languages, inheritance and easy extensibility. It should be noted that the use of the ORM (Object-Relational Mapping) method was preferable for working directly with the database. It is a programming technology that allows associating databases with concepts of object-oriented programming languages such as Java, and work with database tables as classes, and with records in tables as objects. This approach allows avoid binding an application to a specific database, but instead, using the application for various database solutions.

"Swing" is a powerful library of graphical components (buttons, input fields, tables, etc.) for creating an advanced user interface. "Swing" refers to the Java Foundation Classes (JFC), which is a set of libraries for developing graphical shells [19]. "Swing" library components support specific dynamically-connected views and behaviours that make it possible to adapt to the graphical interface of the platform, i.e. to the component you can dynamically connect another one, which is specific to the operating system, including the type and behaviour created by the programmer. Therefore, applications that use the Swing library look like native applications for this operating system. Thus, the positive side of such components is the versatility of the interface of the created applications on all platforms.

"JGraphX" is a freely distributed library written in Java and fully compatible with "Swing", which provides the mathematical apparatus of the graph theory [20]. This library is designed to visualize various representations of entities and their relations, including undirected graphs, oriented graphs, subgraphs, multigraphs, graphs with

parallel arcs, etc. The main advantages of JGraphX include the fact that the library allows using different vertex positioning algorithms, as well as creating graphs based on widely used formats, such as XML-documents.

III. DATABASE STRUCTURE

During software development stage tables that comprise the required database were designed. The database schema of the software has the form shown in "Fig. 1".

The database has a relational structure, so data on entities of a formalized description of the technologies are stored in the form of tables consisting of rows (records) and columns (fields). The concept of the primary key (PK), which is a set of fields that uniquely defines a record was also used in the implementation of the design process.

The design of the tables that make up the database of the software was conducted in two stages:

1) Object and connected tables were constructed to describe real entities and their relations with each other;

2) The decomposition of the obtained tables was carried out in accordance with the rules of normalization. As a result, each table began to correspond to three normal forms: 1NF, 2NF and 3NF [21].

To link most of the tables of the designed database, one-to-many relations were used, each of which was represented graphically as a line with symbols at opposite ends "1" и "∞". In cases where two tables were in the potential many-to-many relation, in order to preserve the integrity of the data, link tables were created. Such link tables mainly consisted of two tables' records identifiers only.

Examples of such kind of tables in the presented database schema are:

- "Links_concepts_preceding_ concepts";

"Resulting_components_private_ concepts";

- "Links_concepts_source_ components";
- "Links_concepts_invariant_ components";
- "Values_cost_characteristics".

The database structure of the software is determined by the following tables.

Table I stores information about projects that contain formalized descriptions of technologies. Each record in this table contains information about one project, which is one formalized description of a certain technology.

TABLE I. PROJECTS

Field name	Data type	Description
project_id (PK)	bigint	project identifier
project_name	character vary-ing(255)	name of project
root_concept_id	bigint	identifier of the root concept of the DST

Table II stores information about concepts that are located in the nodes of the DST. Each record in this table contains information about one of these concepts.

TABLE II. CONCEPTS

Field name	Data type	Description
concept_id (PK)	bigint	concept identifier
concept_name	character vary-ing(255)	name of the concept
level	integer	the level of the DST to which the concept belongs
number	integer	an index that uniquely deter-mines the place of the concept in the DST
own_characteristic	character vary-ing(255)	constant character-istic peculiar to the concept
fully_formed	boolean	a logical flag designed to capture the fact that the concept is fully formed
resulting_ compo-nent_id	bigint	identifier of the re-sulting component of the concept
holistic_concept_id	bigint	identifier of the holistic concept for this concept
project_id	bigint	project identifier

Table III stores information about the components of concepts that are located in the nodes of the DST. In this case "components" mean resulting components of the concepts, and their source components. Each record in this table contains information about one certain component.

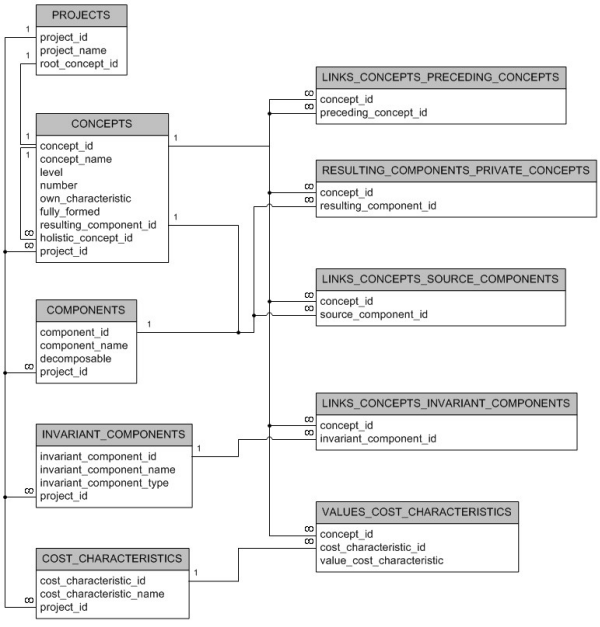


Fig. 1. The database schema of the software

TABLE III. COMPONENTS

Field name	Data type	Description
component_id (PK)	bigint	component identifier
component_name	character varying(255)	name of the component
decomposable	boolean	a logical flag designed to capture the fact of the presentation of a component of one concept as a set of components of other concepts
project_id	bigint	project identifier

Table IV stores information about the invariant components of concepts that are located in the nodes of the DST. Each record in this table contains information about one certain invariant component.

TABLE IV. INVARIANT_COMPONENTS

Field name	Data type	Description
invariant_ component_id (PK)	bigint	identifier of the invariant component
invariant_ component_name	character varying(255)	name of the invariant component
invariant_ component_type	character varying(255)	type of invariant component
project_id	bigint	project identifier

Table V stores information about the cost characteristics of the concepts located in the nodes of the DST. Each record in this table contains information about one cost characteristic.

TABLE V. COST_CHARACTERISTICS

Field name	Data type	Description
cost_characteristic _id (PK)	bigint	identifier of the cost characteristic
cost_characteristic _name	character varying(255)	name of cost characteristic
project_id	bigint	project identifier

Table VI stores information about the links of private concepts of the same level of decomposition, which is identified with the establishment of the relation of immediate precedence between them. Each record in this table contains information about one binary link of private concepts of the same level of decomposition.

links_concepts_preceding_concepts

Field name	Data type	Description
concept_id (PK)	bigint	identifier of the private concept
preceding_concept _id (PK)	bigint	identifier of the preceding private concept

Table VII stores information about links of private concepts of the same level of decomposition with their resulting components. Each record in this table contains information about one such binary link.

TABLE VI. RESULTING_COMPONENTS_PRIVATE_CONCEPTS

Field name	Data type	Description
concept_id (PK)	bigint	identifier of the private concept
resulting_ component _id (PK)	bigint	identifier of the resulting component of the private concept

Table VIII stores information about the links of private concepts of a certain level of decomposition with their source components. Each record in this table contains information about one such binary link.

TABLE VII. LINKS_CONCEPTS_SOURCE_COMPONENTS

Field name	Data type	Description
concept_id (PK)	bigint	identifier of the private concept
source_component _id (PK)	bigint	identifier of the source component of the private concept

Table IX stores information about the links of private concepts of a certain level of decomposition with their invariant components. Each record in this table contains information about one such binary link.

TABLE VIII. LINKS_CONCEPTS_INVARIANT_COMPONENTS

Field name	Data type	Description
concept_id (PK)	bigint	identifier of the private concept
invariant_component _id (PK)	bigint	identifier of the invariant component of the private concept

Table X stores information about the links of private concepts of a certain level of decomposition with their cost characteristics, which allows taking into account specific numerical values of these cost characteristics. Each record in this table contains information about one such binary link additionally with the information about one numeric value of the corresponding cost characteristic.

TABLE IX. VALUES_COST_CHARACTERISTICS

Field name	Data type	Description
concept_id (PK)	bigint	identifier of the private concept
cost_characteristic _id (PK)	bigint	identifier of the cost characteristic of the private concept
value_cost_ characteristic	integer	numerical value of the cost characteristic

THE ARCHITECTURE OF THE SOFTWARE

In accordance with the tasks that the software solves, its modular structure and the data flow scheme presented in “Fig. 2”.

The architecture of the developed software is determined by the following modules.

1) The user interface module. This module allows

user managing the software by means of graphical components in the form of buttons located in the right part of the working window of the program; it allows user to enter, edit and delete data; getting information about the structure of the designed formalized description of a certain technology in two modes of viewing: the main view of each of the UDC and the general view of the entire DST.

2) The dialog module. This module is responsible for displaying auxiliary dialog boxes and for implementing mechanisms for their program interaction with the main user form.

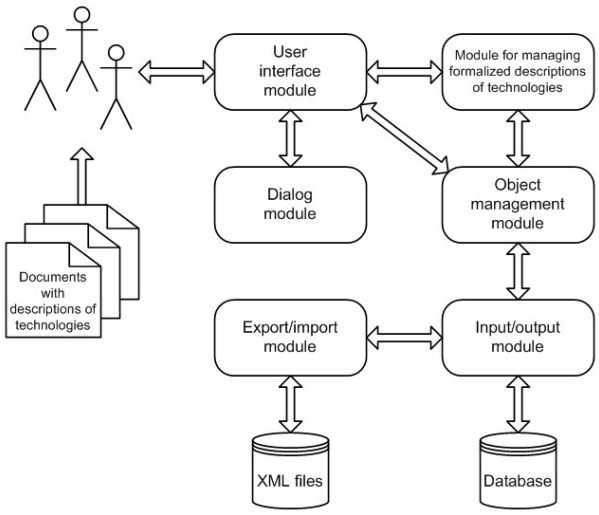


Fig. 2. Data flow scheme

3) The module for managing formalized descriptions of technologies. This module determines the implementation of the basic algorithms of management (formation, structuring and processing) by ontological representations of technologies including all algorithms that implement the method for constructing formalized description of technologies.

4) The object management module. This module allows working with the entities of the ontological representations of technologies as with objects, and allows user to abstract from the structure of storage of the corresponding information in the database.

5) The input/output module. This module is responsible for saving all information about the created or modified formalized descriptions of technologies in corresponding tables of the database. It also allows loading necessary ontological representations of technologies and all related information from the database.

6) Export/import module. This module provides export, i.e. the functionality to save formalized descriptions of technologies in the widely known and practical format of the XML markup language. This allows user to work with the created ontological representations of technologies in other software products that support this format. This module also provides import, i.e. downloading, formalized descriptions of technologies, which are prepared either by the same software, but installed on other workstations, or by other software products that support XML documents format. So, this export/import module provides universal portability of the received design results.

IV. CONCLUSION

The developed software "Ontotechnology" [22], implements new possibilities related to the automation of the process of construction of formalized descriptions of technologies. It supports the procedure for constructing ontological representations of technologies by direct participation of the expert in determining the initial degree of content formation of each of the concepts of technological actions with an explicit indication of their location in the nodes of the DST. The proposed solution allows:

a) Improving the stage of design-technological preparation of production in the part of concentration of processes of information processing necessary for drawing up of the current technical documentation for the technological processes of the enterprise, within one computer program;

b) Increasing the share of automatic procedures in the construction of ontological representations of technologies in comparison with the existing software analogues of this class of systems;

c) Reducing time costs and the need for labour-intensive manual work to obtain aggregate information on technologies, as well as creating new opportunities for rapid obtaining a necessary set of characteristics of the technologies under consideration;

d) Displaying all stages of construction of formalized description of technologies in the form of visual graphic images and providing portability of design results in the format of XML documents supported by most modern information systems.

REFERENCES

[1] J. F. Sowa, „Conceptual graphs as a universal knowledge representation," International journal computers & mathematics with applications, vol. 23 (2-5), pp. 75–94, 1992.

[2] F. F. Vyakkerev, V. G. Ivanov, B. I. Lipsky, and B. V. Markov, Foundations of ontology. Saint-Petersburg: Saint-Petersburg State University, 1997. (in Russian)

[3] A. F. Tuzovsky, S. V. Chirikov, and V. Z. Yampolsky, Knowledge management systems (methods and technologies). Tomsk: NTL, 2005. (in Russian)

[4] T. R. Gruber, „Toward Principles for the design of ontologies used for knowledge sharing," International journal human-computer studies, vol. 43, pp. 907–928, 1992.

[5] I. P. Norenkov, "Intellectual technologies based on ontologies," Information technologies, No. 1, pp. 17-23, 2010. (in Russian)

[6] E. S. Kubryakova, V. Z. Demyankov, Yu. G. Pankratz, and L. G. Luzina, Short dictionary of cognitive terms. Moscow: Philology Department of Moscow State University named after M. V. Lomonosov, 1997. (in Russian)

[7] S. A. Datsyuk, "Ontologization," 2009. [Online]. Available: http://lit.lib.ru/d/dacjuk_s_a/text_0030.shtml [Accessed: Feb. 18, 2019]. (in Russian)

[8] T. A. Gavrilova and V. F. Khoroshevsky, Knowledge bases of intellectual systems. Saint-Petersburg: Piter, 2000. (in Russian)

[9] A. V. Abramov, "Ontology as a method of describing subject areas," Bulletin of Moscow city pedagogical University, Informatics and informatization of education, No. 7, pp. 204-206, 2006. (in Russian)

[10] V. I. Mezhuyev, "Using ontologies as models of subject areas," Artificial intelligence, No. 4, pp. 4–11, 2009. (in Russian)

[11] N. S. Konstantinova and O. A. Mitrofanova, "Ontologies as systems of knowledge storage," 2008. [Online]. Available: http://window.edu.ru/resource/795/58795/files/68352e2-st08.pdf [Accessed: Feb. 18, 2019]. (in Russian)

[12] I. V. Antonov, Method of automated construction of domain ontology, PhD in Technical Sciences [thesis]. Pskov: PSPI, 2011.

(in Russian)

[13] O. M. Ovdey and G. Yu. Proskudina, "Overview of ontology engineering tools," Electronic libraries, vol. 7, issue 4, 2004. [Online serial]. Available: http://rcdl.ru/doc/2004/paper26.pdf [Accessed: Feb. 18, 2019]. (in Russian)

[14] D. A. Andreev and M. V. Voronov, "Method for constructing an ontology of technological actions," Bulletin of Saratov State Technical University, No. 3 (67), pp. 160–168, 2012. (in Russian)

[15] H. Schildt, Eds., Java. Complete guide. Moscow: Williams, 2012. (in Russian)

[16] D. Carlson, Eclipse Distilled. Moscow: Lori, 2008. (in Russian)

[17] B. Eckel, Eds., Thinking in Java. Saint-Petersburg: Piter, 2009. (in Russian)

[18] G. Smith, PostgreSQL 9.0. High Performance. Birmingham: Packt Publishing, 2010.

[19] I. A. Portyankin, Eds., Swing. Spectacular user interfaces. Moscow: Lori, 2011. (in Russian)

[20] JGraph user manual, JGraph Ltd., 2009.

[21] S. D. Kuznetsov, Eds., Database basics. Moscow: INTUIT; BINOM. Knowledge laboratory, 2007. (in Russian)

[22] D. A. Andreev, "Program for automated construction of a formalized description of the technology of the applied field of knowledge OntoTechnology," Certificate of state registration of computer programs No. 2013660420 the Russian Federation, 5 Nov., 2013. (in Russian)

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II,34-37

The Method of Automated Building of Domain Ontology

Igor Antonov
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
igorant63@yandex.ru

Iuliia Bruttan
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
bruttan@mail.ru

Dmitry Andreev
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
dandreev60@mail.ru

Lilia Motaylenko
Faculty of Computer Science and Electrical Power Engineering
Pskov State University
Pskov, Russia
lvs@mail.ru

***Abstract*—This article is devoted to the tasks of automating the construction of domain ontologies. In the beginning, the limitations and problems of constructing the ontology of the domain using the well-known methods are discussed. Next, a model of the domain ontology is proposed, which provides the ability to automatically build the ontological hierarchy, including the automatic synthesis of generalized concepts. Then, the article discusses the method of building an ontology based on the proposed model using machine learning, and discusses its capabilities and limitations.**

Keywords—domain ontology, information systems, information retrieval, data mining, machine learning.

I. Introduction

The concept of ontology came to information technology from philosophy. In the information disciplines ontology is currently understood as a formalized model of a system of concepts covering a certain subject area. Such model includes a formal representation of the content of concepts and relationships between them. Ontologies have practical application in various fields. Among other things, ontologies are the means of documenting a consistent point of view of specialists on the structure and terminology of the subject area. Ontologies are used as part of software tools that provide information support for the design of complex technical systems. Also, ontologies are used as teaching materials, clearly demonstrating the system of connections between the concepts of the domain, thus presenting it as a coherent systemic whole [1].

Information systems developed on the basis of ontologies have shown their effectiveness in practice and interest in them is constantly growing. However, the existing methods of ontology building require expert knowledge in the studied domain, and the construction of ontologies based on their use takes a significant amount of time, therefore, the actual task is to automate the process of ontology building.

II. Limitations and Problems of The Methods for Automated Ontology Building

There are several areas in which research is being conducted for creating means for building and maintenance of domain ontologies automating.

The authors of [2] - [11] propose methods for constructing ontologies based on automatic processing of the content of web resources and natural language texts. As a priority source of source data, the authors often point out Wikipedia as a source with a high density of meaningful information about the relationship between the concepts of the subject area. These papers consider the problem of statistical, syntactic and lexical text processing in order to extract data used in the construction of ontology. The task of a multi-level ontological hierarchy construction automating in the scope of these works is not being solved. The extraction of ontological information is performed by identifying the semantic relations of the terms of the subject domain based on the analysis of the syntactic structure of sentences. According to the results of statistical analysis, the most frequently used terms are singled out as candidates for inclusion in the ontology. The results of computer texts processing go through the stage of verification, adjustment and additions by the domain expert, after which they are used as an ontological model of the subject domain. The main difficulty of the practical use of the solutions considered in these works is that the ambiguity and contextual dependencies of natural language constructs limit the possibilities of their computer processing and do not provide completeness and correctness of extracting semantic information from texts using formal methods.

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4065
© 2019 Dmitry Andreev, Sergey Lyokhin, Olga Poletaeva, Victor Nikolaev.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

The works [12] - [18] analyze the possibility of using the tools developed within the framework of the Formal Concept Analysis (FCA) method for the automatic construction of the domain ontology. As initial data for the ontological analysis of the subject area, these works suggest using a two-dimensional "object - attributes" table, which is based on empirical data about the objects of the subject area or with results of performing certain procedures that reveal the attributes of the objects of the subject area, including data extraction operations from text arrays. Such "objects - attributes" table is transformed into the formal context of the FCA method by combining objects with homogeneous composition into classes. Based on the analysis of the formal context matrix, algorithms that implement the FCA method generate all possible formal concepts. Formal concepts are connected by hierarchical relations. Each concept in the hierarchy represents the objects sharing some set of properties; and each sub-concept in the hierarchy represents a subset of the objects (as well as a superset of the properties) in the concepts above it. The authors of the noted papers do not consider the problem of assessing the semantic significance of concepts included as a result of applying the FCA method to the domain ontology. When constructing ontologies intended for practical use, one should take into account that the hierarchy of concepts synthesized on the basis of the formal method will include formal concepts that do not have semantic significance and correspond to any repeated combinations of lower-level concept attributes. Therefore, the methods considered in these works do not allow synthesizing ontology directly suitable for use as domain ontology in real-world applications.

In general, we can conclude that at present there is a certain lack of technology and tools that provide automated construction and maintenance of domain ontologies. Further progress in this area requires the development of new methods for the automated construction of domain ontology and software tools that provide opportunities for automating the construction and maintenance of ontologies.

III. Domain Ontology Model

As a possible solution to a part of the problems discussed above, this paper proposes a domain ontology model that provides the ability to effectively automate the ontology construction process, and proposes a method for automated ontology construction. The ontology model is represented by a tuple (1).

$$Q=<C,M,R> \quad (1)$$

where $C=\{c_i\}$ - the set of concepts forming ontology, $1,I$;

$M=\{M_i\}$ - the set of sets of concept attributes;

$M_i=\{m_{1_i},...,m_{d_i}\}$ - the set of attributes describing the -th concept,

d_i- number of attributes describing the concept ;

$R \subseteq C \times C$- the relationship of the direct inheritance of concepts.

The relationship R can be defined by a matrix with dimension $I \times I$. If the concept c_k directly inherits from concept c_i, i.e. $(c_i,c_k) \in R$, then the matrix element $r_{ik}=1$, otherwise, if $(c_i,c_k) \notin R$, then $r_{ik}=0$. The relationship of direct inheritance between two concepts means that there are no intermediate concepts between them:

$$\exists c_i \exists c_k ((c_i,c_k) \in R) \Rightarrow \exists c_h (((c_i,c_h) \in R) \wedge ((c_h,c_k) \in R)) \quad (2)$$

If two ontology concepts are in inheritance relation, then the corresponding parent and child concepts in the proposed model are related by a strict partial order relation, denoted by '<'. This relationship is transitive:

$$\forall c_i \forall c_j \forall c_k ((c_i < c_j) \wedge (c_j < c_k) \Rightarrow c_i < c_k) \quad (3)$$

The proposed model differs from the traditional ontology model by the presence of a unique correspondence of a certain set of attributes to a certain ontology concept. The concepts related by inheritance in the proposed model always differ in the composition of their attributes. This solution allows to formalize the operation of incorporating a new concept into the ontological hierarchy. At the same time, mentioned above properties of the model do not create restrictions for the use of ontology based on this model within information systems.

IV. Method of Building of Ontology

The method of constructing an ontology based on the model discussed above allows to automatically determine the position of the inclusion of a new concept in the ontological hierarchy and automatically generate new generalized concepts caused by the addition of a new concept to the ontology.

The proposed method provides the implementation of an iterative ontology construction by successively entering into the ontology model new concept models c_x represented by a set of attributes M_x, where $M_x \subset M$. For newly introduced concepts, the place of the concept in the ontological hierarchy is determined automatically.

If the set of attributes of a new concept is not a superset of the set of attributes of any of the ontology concepts, then the concept will occupy a position at the level immediately following the root ontology concept c_0 (Fig.1):

$$\exists c_j \left((c_j \in C) \wedge (M_j \subset M_x) \right) \quad (4)$$

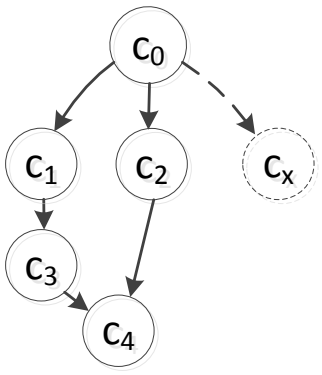


Fig. 1. Adding a new top-level concept.

In other cases, the new concept will have to inherit from any other concepts present in the ontology and inheriting from the concept c_0. To automatically determine the parent concepts, a sequential "top-down" traversal of the graph representing the ontology concept links is performed. In the process of traversing the graph, the parent concept for the concept c_x is determined based on the following criterion: "The concept c_x will have the parent c_p for which a set of attributes M_p is a subset of the set of attributes of the concept c_x , and the set of attributes of any of the child concepts of a concept c_p is not a subset of the set of attributes of the concept c_x":

$$((M_p \subset M_x) \wedge \bar{\exists} c_d((r_{pd} = 1) \wedge (M_d \subset M_x)) \Rightarrow r_{px} := 1 \quad (5)$$

If (5) is violated and $\exists c_d((r_{pd} = 1) \wedge (M_p \subset M_x) \wedge (M_d \subset M_x))$ then c_d is considered as the next candidate for parent concepts for the concept c_x and replaces c_p in the condition check operation (5). In this case, the search for the parent concept for the concept c_x in the general case is carried out along several branches of the graph. If condition (5) is fulfilled when traversing a branch of a graph, the search for this branch is complete at the concept c_p level, and the concepts indicated in (5) as c_p form a set of parent concepts of the concept c_x . If the set of attributes of the concept c_d in the case of truth of condition (5) is a superset of the set of attributes of the concept c_x , then c_d is a child of the concept c_x and is included in the corresponding set:

$$\forall c_d((r_{pd} = 1) \wedge (M_p \subset M_x) \wedge (M_x \subset M_d) \Rightarrow (r_{px} := 1, r_{xd} := 1, r_{pd} := 0)) \quad (6)$$

On the graph the concept c_x is placed in the «gap» between the concepts c_p and c_d (Fig. 2).

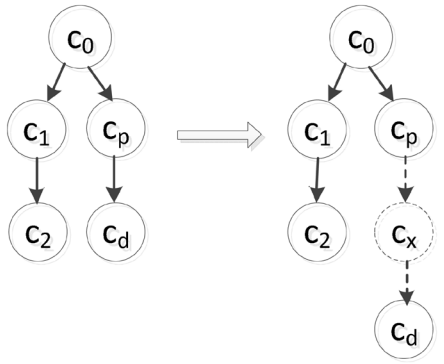


Fig. 2. Adding a new concept with breaking the inheritance link.

If the condition specified for c_d is not fulfilled, then c_x becomes a neighbour concept for the c_d, being located at the same level of the hierarchy of concepts and also inheriting from c_p (Fig. 3).

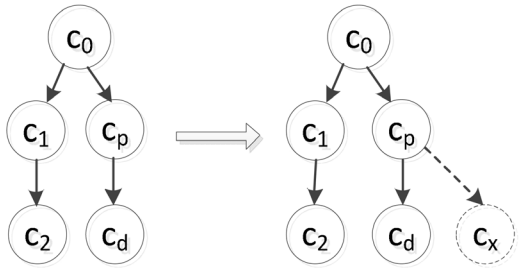


Fig.3. Adding a new child concept without breaking the inheritance relationship.

After determining the place of a new concept in the ontological hierarchy, new generalized concepts of higher levels can be automatically determined on the basis of identifying sets of attributes that recur in various concepts. The discovery of such sets of attributes that are not indicated in the ontology by any concepts means the appearance of new generalizations that can be included in the ontology as new concepts. Fig. 4 shows the inclusion of the concept c_x in the ontology with the formation of new generalizations c_g and c_g'.

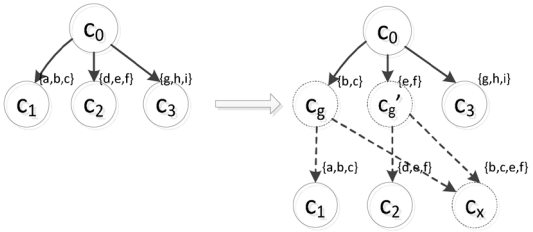


Fig.4. Inclusion of the concept in ontology with the formation of new generalizations.

Automatically detected generalizations before being included in the ontology should be evaluated by an expert for their semantic significance. If the expert considers the generic concept to be of no value, it is not included in the ontology and automatically considered later as ignored when performing the following inclusions of new concepts in the ontology. Iterative construction of a list of generalizations excluded from ontology implements machine learning for automatic recognizing concepts that are not of semantic significance.

When using the proposed method, the domain ontology is formed in the process of successively adding to the ontology models of concepts represented by a set of attributes characterizing the concept. Based on the formal criteria, an analysis of the composition of the new concept is performed and the place of the new concept in the ontological hierarchy is objectively determined. When a new concept is included in ontology, its links with parent and child concepts are automatically determined and new generalized concepts are formed, which, after evaluating their semantic significance, are either included in the ontology or supplement the list of non-valuable generalizations for the ontology. The method reveals the position of the concept in the ontology, regardless of the level that the concept should occupy in the general hierarchy. The position of a new concept can be located at the top level, at the bottom level, or at any of the

intermediate levels of the ontology. The method can be used to form the ontology from scratch, starting with an empty ontology, and to expand an existing domain ontology.

V. Conclusion

On the basis of the model and method proposed in the work, ontology editors can be created that provide automated support for ontological engineering operations. The proposed method provides ontology machine learning for automatic recognition of semantically insignificant generalized concepts, which makes it possible to exclude multiple manual processing of concepts that do not have semantic significance in the process of ontology construction. The method complements the traditional construction of ontology "from top to down" with the possibility of entering into the ontology of new concepts "from below" based on the input into the system the information about the attributive composition of the new concept. The source of information about the structure of the concepts of the lower levels can be empirical data about the objects of the subject area, in particular, the experimental data obtained as a result of measurements and studies of objects of the subject area. The method is limited in application in relation to top-level ontologies and humanitarian ontologies, where the non-formalized interpretation of concepts is mainly used.

References

I. V. Antonov and Iu. V. Bruttan, "Ontological approach to building control systems of electronic learning objects," Bulletin of Pskov State University, Economics. Law. Management, No. 5, pp. 230-234, 2014. (in Russian)E. A. Rabchevsky and G. I. Bulatova, "Automatic ontology construction," Scientific and technical bulletins of SPbSPU, No. 4, pp. 22-26, 2007. (in Russian)

[1] A. E. Ermakov, "Automation of ontological engineering in systems of knowledge extraction from text," Computer linguistics and intellectual technologies: in proceedings of The International Conference Dialogue, Bekasovo, June 4-8, 2004, Moscow: RSUH, 2008, pp.154- 159. (in Russian)

[2] Song, Q., Liu,J., Wang, X. A Novel Automatic Ontology Construction Method Based on Web Data: Proceedings of Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Kitakyushu, Japan, 2014, pp.762-765.

[3] Khalatbari, S., Mirroshandel, S. A.: Automatic construction of domain ontology using wikipedia and enhancing it by google search engine. J. Inf. Syst. Telecommun. 3, 2015, pp. 248-258.

[4] D. Küçük, Y. Arsla, Semi-automatic construction of a domain ontology for wind energy using Wikipedia articles: Renewable Energy, 62, 2014, pp. 484-489.

[5] Damir Jurić, Marko Banek, Zoran Skočir, Automated Construction of Domain Ontology Taxonomies from Wikipedia: International Conference on Database and Expert Systems Applications DEXA, Database and Expert Systems Applications, 2011, pp. 439-446.

[6] P. Buitelaar, P. Cimiano, Ontology learning and population: Bridging the gap between text and knowledge: Frontiers in Artificial Intelligence and Applications Series, Vol. 167, IOS Press, 2008.

[7] P. Cimiano , J. Völker, Text2Onto. NLDB, 2005, pp. 227–238.

[8] Dahab,M.Y.,Hassan, H.A., Rafea, A., TextOntoEx: Automatic Ontology Construction from Natural English Text: Expert Systems with Applications. Volume 34, Issue 2, 2008, pp. 1474–1480.

[9] Azevedo, R. R., Freitas, F., Rocha, R. G. C., Menezes, J. A. A., Oliveira Rodrigues, C. M., and F. P. Silva, G., An Approach for Learning and Construction of Expressive Ontology from Text in Natural Language: Proceedings of 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014, pp. 149-156.

[10] S. V. Smirnov, "Ontological analysis of domain modeling," Proceedings of Samara scientific center of RAS, Vol. 3, No. 1, pp. 62-70, 2001. (in Russian)

[11] V. A. Pronina and L. B. Shipilina, "Using relations between attributes to construct domain ontology," Control problems, No. 1, pp. 27–32, 2009. (in Russian)

[12] Ning, L., Guanyu, L., Li, S., Using formal concept analysis for maritime ontology building: International Forum on Information Technology and Applications 2, 2010, pp. 159–162.

[13] P. Cimiano, A. Hotho, G. Stumme, J. Tane, Conceptual Knowledge Processing with Formal Concept Analysis and Ontologies: Eklund, P.W. (ed.) ICFCA 2004. LNCS (LNAI), vol. 2961, Springer, Heidelberg, 2004, pp. 189–207.

[14] Obitko, M., Snsel, V., Smid, J., Ontology design with formal concept analysis: Proc. of the CLA2004 International Workshop on Concept Latticesand their Applications, 2004, pp. 1377–1390.

[15] H.M. Haav, An Application of Inductive Concept Analysis to Construction of Domain-specific Ontologies: Proceedings of the Workshop of VLDB2003, 2003, pp. 63-67.

[16] R. Bendaoud, A. Napoli, Y. Toussaint, Formal Concept Analysis: A unified framework forbuilding and refining ontologies: A. Gangemi and J. Euzenat editors, Knowledge Engineering: Practice and Patterns – Proceedings of the 16 International Conference on Knowledge Engineering and Knowledge Management, EKAW'2008, pp. 156-171.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 38-43

Research of Human Fatigue and Measurement Parameters for Workability Assessment

Matīss Eriņš
Riga Technical University
Scientific assistantRiga, Latvia
Matiss.Erins@rtu.lv

Oļesja Minejeva
Riga Technical University
Scientific assistant
Riga, Latvia
o.minejeva@inbox.lv

Zigurds Markovičs
Riga Technical University
Professor
Riga, Latvia
Zigurds.Markovics@rtu.lv

Juris Lauznis
Riga Technical University
Researcher
Riga, Latvia
Juris.Lauznis@rtu.lv

Raivis Kivlenieks
Riga Technical University
Scientific assistant
Riga, Latvia
tolocko2@gmail.com

***Abstract*—Human fatigue is reduced working capability for certain period of time as the result of unusual or prolonged workload. Fatigue arises when the body's energy requirements exceed its supply. Fatigue first manifests as reduced concentration capability causing movement coordination and precision disruption leading to decreasing workability. Fatigue is an object of research in physiology, psychology, work ergonomics, medicine, and biotechnology where each domain has a focus on mental fatigue. The functional state in the context of professional activity is defined as a complex of characteristics of functions and qualities that determine the fulfilment of labour activity. Furthermore, a comprehensive estimation of subject functional state in combination with other factors like subject self-assessment and objective performance tests (cognitive load tests) is a necessary input for the evaluation of workability and efficiency on task. The heterogeneous nature of fatigue as a systemic manifestation requires analysis of multiple key parameters which are relevant to the specific type. The current feasibility study focuses on human biological signal from electrical activity of heart, brain, muscles and skin potentials as well as temperature, position, and respiration to obtain diagnostic parameters reflecting the state of cardiovascular, muscles, and central nervous systems for physiological monitoring of vital signs. The fatigue physiological parameter and feature formalization aim to support the development of a platform with complex passive multi-level fatigue monitoring system and workability evaluation system designed in order to provide an integrated service.**

***Keywords*—human fatigue, functional state, workability.**

I. Introduction

Human fatigue is a construct of multiple components that are characterized from experience, physiology or performance [1] leading to the effects of fatigue on operations safety, mental performance, and attention. A wide range of sensors and methods designed for biomedical applications of monitoring vital parameters and algorithms for analysis of human physiological states exist to this date [2]. The areas of research focus on the pathological causes of fatigue as a symptom in differential diagnostics and evaluation of workability based on the physiological states analysis of dynamic monitoring.

The current study aims to formalize the parameters of sensors, subjective questionnaires, and active tests to reflect the physiological, subjective and objective measures of mental fatigue. The listed parameters are significant to the selected method of evaluation of a fatigue component and are categorized by human physiological subsystems. The research analyses methods to gather measurement parameters from different domains of interest in fatigue evaluation. The mental fatigue as a base type for this research is chosen in context with its application in cognitive workability evaluation corresponding the project requirements.

II. Conceptual Method for Mental Fatigue

Task-related fatigue formally is classified as central or peripheral, that later is mental or physical. This paper mainly focuses on mental fatigue, however, some physical fatigue components (psychological, motor-sensory) also used to evaluate the affecting mental performance of cognitive abilities. Mental fatigue is an inability to maintain optimal cognitive performance. The onset of mental fatigue during any cognitive activity is gradual and depends upon an individual's cognitive ability, manifests as a diminished capacity for work and possibly decrements in attention, perception, decision making, and skill performance [3].

Mental fatigue is a transient decrease in maximal cognitive performance resulting from prolonged periods of cognitive activity. It can manifest as somnolence, lethargy, or directed attention fatigue [4]. Fatigue as a multidimensional concept (Fig. 1) implemented in the World Health Organization's Classification of Functioning, Disability, and Health. The multidimensional concept of fatigue is integrated into the World Health Organization's

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4148
© 2019 Matīss Eriņš, Oļesja Minejeva, Zigurds Markovičs, Juris Lauznis, Raivis Kivlenieks.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

International Classification of Functioning, Disability, and Health (WHO-ICF), representing the effect of disease on body function and structure, activity and participation of the patient. Both subjective fatigue and physiological fatigue have an effect on activity and participation and are in most diseases related to health status and disease severity. Psychosocial factors have an influence on fatigue and on activity and participation [5].

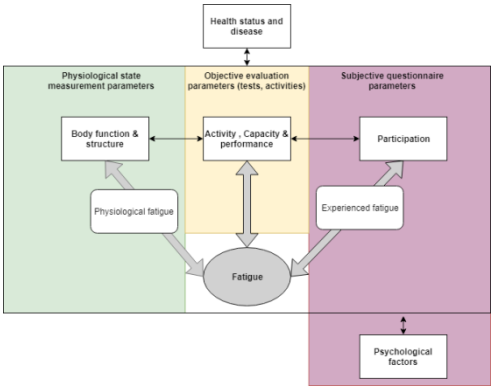


Fig. 1. Fatigue as a multidimensional concept. Parameters from three dimensions reflected in the conceptual model of functioning.

In addition to the model, the current paper focuses on three main workability evaluation perspectives reflecting the overall state of mental fatigue. The physiological states are manifestations known to the literature which serves as a measure for certain mental fatigue related physiological condition. Physiological measurements are grouped under physiological subsystems from which the biological feedback is received and relations to mental fatigue physiological states can be determined (Fig. 2). Further research will focus on the wearable and non-intrusive aspects of the sensor selection.

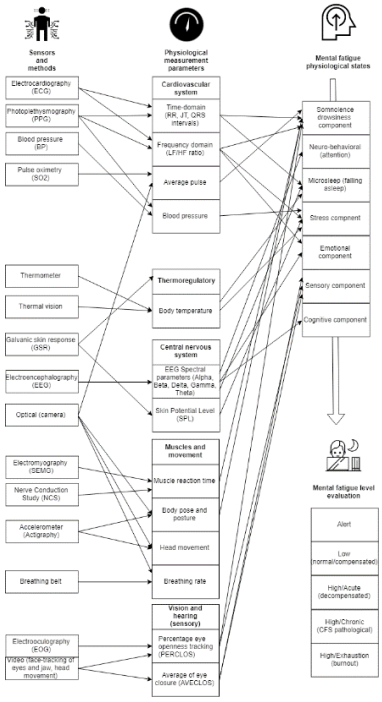
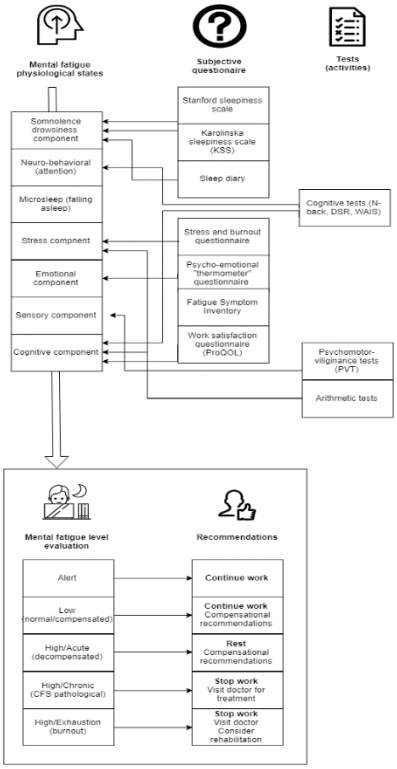


Fig. 1.Part I. Conceptual model of physiological parameters and relation to mental fatigue states.

Mental fatigue physiological states represent a set of indications which can be used in mental fatigue level evaluation. The later stages of the project will focus on the evaluation of fatigue level gradation [6] and possible recommendations (Fig. 3).



Part II. A conceptual model of subjective and test parameters and relation to mental fatigue states, recommendations.

III. Parameters

There are three aspects to fatigue: physiological, objective (work decrement), and subjective fatigue [7]. The process of parameter acquisition in the medical literature follows the steps of symptoms gathering in differential diagnostics [8] where fatigue is generalized as a symptom for at least 10 general diseases, including Chronic Fatigue Syndrome (CFS) [9]. The physiological measurements and tests in practice have dynamic properties, however, a subjective questionnaire can serve as a static measure obtained independently. The combinations of results obtained from three perspectives, which are further discussed in detail, introduce comprehensive human functional state evaluation.

A. Physiological parameters

Generally, human health state is defined by interdependent physiological parameters which respond to mental fatigue and are detectable by sensors. Selected parameters are grouped under corresponding measurable functional systems typically reflected in the literature and are mostly related to symptomatic of mental fatigue.

1) Cardiovascular system

As the heart rate variability (HRV) is coupled to autonomic nervous system activity, it provides a suitable proxy for examining how we feel. There are different methods of HRV analysis. One of the methods is time domain analysis. This method extracts a few special measures using only the temporal RR interval signals. Another method is spectral analysis. This

method interpolates the RR interval at a certain rate and transforms this interval into the frequency domain. There are some standards for these two methods [10]. There are also other methods such as Time-Frequency Domain and Nonlinear method. Frequency domain (LF/HF ratio) method is used in the context of human fatigue monitoring to detect or monitor the level of drowsiness [11]. Across diverse tasks and populations, [12] have found evidence for an association between higher levels of resting HRV and superior performance on tasks that tap executive functions. Of the frequency-domain methods, data related to the amount of low frequency (LF) heartbeats is often used (0.04 to 0.15 Hz) as a measure of sympathetic nervous system activity. Measurements of high-frequency (HF; 0.15 Hz to 0.4 Hz) and very-low-frequency (VLF) are also used. ECG recordings are therefore a clear measure of autonomic nervous system activity [13]. The single most common way to analyse HRV is a time-domain method called RMSSD. This is the Root Mean Square of Successive Differences between each heartbeat. HRV analysis is performed through assessment of time-domain indices, the square root of the mean of the sum of the squares of differences between adjacent normal R-R intervals (RMSSD) measured during short (5 min) recordings and particularly the logarithm of RMSSD (LnRMSSD) has been proposed as the most useful resting HRV indicator [14]. These results suggest that the frequency domain is related to psychological symptoms of mental fatigue [15]. The ECG carries information about a person's vigilance state. Hence, HRV measures could potentially be used to predict when an individual is at increased risk of attentional failure [16]. Respiration rate and blood pressure have a significant difference before and after fatigue: respiration rate decreased after fatigue. Blood pressure showed a significant upward trend after fatigue. The respiratory rate and blood pressure can be used as a field fatigue sensitive indicator [17].

2) Central nervous system

The impact of physical and mental work on the processes of human's nervous system has great practical importance. The activity of the human brain can be determined by electroencephalography (EEG) in the form of electric activity curves of brain cortex cells (neurons). Because the brain is the leading element of the nervous system, it can point to different physiological states of the human being, i.e. sleepiness and fatigue. The EEG is widely considered as the physiological 'gold standard' for the assessment of mental fatigue. Delta rhythm (slow waves) is associated with recovering processes, especially during sleep. Theta waves are associated with changes in the state of human consciousness. Theta rhythm can increase by severe cognitive work and a load of a complex task. The high amplitude of this rhythm may indicate a state of drowsiness and fatigue, as well as chronic stress [18]. Sensory, motor and memory functions may be reflected in the alpha rhythm. The lack of these waves can appear when a person opens his eyes, thinks about a task. At increased brain functional activity, the alpha-rhythm amplitude decreases and may disappear completely. It may also indicate anxiety, fear and other physiological conditions associated with increased activation of the vegetative and central nervous system. The fastest are beta-waves. They are associated with higher cognitive processes and attention focus when human is engaged in mental work and is focused on solving a problem. Rhythm is rapidly increasing with increasing work intensity and cognitive load [19]. Many literature sources show a gamma wave link with increasing attention [20] and semantic operations, cognitive performance [21]. As soon as a person wakes up, the gamma rhythm appears again. In 1937 - 1938 English scientists Lumis, Horvey, Habart, Devis made the first attempts to systematize curves and have described 5 stages of electroencephalographic sleep [22]. A number of researchers later showed that mental fatigue mainly manifests itself in a noticeable increase in alpha and theta rhythms in different areas of the brain, depending on the specificity of the performed work [21], [23]. To simultaneously evaluate fatigue after all EEG curves, the term fatigue index was introduced and represents the relationships between different EEG rhythms. Authors [23] used 3 indexes: theta / alpha, beta / alpha, alpha + theta / beta. As a result, the fatigue condition is best determined by the alpha + theta/beta ratio. This value rises sharply after long mental stress and is most sensitive to changes in human functional condition.

Galvanic Skin Response (GSR) reflects the variation in the conductivity of the skin and is measured in micro-Siemens (μS). GSR originates from the autonomic activation of sweat glands in the skin and is typically recorded from feet, palm or fingers on hands. Skin is responsible for bodily processes such as the immune system, thermoregulation, and sensorimotor exploration together with other organs [24].

3) Muscles and movement

Surface ElectroMyoGraphy (SEMG) - is a non-invasive technique for measuring muscle electrical activity that occurs during muscle contraction and relaxation cycles. Surface electromyography is widely used in many applications, such as physical Rehabilitation, Urology (treatment of incontinence), biomechanics (sports training, motion analysis, research), ergonomics (studies in the workplace, job risk analysis, product design, and certification).SEMG is clinically indicated for biofeedback, relaxation, muscle re-education, treatment of incontinence. The SEMG signal generated by the muscle is captured by the electrodes, then amplified and collected by the sensor before being converted to a digital signal by the encoder. Note that fatigue is not always something that we want to prevent. For instance, in muscular training, short-term fatigue is a necessity for muscle growth and is actually looked for. In polysomnography, EMG is routinely measured from below the chin to detect the REM sleep phase (low tonic level) and sleep onset (continuous reduction of tonus). However, EMG can theoretically also be measured from other muscles, which is less intrusive for the subject and the changes in the signal pattern may be more pronounced at sleep onset at these locations. The problems of EMG are similar to those of EOG. Automatic analysis of EMG signals uses the frequency spectrum and the signal power features [25].

Respiration (Breathing rate) is the first step in the chain of events to transport oxygen to the cells of the body for metabolism to provide the body with energy. Respiration ventilates the lungs with air through inhalation and exhalation. The respiratory rate of a healthy adult at rest is usually between 12 and 20 breaths per minute [26].

Actigraphy is an objective measurement method that assesses limb movement activity and is a part of polysomnography. Recorded data are subjected to a proprietary algorithm that produces estimates of sleep-wake variables. Current standards of practice outline the primary roles of actigraphy in insomnia as the characterization of circadian or sleep patterns and the evaluation of treatment response [27]. Normally, accelerometer-based EE estimation is based on the magnitude of acceleration. It is based on a notion that the higher the resultant acceleration, the more intensive the physical activity. A comparative analysis of actigraphy devices capable of tracking temperature, sleep and activity rhythms with minor discrepancies [28]. Actigraphy is used in Fatigue Risk Management systems for Airline operations as part of the protocol, to gather data used in biomathematical models of fatigue.

4) Sensory(vision)

Use of non-invasive methods, such as making a video of the driver and alerting him/her on using cues that may help in anticipating the presence of a sleep pattern, can be a useful way to detect driver fatigue.

Percentage of Closure of eyes (PERCLOS) is a commonly used method for detection of driver fatigue. It determines the percentage of eye closure by taking the number of frames in which the driver's eyes are closed and dividing this by the total number of frames over a specified period of time.

Average closure of eyes (AVECLOS) is a simple binary measure indicating whether or not the driver's eyes are fully closed. This is a less complex measure of drowsiness than PERCLOS, and, as a result, it permits the use of an automotive-grade data processor as opposed to a high-grade PC processor required for PERCLOS. Validation testing at Delphi has shown a very close correlation between AVECLOS and PERCLOS (Pearson correlation coefficient = 0.95) [29].

5) Thermoregulatory

Thermoregulation is primarily achieved through physiological processes, as a function of the autonomic nervous system. The thermoregulatory function is characterized by body temperature (core, peripheral). Normal human core temperature varies between 36.5 ◦C and 37.5 ◦C. The temperature variation follows the circadian rhythm. The suprachiasmatic nucleus (SCN), which contains the central clock, controls the daily rhythms of sleep-wake behaviour and food intake via hypothalamic connections. The SCN controls the circadian rhythm in the secretion of hormones affecting glucose tolerance, including cortisol, melatonin and growth hormone. Body core temperature follows a sinusoidal circadian cycle and sleep onset is located where core temperature is decreasing and the wakeup is initiated when the core temperature is increasing [30]. Sleep is controlled by a thermo-regulatory process [31] showed that distal skin temperature increased at sleep onset and had a dramatic decrease at wake up (2° C).

B. Subjective parameters

Subjective perception of fatigue expressed as a self-rated effort with the help of questionnaires to evaluate somnolence, stress, cognitive and emotional components of fatigue physiological states.

Stanford sleepiness scale (SSS) - developed by Dement and colleagues in 1972, is a one-item self-report questionnaire measuring levels of sleepiness throughout the day. The scale, which can be administered in 1–2 minutes, is generally used to track overall alertness at each hour of the day. The scale has been validated for adult populations aged 18 and older. The SSS is used in both research and clinical settings to assess the level of intervention or effectiveness of a specific treatment in order to compare a client's progress [32].

Karolinska sleepiness scale (KSS) - has been widely used in studies of shift work, sleep deprivation, and driving. It has been found to correlate well with polysomnographic measurements (PSG), like alpha (8–12 Hz) and theta (4–8 Hz) activity in the EEG, as well as with performance-based measures, indicating that worsening of performance is associated with increased KSS values. A recent review summarizes a number of studies of KSS in different laboratory and field settings [33].

The Multidimensional Fatigue Symptom Inventory-Short Form (MFSI-SF) - a 30-item short form of the MFSI that yield scores only for the empirically derived subscales. Preliminary research suggests that it has acceptable psychometric properties and may be used as a substitute for the MFSI when time constraints and scale length are of concern [34].

Professional Quality of Life Scale (ProQOL) - a 30-question self-report test, that most commonly used to measure the negative and positive effects of helping others who experience suffering and trauma. The ProQOL has sub-scales for compassion satisfaction, burnout and compassion fatigue. This is similar to Burdon (post-traumatic) questionnaire and Stress-burnout adrenal fatigue questionnaire [35].

Sleep diary in aviation is usually concerned with the analysis of the work and rest patterns of aircrew, but they are also helpful in dealing with a suspected disorder of sleep - particularly with a disturbance of the circadian rhythm. Diaries should provide day-to-day details of duty, including time-zone changes, and daily estimates, through subjective, of the quality of sleep [36].

C. Objective parameters

Aspects which characterize workability are evaluated by using tests or activities which exercise physical, self-regulatory, cognitive or psychomotor capabilities. Test results serve as objectively measurable indices of human performance. The activities can be created as games, simulations or tests.

Cognitive tests are assessments of the cognitive capabilities of humans and other animals. Various cognitive capabilities, IQ, arithmetic, memory to measure human mental performance and also evaluate performance decrease caused by fatigue components. Long term mental arithmetic task has a significant effect on psychology, behaviour and physiology of subjects, which induces the changes of subjective sleepiness and mental fatigue, performance, autonomic nervous function and central nervous system [37].

The n-back task is a continuous performance task that is commonly used as an assessment in cognitive neuroscience to measure a part of working memory and working memory capacity also with an audio/visual (sensory) feedback [38].

Wechsler Adult Intelligence Scale (WAIS) is an IQ test designed to measure intelligence and cognitive ability in adults and older adolescents [39].

Psychomotor vigilance task (PVT) is a sustained-attention, the reaction-time task that measures the speed with which subjects respond to a visual stimulus. Research indicates increased sleep debt or sleeps deficit correlates with deteriorated alertness, slower problem-solving, declined psychomotor skills, and increased rate of false responding. The PVT lapse count was significantly associated with MFSI-sf physical fatigue ($r = 0.324$, $p = 0.025$). In hierarchical regression (full model $R2 = 0.256$, $p = 0.048$), higher BMI ($p = 0.038$), and higher MFSI-sf physical fatigue ($p = 0.040$) were independent predictors of the PVT lapse count [40].

IV. Conclusions

A conceptual model for the assessment of human mental fatigue as a multidimensional concept is described in this article. The parameters listed in this paper are required for a comprehensive evaluation of the functional state which indices human performance and workability and can be applied in areas with elevated mental effort and risks associated with low performance or loss of attention or productivity, road safety or critical operations. The observed fatigue assessment problem evolves in a complex environment where the measured parameters have multiple reflections on fatigue thus can be used within a variation method for evaluation of physiological states. Machine-augmented decision making is required in this area of biofeedback, thus it can benefit from a better definition of the measurement parameter effects when the sensor constraints allow modelling of physiological states based on the correlation. The ambiguity of parameter usage in decision making per-subject base can be improved by statistical analysis of characteristics. The reflection of mental fatigue on human performance is guided by the domain of application, it is required to assess the dynamic change of parameters while subject experiences a task-related activity modelled to specific task.

V. Acknowledgements

This research is supported by the EUREKA programme project „Non-intrusive human fatigue assessment" E! 11169 Fatigue.

References

[1] Phillips, Ross. (2014). What is fatigue and how does it affect the safety performance of human transport operators?.

[2] Hatice Ceylan Koydemir and Aydogan Ozcan, Wearable and Implantable Sensors for Biomedical Applications, Annual Review of Analytical Chemistry 2018 11:1, 127-146.

[3] Cercarelli, L. R., & Ryan, G. A. (1996). Long Distance Driving Behaviour of Western Australian Drivers. Proceedings of the Second International Conference on Fatigue and Transportation: Engineering, Enforcement and Education Solutions, Canning Bridge, Promaco, 35-45.

[4] Marcora, Samuele M.; Staiano, Walter; Manning, Victoria (January 2009). "Mental fatigue impairs physical performance in humans". Journal of Applied Physiology. 106 (3): 857–64. doi:10.1152/japplphysiol.91324.2008. PMID 19131473.

[5] de Vries, J. M., Hagemans, M. L., Bussmann, J. B., van der Ploeg, A. T., & van Doorn, P. A. (2009). Fatigue in neuromuscular disorders: focus on Guillain-Barré syndrome and Pompe disease. Cellular and molecular life sciences : CMLS, 67(5), 701-13. doi: 10.1007/s00018-009-0184-2.

[6] Bessonov V.A., Kipor G.V., Goncharov S.F., Ishkov A.V. Assessment of the degree of fatigability of drivers. Bills, A. (1934). General experimental psychology. New York: Longmans, Green & Co.

[7] Lim, Eric KS; Oster, Andrew JK; Rafferty, Andrew T (2014). Churchill's pocketbook of differential diagnosis (Fourth Edition ed.). Elsevier Health Sciences. ISBN 0702054046. Retrieved 1 March 2019.

[8] David Ponka, Michael Kirlew Can Fam Physician, Top 10 differential diagnoses in family medicine: Fatigue. 2007 May; 53(5): 892.PMCID: PMC1949177.

[9] Task force of the European society of cardiology and the North American society of pacing and electrophysiology, "Heart rate variability:standards of measurement, physiological interpretation, and clinical use." Circulation vol. 93, no.5, 1043-1065, 1996.

[10] Bonjyotsna, A., & Roy, S. (2014). Correlation of drowsiness with electrocardiogram: A review. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 3(5), 9538-9544.

[11] Thayer JF, Hansen AL, Saus-Rose E, Johnsen BH.Ann Behav Med. Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. 2009 Apr;37(2):141-53. doi: 10.1007/s12160-009-9101-z. Epub 2009 May 8.

[12] Appelhans, B., & Luecken, L. (2006). Heart rate variability as an index of regulated emotional responding. Review Of General Psychology, 10(3), 229-240. doi: 10.1037/1089-2680.10.3.229.

[13] Schmitt, L., Regnard, J., & Millet, G. P. (2015). Monitoring Fatigue Status with HRV Measures in Elite Athletes: An Avenue Beyond RMSSD?. Frontiers in physiology, 6, 343. doi:10.3389/fphys.2015.00343.

[14] Melo, H. M., Nascimento, L. M., & Takase, E. (2017). Mental fatigue and heart rate variability (HRV): The time-on-task effect. Psychology & Neuroscience, 10(4), 428-436. http://dx.doi.org/10.1037/pne0000110.

[15] Chua, E. C., Tan, W. Q., Yeo, S. C., Lau, P., Lee, I., Mien, I. H., Puvanendran, K., … Gooley, J. J. (2012). Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation. Sleep, 35(3), 325-34. doi:10.5665/sleep.1688.

[16] Meng, Junqing & Zhao, Bi & Ma, Yechao & Yiyu, Ji & Nie, Baisheng. (2014). Effects of fatigue on the physiological parameters of labor employees. Natural Hazards. 74. 10.1007/s11069-014-1235-z.

[17] Mizuhara, Wang, Kobayashi, & Yamaguchi (2004). A long-range

cortical network emerging with theta oscillation in a mental task. Neuroreport, 15(8), 1233.

[18] Niedermeyer & Lopes da Silva (2012, 6th edition). Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Philadelphia, PA: Lippincott Williams & Wilkins.

[19] SpydelJ.D., FordM.R., Sheer D.E. Task dependent cerebral lateralization of the 40 biz EEG rhythm. - Psychophysiology, 1979?16, pp. 347-350.

[20] Pulvermuller, F., Preissl, H., Lutzenberger, W., Birbaumer, N. Spectral responses in the gamma-band: physiological signs of higher cognitive processes? - NeuroReport, 1995, v. 6, pp. 2057-2064.

[21] Jap, B.T., Lal, S., Fischer, P., & Bekiaris, E. (2009) Using EEG spectral components to assess algorithms for detecting fatigue. Expert Systems with Applications, 36, 2352-2359.

[22] Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. Brain Research Reviews, 29(2-3), 169-195. http://dx.doi.org/10.1016/S0165-0173(98)00056-3.

[23] Critchley, H. D. (2002). Electrodermal responses: what happens in the brain. The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry, 8(2), 132–142.

[24] Karlen, Walter. (2009). Adaptive wake and sleep detection for wearable systems. 10.5075/epfl-thesis-4391.

[25] Charbek, E. (2015 08). Medscape Normal Vital Vigns. [Online] Available at: http:// emedicine.medscape.com/article/2172054-overview, [Accessed 8 June 2017].

[26] Katie L.StoneSoniaAncoli-Israel, Principles and Practice of Sleep Medicine (Fifth Edition) 2011, Pages 1668-1675 https://doi.org/10.1016/B978-1-4160-6645-3.00147-X.

[27] Fibion Inc., An Energy Expenditure Estimation Method based on Tri-Axial Accelerometry and Advanced Activity Type Classification, 2018, whitepaper, rerieved on March 3, 2019

[28] Saima Naz, Sheikh Ziauddin*, Ahmad R. Shahid, "Driver Fatigue Detection using Mean Intensity, SVM,and SIFT", 2017

[29] Czeisler, C. A., Zimmerman, J. C., Ronda, J. M., Moore-Ede, M. C., and Weitzman, E. D. (1980). Timing of rem sleep is coupled to the circadian rhythm of body temperature in man. Sleep, 2(3):329–46.

[30] Kräuchi, K., Cajochen, C., and Wirz-Justice, A. (2004). Waking up properly: is there a role of thermoregulation in sleep inertia? Journal of Sleep Research, 13(2):121–127.

[31] Shahid, Azmeh; Wilkinson, Kate; Marcu, Shai; Shapiro, Colin M. (2011-01-01). Shahid, Azmeh; Wilkinson, Kate; Marcu, Shai; Shapiro, Colin M., eds. STOP, THAT and One Hundred Other Sleep Scales. Springer New York. pp. 369–370. doi:10.1007/978-1-4419-9893-4_91. ISBN 9781441998927.

[32] Miley, A. Å., Kecklund, G., & Åkerstedt, T. (2016). Comparing two versions of the Karolinska Sleepiness Scale (KSS). Sleep and biological rhythms, 14(3), 257-260.

[33] Stein, K. D., Jacobsen, P. B., Blanchard, C. M., Thors, C. T. (2004). Further validation of the Multidimensional Fatigue Symptom Inventory-Short Form (MFSI-SF). Journal of Pain and Symptom Management, 27, 14-23.

[34] B. Hudnall Stamm, 2009. Professional Quality of Life: Compassion Satisfaction and Fatigue Version 5 (ProQOL).

[35] Anthony N. Nicholson The Neurosciences and the Practice of Aviation Medicine, page 221, CRC Press, 2011.

[36] Zhang, Chong & Yu, Xiaolin. (2010). Estimating mental fatigue Based on electroencephalogram and heart rate variability. Pol J Med Phys Eng PL ISSN. 16. 67-84. 10.2478/v10013-010-0007-7.

[37] Kirchner, W. K. (1958). "Age differences in short-term retention of rapidly changing information". Journal of Experimental Psychology. 55 (4): 352–358. doi:10.1037/h0043688.

[38] Kaufman, Alan S.; Lichtenberger, Elizabeth (2006). Assessing Adolescent and Adult Intelligence (3rd ed.). Hoboken (NJ): Wiley. p. 3. ISBN 978-0-471-73553-3. Lay summary (22 August 2010).

[39] Lee IS; Bardwell WA; Ancoli-Israel S; Dimsdale JE. Number of lapses during the psychomotor vigilance task as an objective measure of fatigue. *J Clin Sleep Med* 2010;6(2): 163-168.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 38-43

Analysis of Research Trends in Agricultural Engineering

Sławomir Francik
Department of Mechanical Engineering and Agrophysics
Faculty o Production and Power Engineering,
University of Agriculture
Krakow, Poland
slawomir.francik@urk.edu.pl

Adrian Knapczyk
Department of Mechanical Engineering and Agrophysics
Faculty o Production and Power Engineering,
University of Agriculture
Krakow, Poland
adrian.knapczyk@urk.edu.pl

Jarosław Frączek
Department of Mechanical Engineering and Agrophysics
Faculty o Production and Power Engineering,
University of Agriculture
Krakow, Poland
jaroslaw.fraczek@urk.edu.pl

Zbigniew Ślipek
Department of Mechanical Engineering and Agrophysics
Faculty o Production and Power Engineering,
University of Agriculture
Krakow, Poland
zbigniew.slipek@urk.edu.pl

***Abstract*— Nowadays it is necessary to define actual global research trends in disciplines in order to determine the major research topics, discoveries and global scientific networks. The goal of the work was to determine the main topics and research trends in agricultural engineering. For this caser he method of modified bibliometric analysis was applied. Such analysis of the keywords frequency occurrence and the links between the keywords in Biosystems Engineering Journal shows that in the three analysed periods, certain keywords invariably remain dominant. These are "model", "system" or "systems", "temperature" and "water". The highest number of publications was qualified for the CIGR section: Information Technology. It can be, therefore, stated that in the recent period the most intensively developing researches were dedicated to the application of IT tools in creation various types of models, allowing to simulate and optimize agricultural processes. A majority of publications covered decision-making support (Decision Support Systems), image analysis application, and automation of agricultural processes.**

Keywords*— agricultural engineering, bibliometric analysis, biosystems engineering, research trends

I. Introduction

It is important for researchers to know the current research trends in a discipline (such as agricultural engineering) in order to decide in which direction to plan their future research. Regular analysis of research trends related to specific disciplines can help researchers publish more effectively. It is necessary to define a new global research trends that have developed in the discipline in order to determine the major research topics, discover and global knowledge networks. This necessity is emphasized by many authors of scientific publications. Bibliometric studies are widely applied in the analysis of research trends [1]–[7].

Bibliometry was introduced by Pritchard in 1969 as a method of mathematical and statistical analysis applied to books and other communication media. Citation and content analyses are now widely used bibliometric techniques. Broadus defined bibliometry as a quantitative analysis of physically published units or bibliographic units. Bibliometric techniques have many advantages. The most important is the ability to conduct quantitative analyses objectified based on the codified knowledge – measurable, objectified, consistent and accessible data. Hence, bibliometrics is an effective and important tool to determine the trends of research in various fields of science [8]. Bibliometric studies include a number of quantitative and visual procedures to generalize patterns and dynamics of publications [7].

The goal of the work was to determine the main topics and research trends in agricultural engineering. The method of bibliometric analysis was applied.

In line with the CIGR (International Commission of Agricultural and Biosystems Engineering) definition, agricultural engineering "... has been applying scientific principles for the optimal conversion of natural resources into agricultural land, machinery, structure, processes and systems for the benefit of man.

Beside the classical agricultural engineering topics, a comparatively new research area of biosystems engineering was analysed. „Biosystems Engineering" is research in the physical sciences and engineering to understand, model, process or enhance biological systems for sustainable developments in agriculture, food, land use and the environment" [9].

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4169
© 2019 Sławomir Francik, Adrian Knapczyk, Jarosław Frączek, Zbigniew Ślipek.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

II. Materials and Methods

A modified bibliometric analysis method proposed by Rizzi and Leydesdorff [10] was applied in the work. The following stages of analysis were adopted:

1. Identifying journals in the Web of Science "Agricultural Engineering" category from the Journal Citation Reports database.

2. Determining the journals which had the most significant impact on research development in the discipline of agricultural engineering.

3. Selecting one journal for further analysis and determining the length of analysis periods.

4. Uploading all publications of the selected journal in the analysed period of time and extracting bibliometric data (authors, title, year of issue, key words, additional key words, publishing house).

5. Construction and analysis of term maps (VOSviewer software).

6. Identifying the most frequently found key words for the selected journal in the analysed period of time.

7. An topic analysis of the publications which were identified with the use of the most frequently occurring key words (topic: model & system).

One of the most widely used databases for bibliometric analysis of publications in the Web of Science (WoS), which contains articles with the highest level of quality – from magazines with a significant impact factor. Therefore, in this study the WoS database has been used as a data source [8].

All publications in the WoS: „AGRICULTURAL ENGINEERING" category were identified with the use of the tools of InCites Journal Citation Reports. The years 2000-2016 were analysed.

From among the identified journals, the journals which were indexed in the Web of Science – Core Collection (WoS-CC) database for minimum 10 years were selected for further analysis. The analysis of the change of the number of publications in a given period of time was performed on this sample.

It was assumed that to have a significant impact on the development of a given field, journals needed to be indexed in WoS-CC during the whole period or in the recent years (minimum since 2006).

Selection of one journal for further analysis was performer with the use of journal ranking in 2016 in the WoS "Agricultural Engineering" category and the topic range closely connected with the category (the journal linked to maximum two WoS categories).

Bibliometric analyses were performer with the use of the freeware, VOSviewer. This program is used for creating and graphic visualisation of bibliometric maps. VOSviewer utilized the method of „visualization of similarities" VOS. The program allows to create maps of authors, journals etc.[11].

In this work the program was used to create maps of terms in the version of „thermic maps".

Such maps allow for analysing concentrations in a simple way.

The key words with the highest occurrance frequency are displayed in a circles a bigger size.

The words with the lower frequency of occurance, on the other hand, have smaller font and they are presented in a smaller circles.

VOSviewer is widely used for bibliometric analyses, mainly to analyse research trends [12]–[15], but also to visualize certain fields of knowledge [16].

Identifying the most frequent key words for a given journal was also performed with the use of VOSviewer program – 14 most frequently occurring key words in each of the three analysis periods were determined. Out of the lot, only the key words occurring in each of the three periods were used for further analysis.

The next stage covered the quantitative analysis (for the 3 periods) and topics analysis (the most recent period) of publications which had been identified in the Web of Science – Core Collection database with the use of the most frequently found key words. The search was limited to the selected journal (BIOSYSTEMS ENGINEERING) and to the publications which contained two of the most frequently used key words.

The analysis of topics was performer by assigning individual publications on agricultural engineering research topics (agricultural engineering) to one of 7 CIGR Technical Sections:

Section I: Land and Water,

Section II: Structures & Environment,

Section. III: Plant Production,

Section. IV: Energy in Agriculture,

Section. V: System Management,

Section. VI: Bioprocesses,

Section VII: Information Technology

Assigning followed the analysis of the publication contents based on abstracts.

III. RESULTS

A total of 18 journals can be found in the WoS 'AGRICULTURAL ENGINEERING' category in the years 2000-2016. The journals selected for further analysis were those which were indexed in the WoS – Core Collection database for minimum 10 years. For thus selected 7 journals, the number of publications in the analysed years was determined. The analyzes show that 3 journals (Bioresource Technology, Biomass & Bioenergy, and Industrial Crops And Products) show a significant increase in the number of publications. The other journals in the analysed period of time showed an even distribution of the number of publications. In total, the presented journals

published 31402 publications in the researched period.

The journals with the highest rank (the highest citation index / impact factor) had the most significant impact on the research development in the field of Agricultural Engineering. Those were the journals which in the year 2016 belonged to the Q1 quartile (according to JCR): Bioresource Technology (position 1/14), Biomass & Bioenergy (position 2/14), Industrial Crops and Products (position 3/14) and to the Q2 quartile: Biosystems Engineering (position 4/14).

The Biosystems Engineering journal (position no. 4 in the ranking) was selected for further analyses. The journal was categorized as WoS 'AGRICULTURAL ENGINEERING' and 'AGRICULTURE, MULTIDISCIPLINARY'. The journals with the positions 1-3 in the JCR ranking are classified in a higher number of categories. Their range of topics is, thus, very wide. Publications in the Biosystems Engineering journal are the most homogenous with regard to the presented topics.

According to the information on the journal web site," Biosystems Engineering publishes research in engineering and the physical sciences that represent advances in understanding or modelling of the performance of biological systems for sustainable developments in land use and the environment, agriculture and amenity, bioproduction processes and the food chain. The subject matter of the journal reflects the wide range and interdisciplinary nature of research in engineering for biological systems."

The Biosystems Engineering journal has been indexed in the WoS database since 2002 (Formerly known as Journal of Agricultural Engineering Research).

For this reason, it was decided to split three 5-year research periods:

- period I 2002-2006,
- period II 2007-2011,
- and period III 2012-2016.

In the analysed periods of time the Biosystems Engineering journal published a total of 2354 articles which were distributed over the time periods as follows: period I – 751, period II – 807, period III – 796. The bibliometric data was extracted for all the publications (Title, Abstract, Author Key Words, Key Words Plus, year of publishing).

In the next stage, the most frequently occurring key words were determined for the analysed periods. For each period the analysis of all key words (Author Keywords, KeyWords Plus) was performed. The analysis was limited to the key words which occurred minimum 5 times. Results of the simulations are presented in Fig. 1.

In Fig. 1 maps of terms (key words) for the three analysed periods are presented.

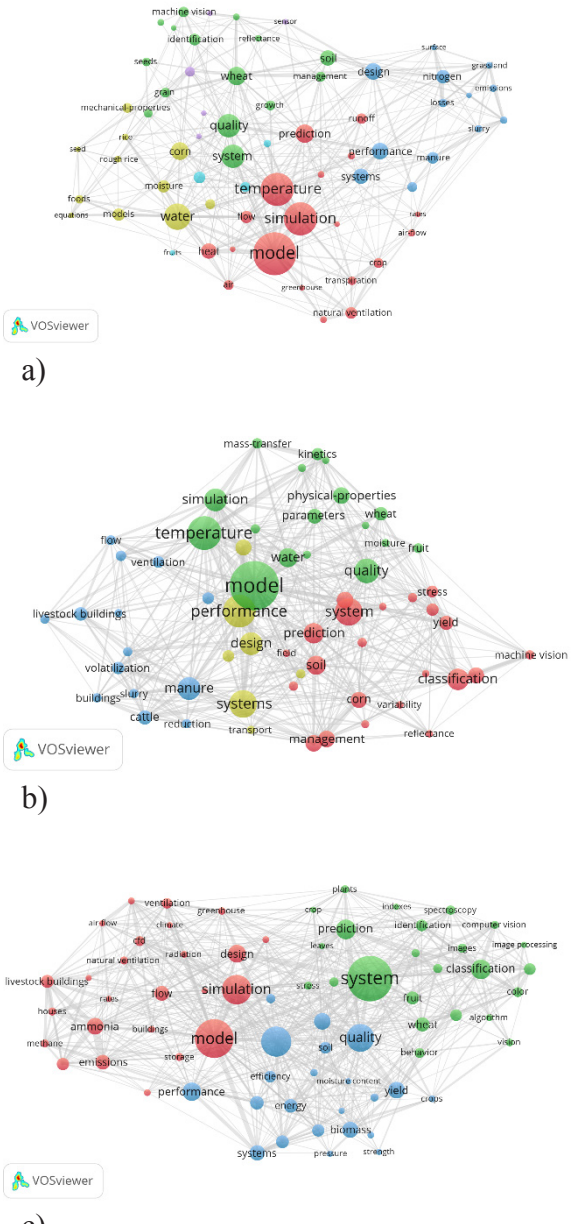


Fig. 1. Set of term maps for the analysed periods: a) period I b) period II c) period III.

In selected periods can be observed several groups around the most frequently used keywords, which is the evidence of the subject-matter linkages between the key words. The main topics in the first period were: "model", "simulation", "temperature", "system", "quality", "prediction", "wheat", "water" and "design". The main topics in the second period were: "model", "temperature", "simulation", "performance", "system", "quality", "prediction", "classification", "management", "manure", "physical properties" and "soil". In the last period the main topics were: "model", "system", "simulation", "design", classification", "prediction", "quality", "ammonia", "emission", "livestock buildings" and "biomass". Change in the interest of researchers from the study of physical properties, prediction, simulation, quality of biological raw materials, classifications, fertilizers, soil and ending with the emission of greenhouse gases.

Then, the most frequently used key words in the analysed periods were listed (Table 1). It can be observed that key words are repeated in the analysed periods but there are changes in their rank (the frequency of use changes). The key words which occur in all the three periods were selected for further analyses. They are shaded in the table of results.

TABLE I. List of the most frequently used key words in the analysed periods

Period 1	*Period 2*	*Period 3*
MODEL	MODEL	SYSTEM
TEMPERATURE	TEMPERATURE	MODEL
SIMULATION	PERFORMANCE	TEMPERATURE
WATER	SYSTEMS	SIMULATION
QUALITY	SYSTEM	QUALITY
SYSTEM	QUALITY	PREDICTION
PREDICTION	MANURE	CLASSIFICATION
WHEAT	SIMULATION	PERFORMANCE
CORN	DESIGN	DESIGN
PERFORMANCE	CLASSIFICATION	AMMONIA
DESIGN	PREDICTION	YIELD
HEAT	SOIL	GROWTH
SYSTEMS	WATER	BIOMASS
SOIL	NITROGEN	SYSTEMS

In the next stage, a series of searches was performed in the Web of Science – Core Collection database with regard to the selected key words and separately for all the 3 periods. The search was limited to the category of WoS AGRICULTURAL ENGINEERING and to the selected journal (BIOSYSTEMS ENGINEERING).

TABLE II. Listing of the number of identified journals with regard to the number of searched words in the period I/ II / III

searche words	*1. +TEM*	*1.2. +SIM*	*1.3. +QUA*	*1.4. +PER*	*1.5. +DES*	*1.6. +PRE*
1. TEM	39/ 36/ 37	15/ 15/ 15	3/ 7/ 11	6/ 16/ 10	8/ 14/ 14	10/ 6/ 10
2. SIM		47/ 42/ 54	4/ 5/ 11	12/ 15/ 10	13/ 14/ 20	14/ 8/ 10
3. QUA			10/ 20/ 42	2/ 3/ 10	0/ 5/ 18	4/ 7/ 10
4. PER				29/ 43/ 52	11/ 16/ 19	5/ 10/ 10
5. DES					38/ 48/ 57	0/ 11/ 9
6. PRE						26/ 34/ 36

Legend: TEM - TEMPERATURE, SIM – SIMULATION, QUA - QUALITY, PER - PERFORMANCE, DES - DESIGN, PRE - PREDICTION

The numbers of publications meeting the search criteria (different configurations of key words) are presented in Table 2. The criteria always included the words: "model" and "system". The highest number of publications (143 articles) was obtained for the two key word configurations: "model" AND "system" AND "simulation" (143=47+42+54), „model" AND „system" AND „design" (143=38+48+57).

Result of the topic analysis of publications in the third period (2012-2016) :

CIGR Technical Section I – Land & Water :

 # irrigation systems:

- technical solutions in irrigation systems

- modeling and optimization of irrigation

- other subjects: assessing water saving at farm, modeling the impact of climate change on irrigation systems, modeling the embodied energy for irrigation systems

 # study of soil physical and chemical properties

CIGR Technical Section II – Structures & Environment :

 # modeling and optimization ventilation system

- microclimate in livestock buildings

- designing forced air circulation system for plant factories

 # ammonia, greenhouse gas and odour emissions from livestock buildings and greenhouses:

- modeling and optimization

- emission measurement (pig production, cattle farms, greenhouses)

 # other topics: radiometric properties of materials for livestock buildings

CIGR Technical Section III – Plant Production:

 # agricultural machines investigation:

- measurement of processes parameters (disc harrow, tillage machines, combine harvester, seeder, orchard sprayers, tractor rollover)

- technical solutions in plant production systems (tillage machines, machine prototype for non-chemical weed control, semiochemical devices for pest management)

 # modeling and optimization operations in plant production:

- crop harvesting and transport

- horticulture (rose cultivation, grape harvest, sweet cherry picking

 # other topics: water requirements for plants, detection of plant nutrition level, weed control, detection of virus-infected seeds, canopy cover measurement in an onion crop, - modeling of yield maps

CIGR Technical Section IV – Energy in Agriculture:

 # biomass

- measurement of biomass parameters: moisture

content, diameters of stems

- optimisation and investment analysis: comparing of two different biomass energy exploitation systems (domestic pellet boilers, centralised co-generation unit)

 # energy savings

CIGR Technical Section V – System Management:

 # food and grain supply chain management

 # costs analysis for machine systems and technical infrastructure

 # agricultural operations planning

CIGR Technical Section VI – Bioprocesses:

 # aeration, drying and storage

- postharvest quality of fresh fruit and vegetables

- modeling, simulation and optimization of air distribution in grain storage bins

- modeling and simulation of seed drying process,

- heat transport models

 # disinfesting postharvest fresh and stored products

 # other topics: evapotranspiration estimating techniques (agricultural crops, banana plantation, processing tomato crop), chemical oxygen demand

CIGR Technical Section VII – Information Technology:

 # image analysis

- animal production (pigs, cows, chickens),

- plant production (apples, oil palm, vegetables, citrus)

 # 3D models creating

- trees and orchard modeling based on LiDAR system

 # Decision Support Systems:

- web-based DSSs: pest management, crop production, biomass production

- DSSs for: operational planning machine activities, pig production chain

 # Prediction and forecasting system

- airborne spread of livestock infectious disease

- tools for prediction of: size and weight pigs in a herd, yield predictions in agriculture

 # autonomous tractors and machines

- automatic control and navigation

- stability evaluation of robotic agricultural on slopes

- path planning of robot for collection of eggs

 # devices for modeling and simulation

- activity of honeybees

- modeling radio propagation within orchards

- monitoring grain losses in combine harvesters

- automated monitoring tool for pest

- wireless sensor network for monitoring of temperature and oxygen concentration inside silage stacks

The last step was the topic analysis of publications in the third period (2012-2016). The review was performed for the documents which contained the words 'model*' and 'system*'. Out of 178 articles which were generated from the WoS-CC database, 35 were rejected as they did not contain the searched phrases in the title, author key words and additional key words.

In addition, 11 publications whose subject matter did not fit into the discipline Agricultural Engineering were removed.

Finally, 132 documents were analysed with regard to content. The results of topic analysis are presented in the Table 3 and Table 4. The main research topics performed in the framework of each of the research areas (CIGR Technical Sections) were identified. Table 3 and Table 4 contain also references to the publications which pertained to a given topic.

Most of the publications (32) were related to the applications of information technology in agricultural engineering (CIGR Section: Information Technology). A smaller number of publications were related to four CIGR sections: Plant Production (27), Bioprocesses (21), Land & Water (19) and Structures & Environment (19). The smallest number of publications was qualified in the CIGR sections: System Management (10) and Energy in Agriculture (4).

The publications assigned to the Information Technology section discussed mainly: use of image analysis in agriculture (10), creating Decision Support Systems for agriculture (5) and introducing automatically steered machines and equipment (5). In the Plant Production section the highest number of publications discussed agricultural machines (12). In the Bioprocesses section 15 publications were dedicated to the research of aeration, drying and storage. In the Land & Water section the highest number of publications discussed irrigation systems (16). In the Structures & Environment section 13 publications were dedicated to the research of microclimate in inventory buildings. In the other sections there were no dominant scientific research topics.

The obtained results coincide with the activities of the Faculty of Production and Power Engineering of the University of Agriculture in Krakow, whose scientific activity is related to agricultural engineering. For example, in the range: optimization operations in plant production [17], [18] research on physical and chemical properties as well as energetic plant materials and agglomerations [19]–[23], waste treatment and testing of their properties [24]–[27], decision support systems [28], [29] and optimization of the drying process [30] and other.

IV. CONCLUSIONS

Summing up the results, it can be concluded that:

The results of bibliometric analyses show that the strongest influence on the research development in Agricultural Engineering was exercised by the journals with the highest JCR rank: Bioresource Technology, Biomass & Bioenergy, Industrial Crops and Products, and Biosystems Engineering. It should be noted that the subject matter in the Biosystems Engineering journal was the one most closely connected with Agricultural Engineering.

Bibliometric analysis of the key word frequency of occurrence and the linkages between the key words in Biosystems Engineering journal (VOS Viewer program) showed that in the three analysed periods certain key words invariably remain dominant (they are placed in large circles). These are 'model', 'system' or 'systems', 'temperature' and 'water'. In different periods new most frequently occurring key words appear. Linkages between key words change, too. The analysis of the three researched periods showed that the most frequent key words have been: "model", "system", "temperature", "simulation", "quality", "performance", "design" and "prediction". It can be, therefore, stated the biggest number of scientific publications was connected with modelling of agrotechnical (biotechnological) systems and simulating their activities.

A detailed publication topic analysis of the period 2012-2016 with regard to the most frequently used topic key words (143 articles) allowed to identify the dominant research areas – research trends in agricultural engineering.

The biggest number of publications was qualified for the CIGR section: Information Technology. It can be, therefore, stated that in the recent period the most intensively developing research was dedicated to the application of IT tools for creating various types of models, allowing to simulate and optimize agricultural processes. A large number of publications covered decision-making support (Decision Support Systems), picture analysis application, and automation and robotization of agricultural processes. It can be assumed that those trends will continue in the nearest future and that the research dedicated to the application of Information Technology and mechatronics in agriculture will remain dominant.

ACKNOWLEDGEMENT

This research was financed by the Ministry of Science and Higher Education of the Republic of Poland (statutory activities DS-3600/WIPiE/2018, Faculty of Production and Power Engineering, University of Agriculture in Krakow).

REFERENCES

[1] S. Bojović, R. Matić, Z. Popović, M. Smiljanić, M. Stefanović, and V. Vidaković, "An overview of forestry journals in the period 2006–2010 as basis for ascertaining research trends," Scientometrics, vol. 98, no. 2, pp. 1331–1346, Feb. 2014.

[2] J. B. Hernández and S. Chalela, "Research Trends in the Study of ICT Based Learning Communities: A Bibliometric Analysis," EURASIA J. Math. Sci. Technol. Educ., vol. 13, no. 5, pp. 1539–1562, Dec. 2016.

[3] A. Magrí, F. Giovannini, R. Connan, G. Bridoux, and F. Béline, "Nutrient management from biogas digester effluents: a bibliometric-based analysis of publications and patents," Int. J. Environ. Sci. Technol., vol. 14, no. 8, pp. 1739–1756, Aug. 2017.

[4] E. W. Wambu, H.-Z. Fu, and Y.-S. Ho, "Characteristics and trends in global tea research: a Science Citation Index Expanded-based analysis," Int. J. Food Sci. Technol., vol. 52, no. 3, pp. 644–651, Mar. 2017.

[5] J. Wang, Z. Chen, L. Yang, and S. Xi, "Study on trends and performance of landfill research from 1999 to 2013 by using bibliometric analysis," Environ. Prog. Sustain. Energy, vol. 34, no. 5, pp. 1349–1355, Aug. 2015.

[6] X. Wu, X. Chen, F. B. Zhan, and S. Hong, "Global research trends in landslides during 1991–2014: a bibliometric analysis," Landslides, vol. 12, no. 6, pp. 1215–1226, Dec. 2015.

[7] X. Yaoyang and W. J. Boeing, "Mapping biofuel field: A bibliometric evaluation of research output," Renew. Sustain. Energy Rev., vol. 28, pp. 82–91, Dec. 2013.

[8] S. Francik et al., "Bibliometric Analysis of Multiple Critteria Decision Making in Agriculture," Tech. Sci., vol. 20, no. 1, pp. 17–30, 2017.

[9] D. Briassoulis, P. Panagakis, and E. Nikopoulos, Education and research in Biosystems engineering in Europe. ERABEE, 2009.

[10] A. Knapczyk, S. Francik, N. Pedryc, and T. Hebda, "Bibliometric analysis of research trends in engineering for rural development," in 17th International Scientific Conference Engineering for Rural Development, 2018, pp. 700–707.

[11] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," Scientometrics, vol. 84, no. 2, pp. 523–538, Aug. 2010.

[12] F. T. Gizzi, "Worldwide trends in research on the San Andreas Fault System," Arab. J. Geosci., vol. 8, no. 12, pp. 10893–10909, Dec. 2015.

[13] O. Tabatabaei-Malazy, R. Atlasi, B. Larijani, and M. Abdollahi, "Trends in publication on evidence-based antioxidative herbal medicines in management of diabetic nephropathy," J. Diabetes Metab. Disord., vol. 15, no. 1, p. 1, Dec. 2015.

[14] S. H. Zyoud, W. S. Waring, W. M. Sweileh, and S. W. Al-Jabi, "Global Research Trends in Lithium Toxicity from 1913 to 2015: A Bibliometric Analysis," Basic Clin. Pharmacol. Toxicol., vol. 121, no. 1, pp. 67–73, Jul. 2017.

[15] A. Watad et al., "Is autoimmunology a discipline of its own? A big data-based bibliometric and scientometric analyses," Autoimmunity, vol. 50, no. 4, pp. 269–274, May 2017.

[16] J. Zhu and W. Hua, "Visualizing the knowledge domain of sustainable development research between 1987 and 2015: a bibliometric analysis," Scientometrics, vol. 110, no. 2, pp. 893–914, Feb. 2017.

[17] T. Głąb and B. Kulig, "Effect of mulch and tillage system on soil porosity under wheat (Triticum aestivum)," Soil Tillage Res., vol. 99, no. 2, pp. 169–178, Jun. 2008.

[18] T. Głąb, "Impact of soil compaction on root development and yield of meadow-grass," Int. Agrophysics, vol. 27, no. 1, pp. 7–13, Jan. 2013.

[19] K. Kubica, M. Jewiarz, R. Kubica, and A. Szlęk, "Straw Combustion: Pilot and Laboratory Studies on a Straw-Fired

Grate Boiler," Energy & Fuels, vol. 30, no. 6, pp. 4405–4410, Jun. 2016.

[20] M. Niemczyk, A. Kaliszewski, M. Jewiarz, M. Wróbel, and K. Mudryk, "Productivity and biomass characteristics of selected poplar (Populus spp.) cultivars under the climatic conditions of northern Poland," Biomass and Bioenergy, vol. 111, pp. 46–51, Apr. 2018.

[21] M. Wróbel, J. Hamerska, M. Jewiarz, K. Mudryk, and M. Niemczyk, "Influence of Parameters of the Torrefaction Process on the Selected Parameters of Torrefied Woody Biomass," Springer, Cham, 2018, pp. 691–700.

[22] M. Wróbel, K. Mudryk, M. Jewiarz, S. Głowacki, and W. Tulej, "Characterization of Selected Plant Species in Terms of Energetic Use," Springer, Cham, 2018, pp. 671–681.

[23] M. Mierzwa-Hersztek, K. Gondek, M. Jewiarz, and K. Dziedzic, "Assessment of energy parameters of biomass and biochars, leachability of heavy metals and phytotoxicity of their ashes," J. Mater. Cycles Waste Manag., Feb. 2019.

[24] B. Brzychczyk, T. Hebda, and J. Giełżecki, "Energy Characteristics of Compacted Biofuel with Stabilized Fraction of Municipal Waste," Springer, Cham, 2018, pp. 451–462.

[25] B. Brzychczyk, T. Hebda, and J. Giełżecki, "Physical and Chemical Properties of Pellets Produced from the Stabilized Fraction of Municipal Sewage Sludge," Springer, Cham, 2018, pp. 613–622.

[26] K. Grzesik and M. Malinowski, "Life cycle assessment of refuse-derived fuel production from mixed municipal waste," Energy Sources, Part A Recover. Util. Environ. Eff., vol. 38, no. 21, pp. 3150–3157, Nov. 2016.

[27] K. Dziedzic, B. Łapczyńska-Kordon, M. Malinowski, M. Niemiec, and J. Sikora, "Impact Of Aerobic Biostabilisation And Biodrying Process Of Municipal Solid Waste On Minimisation Of Waste Deposited In Landfills," Chem. Process Eng., vol. 36, no. 4, pp. 381–394, Dec. 2015.

[28] S. Francik, N. Pedryc, A. Knapczyk, A. Wójcik, R. Francik, and B. Łapczyńska-Kordon, "Bibliometric Analysis of Multiple Critteria Decision Making in Agriculture," Tech. Sci., vol. 20, no. 1, pp. 17–30, 2017.

[29] M. Cupiał, A. Szeląg-Sikora, and M. Niemiec, "Optimisation of the Machinery Park with the Use of OTR-7 Software in Context of Sustainable Agriculture," Agric. Agric. Sci. Procedia, vol. 7, pp. 64–69, 2015.

[30] T. Lapczynska-Kordon, B; Fraczek, J; Slipek, Z; Ivanowa, "Genetic Algorithms In Optimisation Of Dried Fruits And Vegetables Quality," in Engineering for Rural Development, 2013, pp. 628–632

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 51-55

Data Science Approach for IT Project Management

Janis Grabis
Department of Management Information
Riga Technical University
Riga, Latvia
grabis@rtu.lv

Bohdan Haidabrus
Computer Science Department
Sumy State University
Sumy, Ukraine
haidabrus@gmail.com

Serhiy Protsenko
Department of Electronics, General and Applied Physics
Sumy State University
Sumy, Ukraine
serhiy.protsenko@gmail.com

Iryna Protsenko
Computer Science Department
Sumy State University
Sumy, Ukraine
protsenko3107@gmail.com

Anna Rovna
Computer Science Department
Sumy State University
Sumy, Ukraine
a.rovna@mk.sumdu.edu.ua

***Abstract*—Majority of the IT companies realized that ability to analyse and use data, could be one of the key factors for increasing of number of successful projects, portfolios, programs. Key performance indicators based on data analysis helps organizations be more prosperous in a long term perspective. Also, statistical data are very useful for monitoring and evaluation of project results which are very important for managers, delivery directors, CTO and others high level management of company. The Data Science methods could make more efficient project management in several of business problems.**

Analysis of historical data from the project life-cycle based on Data Science models could provide more efficient benefits for different stakeholders. Differential of the project data vector with target as an integral evaluation of the project success which allow for the complex correlations between separate features. Therefore, the influence of features importance and override creatures could be decreased on the target.

This study propose new approach based on Data Science providing more efficient and accurately project management, taking into account best practices and project performance data.

Keywords*— Machine Learning, Data Analysis, Project Management, Business Processes.

I. INTRODUCTION

Nowadays, the ability to analyse and use data is one of the key factors affecting the organization's ability of the IT companies to work effectively in the long perspective. The main factor affected the successful implementation of the projects could be described with so-cold project management triangle. The triangle consists of following components: time, budget and quality (or scope) [14, 19]. The are several ways to calculate indicators related to the deadlines:

- assessment of the task in hours of the developer who will be engaged in it (previously agreed with the developer himself);
- have the data on how many hours were actually spent on this task (this requires a time tracker);
- due date due to which the task should be ready.

From the other hand, it is possible to convert part of the time metrics to the project budget. The evaluation of the project time scale could be described as - *Start date*, *Due Date*, *Actual Date*. Specifically, the time metric can be measured as the number of deadlines per task, or the ratio of differences between *Start date* / *Due Date* and *Due Date / Actual Date*. The budget indicator is based on a preliminary assessment of the time and the actual time spent to the task [10].

Regarding the project budget, with respect of the price determination our research consists of three project options:

- *The fixed cost project* does not imply a deviation from the budget characteristics;
- *The time and Material project* means that the cost is linearly related to the hours spent, which means that deviations from the budget are identical with deviations from the deadlines. Thus, in these cases, there are virtually no budget metrics;
- *A Flexible fixed cost project*, allowing coordinate the budget changes, it is potentially not clear which part of the budget change should be attributed to the "merit" of a particular developer.

In general, it means that in the project management triangle for an individual developer, budget figures cannot be adequately representing, thus it is necessary to focus on the metrics of project quality and time.

Project success depends on a large number of factors. They for it is necessary to build and analyze the system of key performance indicators at all stages and phases of the project life-cycle. A project team, project

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4163
© 2019 Janis Grabis, Bohdan Haidabrus, Serhiy Protsenko, Iryna Protsenko, Anna Rovna.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

manager, delivery manager, company's CEO and CTO required to have information on the current status of the project. Important to provide all key information about project management and realization: project status, time delay, according to budget, risk monitoring and control, requirements status from the project owner, feedback from the stakeholders and the need to generalize all parameters of the specific project.

The aim of our research is to develop and implement an approach for predicting the evaluation of the project efficiency which based on an integrated estimation of all parameters during the project life-cycle. Analysis of historical data with Machine Learning (ML) models [7, 9] with aim to do initiation and planning phases and better coordination with the different stakeholders.

II. MATERIALS AND METHODS

Using Data Science Approach for IT Project Management can be divided into two parts:

- *The Best Practices Approach:* historical integrated analysis of the best practices and fails on the projects;
- *The Historical Data Approach*: analysis of historical data for project initiation and planning depend on the customer requirements and actual status of the project by using data.

According to *The Best Practices Approach* we have analysed *The Sales, Project, Production Methodology* of the IT Companies MindK (mindk.com) [22] and Already On (alreadyon.com) [23]. The documentation is developed as a support document for MindK and Already On to secure a common framework on how we indemnify an optimal project and development process.

This methodology consists of the three general part: Sales/Initiation Process, Project Process and Production Process. The Sales process define the following steps: Identification, Quality, Offering, Negotiation, Pre Project, Signing, Specification phase. The Sales process always ends with a signed specification and acceptance criteria. The data which we can use to analyze from the Sales/ Initiation Process are at the RASCI-roles matrix (Table 1) (variation of RACI):

- R - Responsible – Persons involved to achieve a task.
- A - Accountable – Persons ultimately accountable for the correct and thorough completion of the deliverable or task, and the one to whom Responsible is accountable. If there is no Responsible on the project, Accountable does the work.
- S - Support – Resources dedicated to Responsible. Unlike Consulted, who may provide input to the task, Support will assist in completing the task.
- C - Consulted – Persons who are not directly involved in a process but provide inputs and whose opinions are sought.
- I - Informed – Persons receive outputs from a process or are kept up-to-date on progress, often only on completion of the task or deliverable.

Roles in the Sales/Initiation Process are:

- Product Owner (PO)
- Sales person
- Business Analyst (BA)
- Project Manager (PM)
- UX specialist (Designer)
- Solution Architect
- QA specialist
- Development Team (Tech Lead, Developer, HTML-Developer, QA, System Administrator)
- HR manager.

Documents of Sales/Initiation Process related to: Contract, Specification, Acceptance criteria, Project Plan.

TABLE 1 RASCI MATRIX

Stage	Planning phase		
	To do	RASCI roles	Artefact/ Outputs
Resource planning	Resource availability analysis (human resources, software, etc.)	PM) – A HR manager – S	Team members list
	Plan acquisition of additional resources (if required)	PM – A HR manager – S	
	Plan education of new human resources (if required)	PM – A	
	Schedule resource involvement	PM – A	
Product backlog planning	Prepare a detailed work breakdown structure	PM – A Team – R	WBS, Gantt Chart
	Prepare Gantt Chart (if required)	PM – A Team – R	
	Assign tasks to team members (if required)	PM – A Team – R	
	Prepare a detailed work breakdown structure	PM – A Team – R	
Budget planning	Prepare a detailed hours break down based on the WBS and the final estimate	PM – A Team – R	Hours break-down structure
Risk management planning	Prepare a risk management plan (mitigation measures, responsible people, deadlines, etc) based on the risk management report	PM – A Team – C (PO - C	Risk management plan Risk status tracking document
	Prepare a risk status tracking document, which will be updated during the project period	PM – A	
	Prepare a risk management plan (mitigation measures, responsible people, deadlines, etc) based on the risk management report	PM – A Team – C PO - C	
Project planning	Prepare a detailed project plan based on the WBS, the resource schedule, the risk management plan	PM - A	Approved project plan
	Present to the client and approve	PM - A	
Final approve	Check that everything is ready to start development and receive final approval from the client	PM – A	Approve from the client

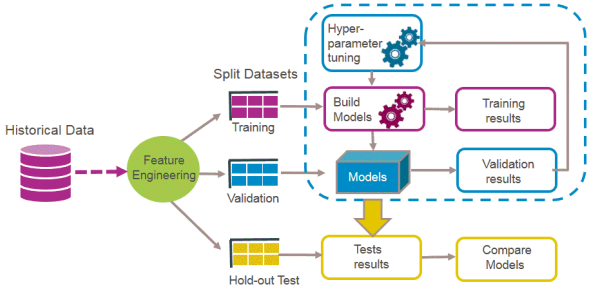
The Project process includes the following activities: Mandate / Scope, Risks, Milestones, Specification, Economy, Change Orders, Test and QA process and period. The documents of this Process are very useful for data analysis: Weekly reports, Change requests (added to specification), Go Live plan and Signed acceptance criteria [20].

And the last, but not list the Production Process is defined by the following activities:

- Acceptance from Customer,
- Invoicing,
- Support,
- Change Orders.

PM's daily duties include many different types of Business Processes and could be very regulated and clear described with using templates and tools according to the IT company standard. In our research we focused on the most important documents from the project life-cycle, which was used as an input data for analysis, in particular:

- *Project Charter*: the document created by PM and PO and consist of: General Information (Project name, Project manager, Data), Project Overview, Project Objectives, Project Type, Resource Costs and Estimates, Main stakeholders, Attachments.
- *Communication and Escalation Plan*: how communication will be managed during Project Lifecycle, customer contacts, corporate and supplier contacts, etc.
- *Risk Management: Risk Matrix*, *Risk Register*. Including a list of risks which described by several criteria's: ID, Risk Description, Impact, Risk Value, Mitigation, Trigger, Owner, Deadline.
- *Stakeholder Sheet:* Formalize the expectation and standards of internal and external customers, suppliers, employees, sponsors, beneficiaries, etc.



ML life cycle model for historical data analysis.

Based on *The Sales, Project, Production Methodology* the massive of data with target as an integral evaluation of the project success which allow for the complex correlations between separate features. This approach, which is presented in "Fig. 1", allows to describe all complex correlations between different project features, to consider the features importance and to zero the importance of those features which have a weak effect on the target. It means, that the PM should consider such features as a low-priority [7].

The disadvantage of this approach may be the complexity of presenting different projects in one vector space. Because there are no identical projects and there is an another problem which associated with small number of observations in multidimensional space. Considering these remarks, using of classical ML models such as linear models, random forest or boosting will not always give good results but certainly more adequate than using deep learning based on small datasets.

III. RESULTS AND DISCUSSIONS

In our research, it is very important step of multidimensional project vectors is the initial preprocessing and features selections for this stage, we have used a statistical approach based on Bayesian statistics Automatic Relevance Determination (ARD) [11, 18]. As a result, we received smaller dimensional vectors with features that were of greater statistical weight and in this case the use of classical ML models gave a more adequate result.

Example 1. We have the original data - only 30 points (data on 30 projects). And each project has 30 signs. And the task was to create a regression model (for example, to predict project sales volume according to location, type, sales area, configuration, number of features and other project parameters).

Building ordinary linear regression under such conditions will be pure insanity. Let us further exacerbate the problem by the fact that only 5 signs really matter, and the rest are completely irrelevant data [21].

Thus, let the real dependence be represented by the equation:

$$Y = W \cdot X + e \qquad (1)$$

where, e - is a random normal error and the coefficients W are equal [1, 2, 3, 4, 5, 0, 0, ..., 0], that is, only the first five coefficients are nonzero, and the signs from the 6th to the 30th generally do not affect the real value of Y. We only have data - X and Y - and we need to calculate the coefficients W:

```
# prepare the data, separate target
and form Xtrain та Ytrain
Xtrain=data.iloc[:,1:-1]
Ytrain=data.iloc[:,-1]
#Check the size of our massive
Xtrain.shape, Ytrain.shape
```

Using the *ARDRegression()* and will see, which features are important and influence to the target:

```
ard = ARDRegression()
ard.fit(Xtrain, ytrain)
# The higher of coefficient value, so
it is more important for the target
print (ard.coef_)
```

And the results are:

```
 [ 1.92557895e-03  -9.51988416e-04
-2.13868725e-04  -7.21210062e-04
 1.51514818e-02  4.13840050e-01
5.14818026e-01  8.82658008e-04
  9.09728493e-04]
```

Sort by decrease the value of coefficients, for analyzing the feature importance:

```
ard_df=pd.
DataFrame(columns=['Features_name',
'best_ard.coef_' ])
ard_df['Features_name']=Xtrain.columns
ard_df['best_ard.coef_']=ard.coef_
ard_df=ard_df.sort_values(by='best_
ard.coef', ascending=False)
ard_df
ard_df.to_csv('features importance_
ard.csv')
ard_df
```

Thus, having only 30 points in a 30-dimensional space, it is possible to build a model that identical the real dependence.

For comparison, the coefficients calculated using ARD regression:

```
array([ 1.92557895e-03, -9.51988416e-
04,
-2.13868725e-04, -7.21210062e-04,
1.51514818e-02,  4.13840050e-01,
5.14818026e-01,  8.82658008e-04,
9.09728493e-04])
```

The ARD are very popular for using in the different kernel-methods [8], for example Relevance Vector Machine (RVM) – this is Support Vector Machine (SVM) with ARD. It is also convenient in classifiers, when it is necessary to evaluate the significance of the available features from the projects.

According to *The Historical Data Approach* it is necessary to analyse the current historical project data from the company's analytical systems: an issue tracking systems, a team collaboration software, internal Customer Relationship Management (CRM) system, proprietary Human Recourses Management (HRM) system (for tracking HRM processes), code repository (GitHub, GitLab, etc.), the internal financial management system and others. The project processes proposed as acyclic graph-theoretical model, specifically Probabilistic Graphical Models (PGM). Such network may represent probabilistic relationships between estimates and inputs, it can be used to calculate the probabilities of evaluation of successful project realization with the different inputs that will be taken into account with their weights.

Disadvantage of the Bayesian network graph is than it does not consider into account the historical changes in the project. It may possibly correct by using the Long Short-Term Memory (LSTM) neural networks, where inputs will be historical data from the Bayesian network [6]. This will allow time to consider changing the status of implementation of our project and make more accurate predictions of its performance, it is this approach we believe can give better results.

IV. CONCLUSIONS

Our results show that proposed method could be the complexity of representing as the one of vector data model and lack of on a many-dimensional space. There for convention Machine Learning models such as: linear models, random forest, boosting are not always give reliable results but definitely more effective than using Deep Learning on the small data-sets. The initial processing and features selection are the important step in the case of many-dimensional vectors of the projects, using statistical approach based on the Bayesian probability Automatic Relevance Determination. Preliminary test shows the perspective of applying many tools and templates for project management, namely: Priority Matrix, Work Breakdown Structure, Gantt Chart, Project Budget, Risk Matrix, etc.

The implementation of our research allows more precise do project planning according the best practices, historical data and data on completed previous projects. In the next step, our research will allow to the project status threats and risks in the real time. The company's high level management would have an access information about the economic analysis, from the project initiation phase to feedback from all internal and external stakeholders.

REFERENCES

[1] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion". IEEE Trans. Pattern Anal. Mach. Intell., 12:629– 639, 1990.

[2] C. L. Philip Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," Inf. Sci. (Ny)., vol. 275, pp. 314–347, 2014.

[3] J. Alex Stark, "Adaptive Image Contrast Enhancement Using Generalizations of Histogram Equalization" IEEE Transactions on image processing, 2000.

[4] N. Bhargava, A. Kumawat, R. Bhargava, "Threshold and binarization for document image analysis using otsu's Algorithm ",International Journal of Computer Trends and Technology (IJCTT) – volume 17 Number 5 Nov 2014.

[5] D. F. Rogers and J. A. Adams, "Matematical elememnts for computer graphics", 2001.

[6] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., and Zhifeng Chen, e. a. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

[7] Aha, D. W., Kibler, D., and Albert, M. K. (1991). Instancebased learning algorithms. Machine Learning, 6(1):37– 66.

[8] Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 46:175–185.

[9] Bishop, C. M. (2006). Pattern Recognition and Machine Learning (Information Science and Statistics). SpringerVerlag New York, Inc., Secaucus, NJ, USA.

[10] Boehm,B.W.,Defense,T.R.W.,Group,S.,Boehm,H.W., Defense, T. R. W., and Group, S. (1987). A Spiral Model of Software Development and Enhancement. Computer (Long. Beach. Calif)., 21(May):61–72.

[11] Burch, C. (2010). Django, a web framework using python: Tutorial presentation. J. Comput. Sci. Coll., 25(5):154– 155.

[12] Freedman, D. (2005). Statistical models: theory and practice.

Inza, I., Larranaga, P., and Sierra, B. (2002). Feature Weighting for Nearest Neighbor by Estimation of Distribution Algorithms, pages 295–311. Springer US, Boston, MA.

[13] McConnell, S. (1996). Rapid Development: Taming Wild Software Schedules. Microsoft Press, Redmond, WA, USA, 1st edition.

[14] Project Management Institute (2004). A Guide To The Project Management Body Of Knowledge (PMBOK Guides). Project Management Institute.Ruder, S. (2016). An overview of gradient descent optimization algorithms. Web Page, pages 1–12.

[15] Tahir, M. A., Bouridane, A., and Kurugollu, F. (2007). Simultaneous feature selection and feature weighting using hybrid tabu search / k-nearest neighbor classifier. Pattern Recognition Letters, 28(4):438 – 446.

[16] The Bull Survey (1998). The bull survey. London: Spikes Cavell Research Company.

[17] Theano Development Team (2016). Theano: A Python framework for fast computation of mathematical expressions. arXiv e-prints, abs/1605.02688.

[18] Wilson, J. M. (2003). Gantt charts: A centenary appreciation. European Journal of Operational Research, 149(2):430 – 437. Sequencing and Scheduling.

[19] Chenarani, A., Druzhinin, E.A., Kritskiy, D.N.: Simulating the impact of activity uncertainties and risk combinations in R & D projects. Journal of Engineering Science and Technology Review 10(4), 1-9 (2017).

[20] Zou, K. H., Tuncali, K., and Silverman, S. G. (2003). Correlation and simple linear regression. Radiology, 227(3):617–628.

[21] MindK IT Company, March 2019. [Online]. Available: https://www.mindk.com/ [Accessed: March. 01, 2019].

[22] AlreadyOn IT company, March 2019. [Online]. Available: https://www.alreadyon.com/ [Accessed: March. 01, 2019].

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 56-60

Informational Warfare – Influence on Informational Structures

Peter Grabusts
Engineering Faculty
Rezekne Academy of Technologies
Rezekne, Latvia
peteris.grabusts@rta.lv

Aleksejs Zorins
Engineering Faculty
Rezekne Academy of Technologies
Rezekne, Latvia
aleksejs.zorins@rta.lv

Artis Teilans
Engineering Faculty
Rezekne Academy of Technologies
Rezekne, Latvia
artis.teilans@rta.lv

***Abstract*— The concept of information warfare encompasses the use of information and communication technologies to gain an advantage over a potential opponent. The information warfare is the manipulation with the information that trusts the goal, so that the goal should make decisions about its interests in the interests of opponents. Information structures are treated as systems that process different types of information, provide storage and access to users. Such structures may enclose neural networks, self-learning systems etc. They need to be ready to learn, respond to threats and ensure their safety, which is topical in today's information warfare. This paper will address aspects related to the security of information systems from a system theory point of view. The knowledge base of information structures can be elements of artificial intelligence, which security must be protected against various threats. The authors considers artificial neural networks to be one of the potential threats in the context of information warfare.**

Keywords*— *artificial neural networks, information structures, information warfare, neural networks.

I. Introduction

The information warfare have always existed - between individuals, groups, races, religions, states, cultures, civilizations. It is always the forerunner and driving force of different wars.

H. Lasswell [1] can be named the information warfare theorist of the first half of the 20th century. He actively used the methods of social psychology, psychoanalysis in the study of political behavior and propaganda, identifying the role of mass communications during information warfare of various states of the world for power. He outlined four main media functions:

- collection and dissemination of information;
- selection and commenting of information;
- formation of public opinion;
- spread of culture.

Obviously, all these components are active parts of the information warfare.

The strategy of waging the information warfare by targeting public opinion presupposes awareness of the moods of all social and ethnic groups, awareness of the real state of things. Consequently, on the one hand, informational and psychological effect through all possible channels, and on the other hand, a meticulous study of public opinion, that is, the revealing of the reaction - the relationship of the elite and the population to informational and psychological impacts, so that you can make adjustments to the impact parameters.

In order for the public to survive in the condition of information warfare, it needs an understanding of the information structures and their ability to counteract the effects of the information warfare. They try to store information so that it can be easily oriented in it, that is, to quickly find the necessary information element. Therefore, information is structured, that is, recorded in a specific pattern.

Information structure is now the most common term for those aspects of a sentence's meaning that have to do with the way in which the hearer integrates the information into already existing information.

An information system is a system that performs: obtaining input data; processing the data; issuing a result or changing its external state. Information warfare between two information systems is open and hidden targeted informational effects of systems on each other in order to obtain a certain gain.

Information impact is realized with the use of information weapons, i.e. such means that allow the planned actions to be realized with the transmitted, processed, created, destroyed and perceived information.

The aim of the work is to study the impact of information warfares on information structures.

II. The essence of the information warfare

The term "information warfare", the 4th generation war, was introduced in the late 80s and became widely used. Then, in the beginning of the 90s, the first theoretical and later practical works were published, where numerous definitions of the "information warfare" were presented.

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4035
© 2019 Peter Grabusts, Aleksejs Zorins, Artis Teilans.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

Nowadays, the concept "cyber war" is also widely used, which quite often is endued with content and meanings that are assigned to "information warfares".

The first deep definition of the concept "information warfare" was given in the 1996, in a report of the American RAND corporation "Strategic Information Warfare and the New Face of War" [2]. According to it, "Information warfare is a war in the information space". At that time 3 military spaces (land, naval and air) existed and a new one - information space - was added.

Afterwards, the headquarters of the "Joint Doctrine for Information Operations", 1998 [3], worked out the joint document where a definition of "information warfare" as "information operations - a conflict in which a critical and strategically important resource is information that must be exploited or destroyed " was given. It is a multidimensional concept, which represents only one aspect, the dimension of which is mainly military. The concept "information operations" gives a chance, more accurate than the traditional term "information warfare", investigate the place and role of information confrontation as elements of global confrontations.

There are a lot of other definitions of "information warfare", both official and non-official. D. Denning [4] in his work "Information Warfare and Security" noted: "Information warfare is a number of operations that have the purpose to gain or operate the information resources". Another researcher, G. Stein in his work "Information Warfare" [5] asserted: "Information war - is the use of information to achieve our national goals".

The most profound definition of "information war" was offered by the American theorist M. Libicki in his work "What Is Information Warfare?" dated 1995, where he distinguished 7 types of information warfares [6]:

- military confrontation for monopolizing command-control functions;

confrontation of intelligence service and counterintelligence;

- confrontation in the electronic sphere;
- psychological operations;
- organised spontaneous hacker attacks on information systems;
- informational-economic wars for controlling the trade of information products and monopolizing the information that is necessary to overcome the competitors;
- cybernetic wars in virtual space.

Information warfare can be used among the military and among civilians. For this purpose one of the types of information warfare or a set of actions can be used. The following types of information standoff can be defined:

- Information warfare on the Internet - different and often conflicting pieces of information are offered, which are used to confuse the enemy.
- Psychological operations - screening and supplying of such information, which sounds like a counter-argument on the mood that exists in society.
- Misinformation - stimulation of false information with a purpose to guide the enemy side to the wrong way.
- Destruction - the physical destruction or blocking of electronic systems that are important to the opponent.
- Security measures – intensification of the security of the resources with a purpose to save plans and intentions.
- Direct information attacks - combination of false and truthful information.

G. Stein issued the study "Information Warfare" [5], where he highlighted that information warfare operates with ideas. With regard to more specific aims, he claims the following: "The goal of the information warfare is the human mind, especially the one that makes the key decisions of war and peace, and the one that makes the key decisions about where, when and how to use the potential and opportunities that appear in their strategic structures".

In his book "War and Anti-war", A. Toffler [7] gives some examples of what is most often used to make an impact on others:

- accusations of atrocities;
- bid hyperbolization;
- demonization and dehumanization of the opponent;
- polarization;
- divine sanctions;
- meta-propaganda, which discredits the propaganda of the other side.

Presently, there are several means and methods of information warfare [8], [9]. The authors differentiates software and media.

Software can be categorized according to the tasks performed with their help. The following can be distinguished: information collecting tools, distorting and destroying information tools and tools influencing on functioning of information systems. Some tools can be universal and used to distort or destroy information, and to impact the functioning of information systems.

The main techniques and methods of applying information weapons are following:

- damage to particular elements of the information infrastructure;
- destroying or damaging of enemy information and software resources, overcoming protection systems, implantation of viruses and logic bombs;
- influence on software and databases of information systems and control systems with the purpose of their distortion or modification;

- taking over media channels with a purpose to spread misinformation, rumors, demonstrate power and bring in their demands;
- destruction and repression of communication lines, simulated overloading of switching nodes;
- influence on computer equipment with a purpose to disable it.

III. Information structures

Information weapon is directly related to the algorithms. Therefore, it is possible to call any system an information system – the object of information warfare – if it is capable of processing an algorithm according to the input data.

One of the crucial questions that shows the undecidability of the problem of winning an information warfare is the following: "Is the information system able to determine that information warfare has been launched against it?"

Why is it important to protect the information structure from information? It is vital because any information entering the system necessarily changes the system. Focusing informational impact can lead to irreversible changes and self-destruction.

Therefore, information warfare is nothing but explicit and hidden targeted informational actions of systems on each other with a purpose to get a certain win. Practising of information weapons means the supplying to the input of the information self-learning system such a sequence of input data, which activates certain algorithms in the system.

The conclusion is, that information weapon is, first of all, an algorithm. Applying an information weapon means choosing the input data for the system in such a way that it could activate certain algorithms in the system, and in the case of the absence of necessary ones activate the algorithms for generating the necessary algorithms.

Further it goes about information structures - learning systems - in the simplest assumption it could be an artificial neural network (ANN) and social networks. It is assumed that the information structure is the bearer of knowledge and knowledge of the information system is expressed through its structure. Then, to estimate the amount of information received by the system, it could be logical to use such a concept as the degree of structure modification by the input data.

It can be said that the information structure is stable against external influences if the number of its elements does not have sharp fluctuations from these influences.

What structure should the system have so that the number of its elements do not experience sharp fluctuations? This is a structure in which there are several groups of elements that are closely related to each other, but the connections between the groups are very unstable, for example: structure A: 1 – (2, 3, 4), 2 – (1, 3, 4), 3 – (1, 2, 4), 4 – (1, 2, 3, 5), 5– (4, 6, 7), 6 – (5, 7), 7 – (5, 6).

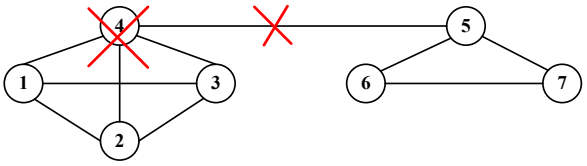


Fig. 1. An example of stable information structure.

In structure A, it is enough to destroy the element with number 4 and, as a consequence, the number of elements of the system will be reduced by half. Clear, that this structure is not stable (unstable is any structure in which there are single elements that carry out a bundle of groups of elements) [10].

Conversely, the most stable system can be considered a system in the structure of which there is the maximum number of connections — each is connected to each, i.e. each element is basic.

Different information systems may contain various information structures that can be described in an analytical or graphical way that describes the knowledge base of structures. It is necessary to formulate task classes necessary for describing information structures [10]:

- Assessment of information structure capabilities.
- Assessing the form of information structure, which is maximum resistant to external threats at the moment.
- Determining the minimum configuration of an information structure that is resistant to threats.
- Predicting the effect of potential changes in information structures.
- Attempts to predict a factor that could affect the stability of the structure and, as a result, it could be modified or collapse.

For example, what should be the impact strategy on the structure shown in Figure 2, so that the realization of the external impact could lead the system to destruction (structure A: 1 – (2, 3), 2 – (1, 4, 5), 3 – (1, 6, 7), 4 – (2), 5– (2), 6 – (3), 7 – (3))?

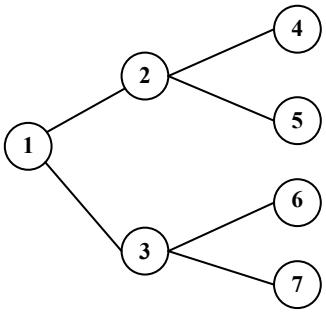


Fig. 2. An example of information structure.

IV. Risk of information structures

In general, artificial neural networks cannot be considered as stable information structures. It is related to various learning algorithms, which mostly work on the "black box principle", which can make them unprotected to various external threats.

A popular approach in machine learning and perception is artificial neural network. Traditionally, they are assigned to the properties of self-learning, self-organization, ability to process figurative information as opposed to the usual algorithms, which are also traditionally considered to be rigidly defined, untrained, and intended for processing symbolic information.

The more complex the network, the more parameters it contains, the more data is required for its training. Usually we do not understand what connection the trained neural network has with the simulated phenomenon. It is unclear in details why it works and we can not predict in which cases it can fail.

In recent years, the issue of restricting artificial intelligence (AI) has become topical [11], [12].

An AI box is a hypothetical isolated computer system where a possibly dangerous AI, is kept constrained in a "virtual prison" and not allowed to manipulate events in the external world. Such a box would be restricted to minimalist communication channels. Unfortunately, even if the box is well-designed, a sufficiently intelligent AI may nevertheless be able to persuade or trick its human keepers into releasing it, or otherwise be able to "hack" its way out of the box [11].

The authors offers his own viewpoint of AI as a protection of information structures during information warfare [13].

In terms of information warfare, a certain threshold is set against an AI system (ANN or social network based on it), which, apparently, should be counted according to some methods taking into account certain activities within the system (fake news, social polls, etc.). The importance of the problem must be recognized by the corporation and, in case of a critical situation, the government.

In any case, the system should have a built-in mechanism that could be called a trigger, which should respond to an emergency intrusion into its structure in the context of an information warfare. At the same time, the system is learning, re-learning, and self-learning.

If, in case of an information warfare attack against the information structures the trigger had to respond, five situations would be possible (see Fig. 3):

- trigger "ON" – the self-destroyed mechanism is launched – the network activity is paralyzed, links are destroyed. The AI box protocol is interrupted;
- trigger "OFF" – the attack is treated as false alarms and the system continues to work in the previous mode under the AI box protocol;
- trigger "NEUTRAL" – the attack is treated as an unknown alert and the system continues to work in the previous mode under the AI box protocol, but by intensifying the analysis of the causes of the attack and trying to identify and prevent future threats;
- trigger "COUNTERATTACK" – self-learning allows the system to exit the AI box protocol framework and the effects are not predictable.
- trigger "UNKNOWN" – the effects are not predictable.

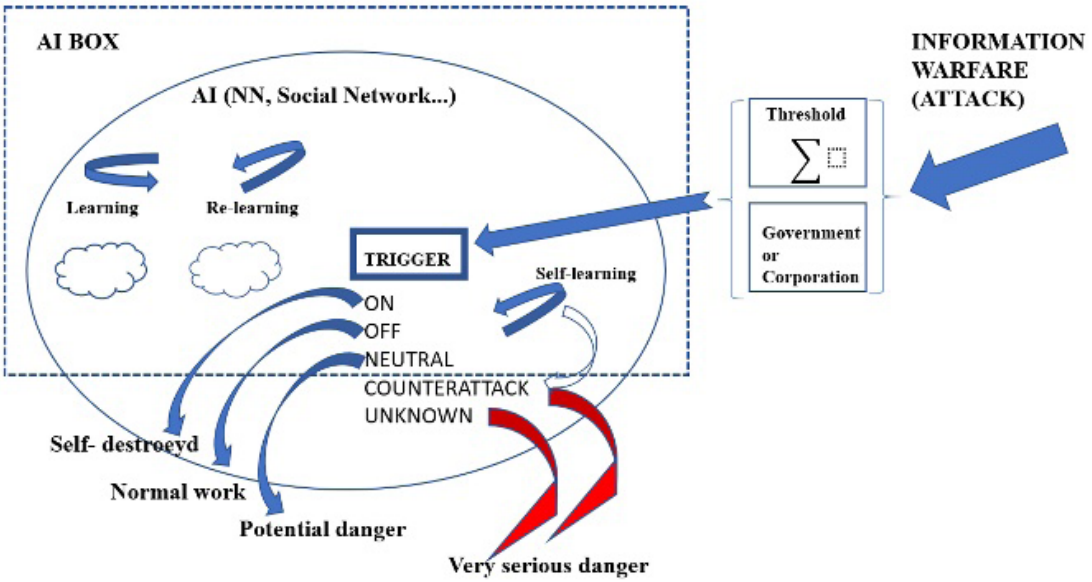


Fig. 3. Potential counteraction structure in case of information warfare attack.

V. Conclusions

Information warfare in essence is a war of technology. It is a war in which exactly the structures of systems, as knowledge holders, interfere. It is necessary to talk about the methods of information warfare because an understanding of the techniques of information warfare makes it possible to transfer it from the category of hidden threats into explicit ones that can be dealt with.

Consequences of information warfare:

- death or emigration of part of the population;
- the destruction of industry;
- loss of part of the territory;
- political dependence on the winner;

- the destruction (sharp reduction) of the army or a ban on its own army;
- export of the most promising and knowledge-intensive technologies from the country.

Information warfare is nothing but explicit and hidden targeted informational actions of systems on each other with a purpose to get a certain win. Practising of information weapons means the supplying to the input of the information self-learning system such a sequence of input data, which activates certain algorithms in the system.

The research presents a description of a potential counteraction against the threats of information warfare against information systems (AI based on artificial neural networks).

References

[1] H. Lasswell, "The Structure and Function of Communication in Society," in The Communication of Ideas, L. Bryson, Ed. Institute for Religious and Social Studies, 1948, p. 117.

[2] R. C. Molander, A. Riddile and P. A. Wilson, "Strategic Information Warfare: a New Face of War," RAND Corporation, 1996. [Online]. Available: https://www.rand.org/pubs/monograph_reports/MR661.html [Accessed: March 10, 2019].

[3] "Joint Publication 3-13/Information Operations," Oct. 9, 1998. [Online]. Available: http://www.c4i.org/jp3_13.pdf. [Accessed: March 10, 2019].

[4] D. E. Denning, Information Warfare and Security. Addison-Wesley, 1999.

[5] G. J. Stein, "Information Warfare," 1995. [Online]. Available: http://www.iwar.org.uk/iwar/resources/airchronicles/stein.htm[-Accessed: March 10, 2019].

[6] M. C. Libicki, What Is Information Warfare?, National Defense University, Institute for National Strategic Studies, 1995.

[7] A. Toffler, War and anti-war. Survival at the dawn of the 21st century, Little Brown & Co., 1993.

[8] A.Nestoras, "Political Warfare: Competition in the Cyber Era," IEEE International Conference on Big Data, Big Data 2018, 10-13 December 2018, https://doi.org/10.1109/BigData.2018.8622490.

[9] V.Duddu, "A survey of adversarial machine learning in cyber warfare," Defense science journal, Volume 68, Issue 4, July 2018, Pages 356-366, doi: 10.14429/dsj.68.12731

[10] S. P. Rastorguev, Information Warfare, M: Radio and Communication, 1998 (in Russian).

[11] D. Chalmers, "The Singularity: A Philosophical Analysis," Journal of Consciousness Studies, vol.17, no. 7-65, Jan. 2010.

[12] R. V. Yampolskiy, "What to Do with Singularity Paradox?," in Philosophy and Theory of Artificial Intelligence, vol. 5, V. C. Muller Ed. Berlin, Germany: Springer-Verlag, 2013, pp. 397-413. https://doi.org/10.1007/978-3-642-31674-6_30

[13] P.Grabusts, "Ensuring the security of information structures in today's environment," 59th International Scientific Conference on Information Technology and Management Science of Riga Technical University, ITMS 2018 – Proceedings, 29 November 2018, https://doi.org/10.1109/ITMS.2018.8552976.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 61-64

Modern Algorithms to Identify Plagiarism

Aleksejs Grocevs
Software engineering departmentRiga Technical University
Riga, Latvia
Aleksejs.Grocevs@rtu.lv

Natālija Prokofjeva
Software engineering department Riga Technical University
Riga, Latvia
Natalija.Prokofjeva@rtu.lv

***Abstract*—Nowadays information technology industry is growing extremely fast. To solve business needs, address researcher demand for problem solving a lot of programs are built from scratch. However, not all developers are fair enough to align their products with the corresponding library or another (open source) product licenses, i.e. copyrights are being violated, intentionally or due to familiarization with another source code. To address this issue in past decade multiple plagiarism detection techniques and algorithms were invented. Despite the fact, that many of them are capable of code comparison on meta-level, modern Integrated Development Environments (IDEs) provide convenient way to modify program source code without actual re-writing, preserving the original code workflow and avoiding plagiarism detection. This paper will compare and identify available approaches to apprehend this issue, as well as provide insights for the future this problem mitigation.**

Keywords— plagiarism, algorithm, abstract syntax tree, graph.

I. Introduction

This document oversees the existing plagiarism detection algorithms, the main idea of the article is in comparison of most popular algorithms with insights towards their types and structures they do operate with. During the document the text, structure based and semantic algorithms have been covered.

With the purpose of diving in details of plagiarism detection algorithms the analysis of common structures like AST (abstract syntax tree) and Graph structures have been performed, hence mentioned structures are quite common solutions of building algorithms of mentioned types are being used those days.

The following specific algorithms have been reviewed in the document: The Greedy String Tilling, Kolmogorov complexity algorithm and Fingerprint method. Additionally, the abovementioned approaches were comparative analyzed.

II. ALGORITHM TYPES

Initially the plagiarism detection algorithms were based on quantitative comparison of different program's characteristics like: average string length within the program, number of variables are being defined and are being used, average variable naming length, number of "if else" operators. So the two similar programs would have close or equal number of mentioned constructions. Such algorithms were based on correlation counting principle and were well renown as "attribute counting systems".

The main disfunction of such algorithms happened because the only overall number of certain constructions haven't been analyzed, instead of in-depth analysis of program's logic to be performed. So the similar inclusion of code being plagiarized within the original program couldn't be discovered.[3]

Algorithms shown above provides just the overall, quite approximate result of correlation within the program's syntax. The better result could have been provided by the so cold "control-flow graph" algorithms which are being oriented to compare the changes within the programs structure.

The above stated analysis led the authors towards the comparison of the most modern algorithms from the text based, structure based and semantic algorithms types.

In the first part of this article the major plagiarism detection algorithms will be reviewed with the detailed comparative analysis afterwards. Nowadays there are three main types of plagiarism detection algorithms: those are based on the text, structure or semantics analysis.

Text algorithms are based on the text based program code perception, as an alphabet where one symbol is equal to the proper operator from the programming language, named token. All the text algorithms are based on the idea of such token or token groups comparison within the programming code.

Granularity of the token groups within the code to be inter-compared proportionally equals the effectiveness of the algorithm.

Historically text algorithms were the first widely used ones, their effectiveness is based on investment should be made in recurrent comparison constructions of the interchangeable symbol blocks to assure the most frequent combinations are being checked.[2]

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4058
© 2019 Aleksejs Grocevs, Natālija Prokofjeva.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

Structure based algorithms are based on the analysis of the program's structure, where the program logic interpreted in the AST (abstract syntax tree) are being compared. AST (abstract syntax tree) or the Graph of the program's flow are the two names of the program logic tree implementation [1]. Both in AST (abstract syntax tree) or in the Graph the algorithm operates with the program logic as a tree, where such constructions as "if. else." are being perceived as the branches of the tree. Therefore, each program is being perceived as a tree structure what results in quite convenient recurrent comparison between different branches. Important to note that the algorithms using AST (abstract syntax tree) structure are being quite complex in realization and quite advanced supporting technical infrastructure is being required [1].

Semantic algorithms are the algorithms using the Graph relations with two different types of elements, where one element represents the operator type and the connected element represents the type of relation, they are being in. Based on that Graph there is possibility to check each tree limb independently resulting with the quite effective way of comparison between the different code structures.

Each of the main plagiarism detection algorithm types are quite similar and gives the opportunity to analyze the program or it's parts. Such algorithms work well against typical ways of plagiarism, where students do try to change the places of operators within the code.

III. Common structures are being used within the algorithms

Alongside with review of certain algorithms author would like to review common structures they are being based on: AST (abstract syntax tree) and Graph structure.

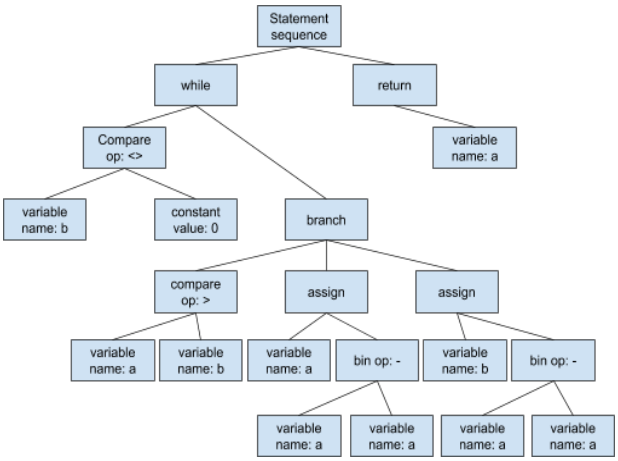


Fig. 1. AST (abstract syntax tree) structure [1]

Figure above (Fig 1.) shows AST (abstract syntax tree) represented in the form of interdependent blocks of code. On the Fig. 1. there could be seen quite different constructions even in such short code example. Comparing part of such tree with the another one there is possibility to successfully fight such simple plagiarism ideas as replacement of the code parts within one program, quite common case in the coursework plagiarism in the high schools.

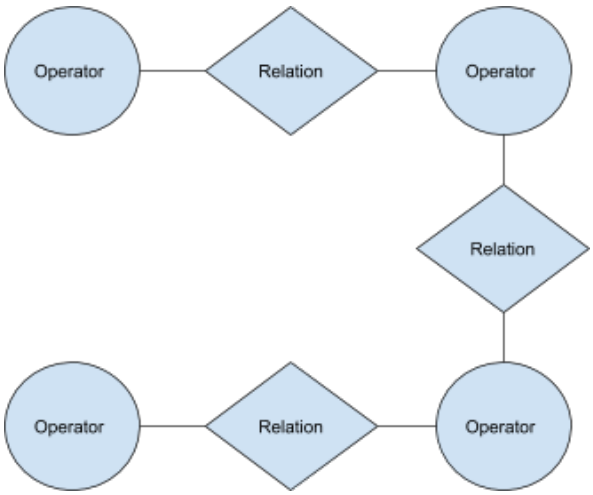


Fig. 2. Graph structure [1]

The Graph structure provides the different idea to the Structure based and Semantic plagiarism detection algorithms types.

Within the Graph structure program code is been transferred to the graph where vertex and edges are being defined by the operators and their relations.[9]

Mentioned structures are being widely used within the various specific plagiarism detection algorithms.

Use of this structure is useful in searching the core functions, blocks in code within the different programs, may result in finding plagiarism even on idea level within students solving one complex problem.

IV. Specific Algorithm Realizations

"The Greedy String Tiling" is the subset of the "Text" plagiarism detection algorithm type. This algorithm interprets to programs to be compared as two lines of text based on the unified alphabet (usually named the token variety) and results with the one line of text - the variety of the unique tokens.

The main two ideas of the comparison are: algorithm does not count in coincidences with the too small number of same tokens. The major findings to count in examining the result are the longer coincidences even if summarized length of the small coincidences is bigger. Such logic could be easily explained, so there could be same parts of code constructions to be used by two different developers, but they are many in numbers and short, on the other hand the long lines of same tokens is the clear sign of possible plagiarism within the programmed solution. [3]

Therefore, not counting in the smaller coincidences just protects the algorithm from the possibility of random matches.

"Kolmogorov complexity algorithm" : "K (s) = min{|p|, U(p) = s}", the algorithmic complexity *K (s)* of a string is the length of the shortest program *p* that produces *s* running on a universal Turing machine *U*. Algorithm counts the length of non-matched tokens between the matching sequences. [5]

This method is being called (an information-based sequence distance). The main advantage of this method

is its versatility, since it will understand the appearance of matching element based on every possible match principle, what makes it much more universal than other plagiarism detection algorithms. So, as the "Kolmogorov complexity algorithm" states, the smaller is the average length of non-matched tokens the higher is the probability of plagiarism within two programs.

"Fingerprint method" - specific realization of "Text" type plagiarism detection algorithm. In "Fingerprint" method there are stored vocabulary of "token combinations" are being named "Fingerprints".

So, in this method the search and comparison are not being handled on token to token principle, but on the other hand by searching specific token, or token combination in codes are being compared.[8]

The "Fingerprint" is much more convenient for various interdependent searches, where search is being handled against the "Fingerprint" base, could be stored in format of simple DB solution. [4]

Usually the "Fingerprint" plagiarism detection algorithm is being realized in following steps:

1.) use of hashing principle for the sequence of tokens (program);

2.) received subset of hash-codes to be put within the hash table;

3.) Comparison of hash tables with the "Fingerprint base" the subsets with the higher risk for plagiarism will be defined.

TABLE I. Plagiarism detection algorithm comparison

	Algorithm name	***Pros***	***Cons***
1.	Greedy String Tiling	Effective in comparing one-to-one program codes	Is not effective enough to use this type of text algorithms with DB of code samples
2.	Kolmogorov complexity algorithm	Versatility, since it will understand the appearance of matching element based on every possible match principle, what makes it much more universal than other plagiarism detection algorithms	Is not effective enough to use this type of text algorithms with DB of code sample
3.	Fingerprint method	"Fingerprint" is much more convenient for various interdependent searches, where search is being handled against the "Fingerprint base", could be stored in format of simple DB solution.	Infrastructure with quite high-performance characteristics is being required to support this solution

Comparison of the "Greedy String Tiling", "Kolmogorov complexity algorithm" and "Fingerprint method" algorithms has shown the different specifics of the, above mentioned algorithms, there are different situations and circumstances they could be applicable and the most effective in.

Each of the reviewed plagiarism detection algorithms could be used in the high school environment, but the "Fingerprint method" algorithm looks to be the most useful one while checking the student's developed program against the DB of the program samples from the previous courses. On the other hand, "Kolmogorov" algorithm should be chosen for the more complicated, choice situations.

V. Other Conclusions

Concluding the article following findings should be highlighted:

Major plagiarism detection algorithm including their specific realizations can handle the plagiarism problem effectively, on the other hand there are proper pros and cons against choosing the one or another of them.

For example, no other algorithm except the "Fingerprint" algorithm can work effectively with the big number of code examples to perform the comparison with. If this type of solution is being required the "Fingerprint" algorithm is the most effective one, since it's speed of work is directly dependent from the number of code examples to compare with.

There are big difference in terms of quality and quantity within the performed analysis, since the "text" type plagiarism detection algorithm are not so precise in terms of detecting the certain logical constructions in comparison with "structure-based algorithms" or "semantic algorithms".

On the other hand, "structure-based algorithms" which use AST (abstract syntax tree) or "semantic algorithms" using the Graph structure do provide the much clearer picture of logical structures are being used within the compared programs code.

Within the big number of existing plagiarism detection algorithms there are the most effective ones like: "Fingerprint" algorithm and algorithms based on the AST (abstract syntax tree), since they are the only ones which can effective and automate the plagiarism check using the DB of comparative examples in the process.

Mentioned algorithms go assure the expected speed based on hash-tables usage in comparison process. This is the main reason of these algorithms' usage in the high schools, where are lots of student's generated code examples on a quite limited number of topics within the disciplines.

Here is very important to understand that during each year students of different courses do repeatedly work on quite similar problems within their courses, since there is no real possibility to significantly differ the course content as frequent as new groups of students do arrive, on the yearly basis at least.

This is the main reason for using "Fingerprint" and AST (abstract syntax tree) based algorithms against the Database of the student works submitted from the previous years.

References

[1] Baxter I., Clone Detection Using Abstract Syntax Trees. ICSM. 2008.

[2] Mishne G., Source Code Retrieval using Conceptual Similarity, RIAO, Vaucluse, 2004.

[3] Chen X., Francia B., Shared Information and Program Plagiarism Detection, IEEE Information Theory. 2004.

[4] A.Grocevs, N.Prokofjeva, "Modern programming assignment verification, testing and plagiarism detection approaches." Proceedings of the IVUS International Conference on Information Technology, pp. 61-64, 2017.

[5] Prechelt L, Malpohl G, Philipsen M. Finding plagiarism among a set of programs w/th JPlag. J. UCS. 2002 Nov 28;8(11): 1016.

[6] Whale G. Plague: plagiarism detection using program structure. School of Electrical Engineering and Computer Science, University of New South Wales, 1988.

[7] Joy MS, Sinclair JE, Boyatt R, Yan JK, Cosma G. Student perspectives on source-code plagiarism, International Journal for Educational Integrity. 2013;9(1):3-19.

[8] A. Parker and J. Hamblen. Computer algorithms for plagiarism detection. IEEE Transactions on Education, 32(2):94099, 1989.

[9] Fintana FA, Mangiacavalli M, Pochiero D, Zanoni M. On experimenting refactoring tools to remove code smells. InScentific Workshop Proceedings of the XP2015 May 25 (p. 7). ACM

Comparative Evaluation of the Rule Based Approach to Representation of Adaptation Logics

Lauma Jokste
Information Technology Institute
Riga Technical University
Riga, Latvia
lauma.jokste@rtu.lv

***Abstract*—Due to the rapid growth of business processes digitalization, enterprise applications cover more and more business and daily life functions thus becoming more complex. Complex enterprise applications often deal with low users' satisfaction of usability. This problem can be solved by implementing adaptation algorithms in enterprise applications, so they can be adjusted for specific context situations and specific users' needs. Some adaptation logics representation techniques are complex and require specific knowledge and skills to manage and modify adaptation process. In this paper rule based adaptation approach is introduced where rules are used as means to manage and modify adaptation process. Rules are easy to read and understand, thereby rule based adaptation should ensure elastic, transparent and easy administrable adaptation process. The goal of this paper is to test this statement by carrying out a comparative adaptation logics representation evaluation experiment. During the experiment participants are required to complete tasks which include different forms of adaptation logics representation (code, rules and models). Experiment results are analyzed by qualitative and quantitative measures such as users' understandability of applications behavior when adaptation case occurs and users' satisfaction with adaptation logics representation. Experiment results are summarized and are to be used for further development of the study.**

***Keywords*—Adaptation process evaluation, Adaptive enterprise applications, Adaptivity, Rule based adaptation.**

I. Introduction

Adaptation can be defined as any process which modifies or extends the behavior of the system in order to improve its interactions with the surrounding parts of the system [1]. Adaptation of enterprise applications is one of the solutions how to improve applications' performance [2] and users' satisfaction [3]. Adaptation algorithms can be represented and implemented in different forms, e.g. software code, models and business rules.

McKinley et al. [4] summarize different software tools and technologies which allow to write self-modifying code and develop built-in adaptation mechanisms. For such systems adaptation code can be implemented in applications middleware level. Such approach ensures transparent adaptation process but leads to limitation that this approach can be used only if applications are written against the specific middleware platform. Adaptation algorithms can be also implemented in applications functional code, but this approach leads to the complex functional code management.

Models can help to bridge the gap between developers and stakeholders as they can be understood by stakeholders without specific IT knowledge as well as can be executed by system components typically in model driven development approaches [5]. Understandability level can be increased by using text-based syntaxes in models thought developers prefer them to graphical editors [6].

Rule based systems development began in the 1960's [7] and are still evolving in different directions such as expert systems [8], software configuration management systems [9] and adaptation systems [10]. Rules are attributed with a lot of advantages mainly related to their form which is easy to read, easy to understand by both – humans and computer system components, easy to modify and manage [11] – [13]. Rule based adaptation systems also have some shortcomings and limitations. Zacharias in [14] reports rule based system developers survey results which outline problems related to tool support for such systems development, debugging difficulties and run-time performance problems. Also problems with managing large amount of rules and dealing with conflicting rules should be considered when developing rule based adaptation systems [15]. Each adaptation logics representation format has some advantages and disadvantages. Adaptation logic expressed in a form of code is easier to develop and implement, while models and business rules improve communication between humans and software components, but often deal with implementation and development difficulties. In this paper emphasis is put on rule based adaptation by evaluating such adaptation approach's comprehension and ease of use in comparison of other adaptation logic representation

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4156
© 2019 Lauma Jokste.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

formats. The objective of this paper is to report the results of an experiment during which comparative evaluation of different adaptation logics representation (code, models and business rules) was carried out. Experiment results are analyzed by qualitative and quantitative measures.

The paper is structured as follows: Section II describes the rule based adaptation approach which ensures adaptation process in a form of business rules. In section III experiment design is given. Section IV summarizes experiment results and in Section V concluding remarks and future challenges are discussed.

II. Rule based adaptation approach

We have developed a context-aware rule based adaptation approach [16] which combines context monitoring, defining software entities dependencies of contextual situations data and software entities adaptation in a form of business rules. For applications adaptation purposes a separate adaptation administration module can be built and integrated with enterprise applications. This approach does not require interfering in applications code thus ensuring independent and easy administrable adaptation process.

Adaptation process includes two different types of adaptation rules: context dependency rules describe software entities dependencies of context data and can be used to determine the amount of context data that needs to be monitored and stored for further processing in adaptation process. Context dependency rules can be derived from context sources by data mining methods. Adaptation rules define an adaptation task which should be executed when particular context situation occurs. Main concepts for this adaptation approach are summarized in concepts model shown in figure 1. Adaptation rules allow to manage adaptation process by defining new rules, modifying existing rules, deleting rules or changing rules weights. Adaptation rules can be manually defined and when there will be accumulated enough historical data about adaptation rules performance, machine learning methods can be applied to generate new adaptation rules automatically.

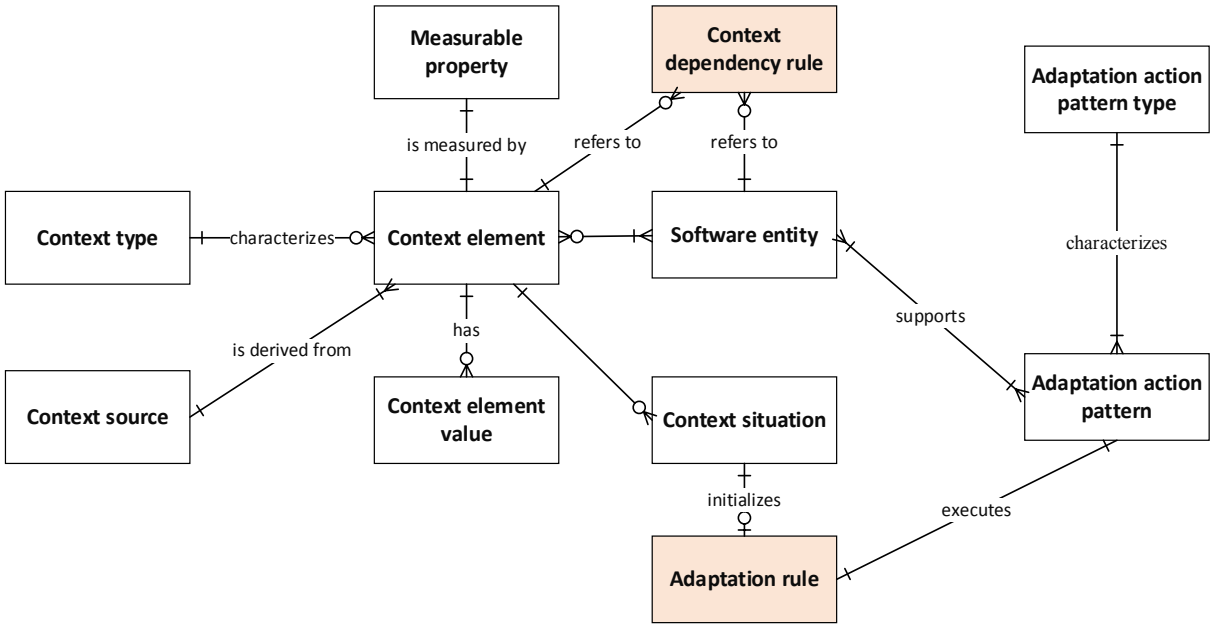


Fig. 1 Context-aware rule based adaptation approach concepts model.

In the proposed approach semantic form of event-condition-action rules is used for adaptation rules. Each adaptation rule can be described as *IF (context situation) THEN (adaptive action)*. Adaptation rules also include parameters such as adaptive action input data and expected user action. Adaptation rule concepts model is shown in figure 2.

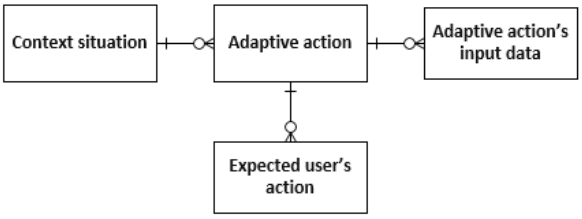


Fig. 2 Adaptation rule concepts model.

Expected user action concept allows to evaluate if adaptation rule execution gives an expected result. After executing adaptation rule, adaptation algorithm searches in data source (typically users' action logs) if user have completed expected action in expected time period. Depending on the search results, adaptation rule weight is recalculated thus ensuring automatic adaptation rules quality management.

III. Experiment design

An experiment was carried out to test the hypothesis regarding the rule based adaptation: Rules as adaptation format ensure transparent, elastic, easy understandable, modifiable and manageable adaptation process which can be administered by users without specific IT skills. The goal of the experiment was to test how different user groups understand different adaptation logic representation formats and how they evaluate simplicity and perceptibility of these formats.

The experiment was developed using online survey tool. It consisted of 6 different tasks. In each task an

adaptation case was described in different adaptation logics representation form – software code, model or business rule (see task examples in figures 3 – 5). Each adaptation task included context situation in which adaptation task should be initiated and executed in application with given adaptation action input data. For each task participants were asked to: (1) describe in free text form applications behavior when adaptation case occurs; (2) evaluate how easy it was to understand applications behavior in given adaptation task representation format. For (1) participants answers were later analyzed by qualitative measures and was rated by 3 points system where 0 refers to complete incomprehensibility, 1 refers to partial comprehensibility and 2 refers to complete comprehensibility of adaptation task. For (2) a linear Likert scale was used with values from 1 to 5 where 1 refers to strongly disagree that adaptation case was easy comprehensible and 5 refers to strongly agree that adaptation case was easy comprehensible.

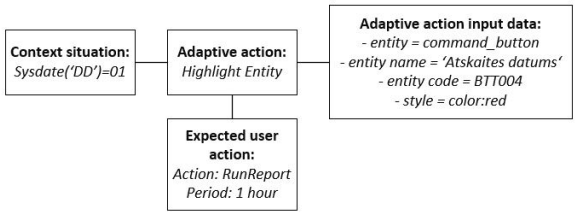


Fig. 3 Adaptation case example – adaptation rule.

In total 20 participants completed the experiment. Testable hypothesis determines that rule based adaptation is easier to understand and manage by users without specific IT knowledge and skills because adaptation cases are described in natural language form – business rules. Accordingly 10 participants were students and/ or specialists from IT field (further group_1) and 10 participants were specialists from other fields without specific IT knowledge and skills (further group_2).

Context situation	Adaptive action	Input data

IF (select count(1) from audit_json_logs a where a.stamp >= ('now'::timestamptz - INTERVAL '60 minutes') and a.proc_name like '%delete%') > 10	
	<script type="text/javascript">
	appCasea.controller("appCaseaCntr", function ($scope, appRentService, $interval, $window) {
	$scope.loadCaseaData = function () {
	var obj = {};
	obj.system = 'RENT';
	appRentService.executeHttp('app_casea_pck.getCaseaData', obj).then(function (data) {
	if (data && data.adaptionExists) {
	if (data.adaptionType == 1)
	Udv.Common.ShowInfoDialog(data.adaptionText);
	} else if (data.adaptionType == 2) {
	Udv.Common.ShowInfoDialog(data.adaptionText + ' ' + data.adaptionLinkName + '?', {
	buttons: {
	'No': function () {$(this).dialog('close');},
	'Yes': function () {$(this).dialog('close');
	$window.open(data.adaptionLink, '_blank');}}});
	} else if (data.adaptionType == 3) {
	$window.open(data.adaptionLink, '_blank');};}});}
	//after each 30 seconds
	setInterval(function () {
	$scope.loadCaseaData();
	}, 30000);});
	angular.bootstrap(document.getElementById("appCasea"), ['appCasea']);
	</script>
	
		data.adaptionType == 2
		data.adaptionText = 'www.usermanual.lv'
		data.adaptionLinkName = 'User manual'

Fig. 4 Adaptation case example – software code

Context situation	Adaptive action	Input data
IF (Number of deleting actions per user per hour) > 200	Send e-mail	E-mail subject = Context monitoring data E-mail text = High deleting action count fixed during last hour Receivers e-mail: test@test.lv

Fig. 5 Adaptation case example – adaptation rule.

Online experiment tool also registered the time each user spent on execution of the experiment with a purpose of determining whether there is difference between two participant groups regarding the time needed to be used to understand and complete adaptation task.

IV. EXPERIMENT RESULTS

It took on average 18 minutes to complete the experiment (group_1 average 21 minutes and group_2 average 15 minutes). Difference between groups regarding average time spent to complete the experiment can be explained by the fact that participants with IT skills and knowledge tried to understood adaptation tasks in form of software code while participants without IT knowledge mostly marked they do not understand the adaptation task and did not spent time on examining the software code.

In each task participants were asked to describe applications behavior based on given adaptation case. Each answer was subsequently analyzed by qualitative measures and rated from 0 – 2 where 0 – complete incomprehensibility, 1 – partial comprehensibility and 2 – complete comprehensibility. Mean scores for each adaptation format were calculated and results are summarized in figure 6. Results showed that group_1 demonstrated a good understanding of software code and participants were able to describe applications behavior when defined context situation will occur and adaptation task will be executed. Adaptation cases in a form of models created a lot of confusion for both groups while all participants demonstrated very good understanding of adaptation cases described in a form of business rules.

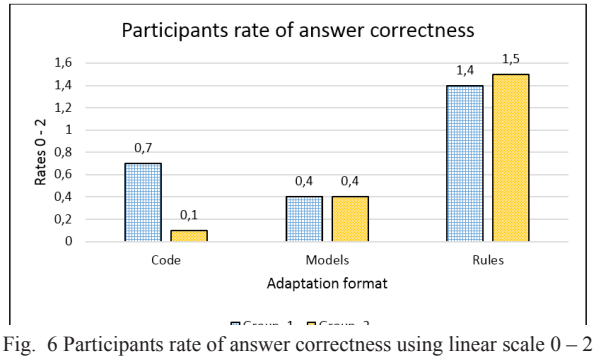


Fig. 6 Participants rate of answer correctness using linear scale 0 – 2

After completing each task of the experiment, participants were required to evaluate the comprehensibility of the given tasks adaptation logics representation format by rating their agreement to the statement *I understood applications behavior in the given adaptation case* using linear scale from 1 – 5 where 1 refers to *strongly disagree* and 5 refers to *strongly agree*. Figure 7 shows mean values of evaluation results.

Results showed that participants with IT knowledge (group_1) evaluated adaptation tasks in a form of code more understandable and simple than participants without specific IT knowledge (group_2), but still rate was low (average 1,95 out of 5). In order to read, understand and modify built-in adaptation process, adaptation administrator should be familiar with specific programming language and development technology which limits the simplicity of such adaptation logics representation format management even for users from IT field.

Participants evaluation of adaptation models comprehensibility were average 2,4 out of 5, but the actual correctness of the answers to tasks where adaptation was described in form of models were low (see figure 6). Comprehension level of adaptation models can be dependent of modelling syntax. In order to increase the level of comprehension for adaptation models as adaptation logics representation format, adaptation tasks should also include some explanation about model components and syntax.

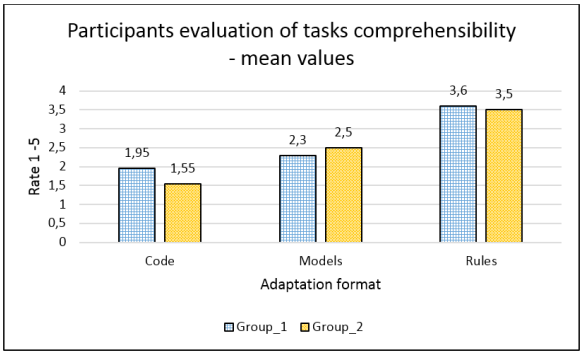


Fig. 7 Participants evaluation of tasks comprehensibility – mean values.

Experiment results showed that participants did well with tasks where adaptation logics were represented in a form of business rules. Participants demonstrated good comprehension of applications behavior when defined context situation will occur as well as highly appreciated tasks simplicity and intelligibility. Results quite slightly differed between both participant groups as business rules are described in a form close to natural language and there are no specific knowledge required to read, understand and manage them.

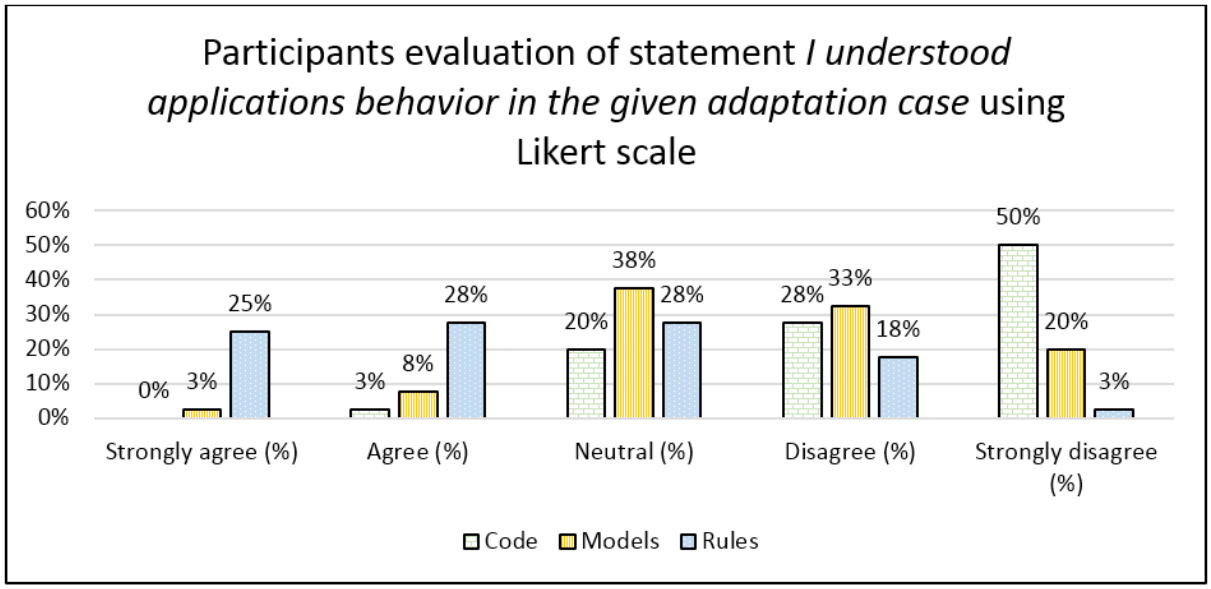


Fig. 8 Participants evaluation of statement *I understood applications behavior in the given adaptation case* using Likert scale.

In figure 8 participants agreement to the statement *I understood applications behavior in the given adaptation case* using Likert scale is shown. Diagram shows that 25% of participants strongly agreed to the statement regarding business rules while none of the participants strongly agreed to the statement regarding software code which certifies rules comprehensibility.

Overall experiment results allowed to test the hypothesis regarding rule based adaptation process transparency and comprehensibility which are important characteristics of adaptation process management. Experiment results also allowed to compare adaptation process comprehensibility for different adaptation logics representation formats.

V. CONCLUSIONS AND FUTURE WORK

In this paper we report the results of experiment during which the comprehensibility and simplicity of different adaptation logics representation formats were tested and analyzed. Results outlined an important advantage of using business rules as adaptation means for adaptive enterprise applications. Comparing to software code and adaptation models, rules demonstrate an ability for fast and simple adaptation management process since

rules are expressed in natural language form thus are easy understandable for humans without any specific IT knowledge in the same time being readable and executable by software components. Nevertheless code would be more expressive than rules if complex adaptation cases should be implemented.

Despite the good characteristics of rules as adaptation means, they still lead to various difficulties such as relatively complicated development of rules execution mechanisms as well as challenges with rules run-time performance measures and assessment of the efficiency of the adaptation process.

So far in our research we have developed context aware rule based adaptation approach and tested the comprehensibility of such adaptation process. Further work will focus on rules run time performance evaluation and adaptation approach's overall evaluation regarding its effectiveness and usefulness.

References

[1] S. Kell, "A Survey of Practical Software Adaptation Techniques", Journal of Universal Computer Science, vol 14, issue 13, 2008. DOI 10.3217/jucs-014-13-2110

[2] F. D. Macias-Escriva, R. Haber, R. Toro and V. Hernandez, "Self-adaptive systems: A survey of current approaches, research challenges and applications", Expert Systems with Applications, 40(18), p. 7267-7279, 2013

[3] D. Te'eni, R. Feldman. "Performance and satisfaction in adaptive websites: an experiment on searches within a task-adapted website", Journal of the Association of Information Systems, vol. 2, Article 3, 2001.

[4] P. K. McKinley, S. M. Sadjadi, E. P. Kasten and B. K.C. Cheng, "Composing Adaptive Software", Computer, vol. 37, Issue 7, 2004, DOI: 10.1109/MC.2004.48.

[5] L. Jokste, "Integration of Business Rules and Model Driven Development", Short Paper Proceedings of the 6th IFIP WG 8.1 Working Conference on the Practice of Enterprise Modeling (PoEM 2013), vol. 1023, 2013.

[6] A. Uhl, "Model-Driven Development in the Enterprise", IEEE Software, 2008

[7] D. A. Partridge, K. M. Hussain, K. M. Hussain, "Knowledge Based Information Systems". Mc-Graw Hill book comp., 1994

[8] A. Abraham, "Rule-based Expert Systems", Handbook of Measuring System Design, edited by Peter H. Sydenham and Richard Thorn, 2005.

[9] G. Ge, E. J. Whitehead, Jr. (Advisor), "Automatic Generation of Rule-based Software Configuration Management Systems", Proceedings of 27th International Conference on Software Engineering, 2005. L. Jokste and J. Grabis "Rule Based Adaptation: Literature Review", Environment. Technology. Resources: Proceedings of the 11th International Scientific and Practical Conference, vol.2, 2017

[10] T. Zhao, H. Zhao, W. Zhang, and Z. Jin, "User Preference Based Autonomic Generation of Self-Adaptive Rules", Internet ware 2014 Proceedings of the 6th Asia-Pacific Symposium on Internet ware, p. 25-34, 2014.

[11] T. Tran, P. Cimiano and A. Ankolekar, "Rules for an Ontology based Approach to Adaptation", Proceedings of the 1st International Workshop on Semantic Media Adaptation and Personalization, 2006.

[12] J. Rubart, "Semantic Adaptation of Business Information Systems using Human-Centered Business Rule Engines", Proceedings of 10th International Conference on Semantic Computing, p. 187-193, 2016

[13] V. Zacharias, "Development and Verification of Rule Based Systems – a Survey of Developers", Lecture notes of international workshop of Rules and Rule Markup Languages of Semantic Web, 2008. https://doi.org/10.1007/978-3-540-88808-6_4.

[14] M. Zeeshan and S. A. Khan, "A novel algorithm for link adaptation using fuzzy rule based system for wideband networking waveform of SDR", International Journal of Electronics and Communications (AEÜ), vol. 69, p. 1366- 1373, 2015.

[15] L. Jokste and J. Grabis, "Context-Aware Adaption of Software Entities using Rules", ICEIS 2017: Proceedings of the 19th International Conference on Enterprise Information Systems, vol. 3, SciTePress, 2017.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 70-73

Process-Event Approach for Operational Risk Estimation

Vasily Karasev
Institute of Problems of Mechanical Engineering of Russian Academy of Sciences
St. Petersburg, Russian Fereration
inorisk@gmail.com

Ekaterina Karaseva
State University of Aerospace Instrumentation
St. Petersburg, Russian Ferderation
matatkakate@gmail.com

***Abstract* - Authors propose a new process-event approach for quantitative estimation of operational risk in a bank and calculation the amount of economic capital in dynamics. The proposed approach, according to the Basel II Capital Accord, belongs to the category of "advanced methods". Operational risk is not financial risk and appeared in unfavorable events mainly. A number of business processes are performed in banking activity. Every business process contains a set of routines (operations), which can be interrupted by operational risk events with certain losses. The main idea of the approach is to describe processes as chains of casual events instead of a traditional graphic description as diagrams. Authors introduce new concepts: an elementary process event, a chain of process events, a time diagram of enterprise's event flow, and build logical and probabilistic risk models. Methods and formulas for calculation the current and integrated operational risk in dynamics, the amount of economic capital, upper and lower limits of reservation, are given. The value of integrated operational risk can be used as a rating of the current reliability of the bank. The paper outlines the advantages and disadvantages of proposed process-event approach. Research results can be implemented as analytical tool in risk management technology, "process mining" technology and in bank intelligent management systems.**

Keywords - process, event, logic, probability, risk, management, event chain, event flow.

I. INTRODUCTION

Operational risk is connected with current functioning of bank or company, fraud risk and external events. In most publications the operational risk is considered as one of main risks in commercial bank and there is standard operational risk definition, given in Basel II Capital Accord [1].

In global banking practice, operational risk management is a key and paramount task. Operational risk penetrates all aspects of other possible risks and, more precisely, it is interrelated with all other types of risk (for example, market risk, credit risks and liquidity risks), making them more difficult for estimation. In the absence of operational risk failures, all other types of risk are significantly less important.

In fact, operational risk is not bank phenomenon only. Operational risk is a complex systemic phenomenon in all existing socio-economic systems. This is a non-financial risk and it influences on activity of any economic subject - a commercial enterprise, an investment company, an industrial enterprise, a social institution, executive and legislative authorities, army.

For banks, the Basel II agreement defines a choice of three approaches for operational risk estimation and capital reservation: the base indicator method, the standardized method, and the advanced method [1]. But for other economic subjects there are no standards and regulatory documents, despite the importance and relevance of the problem.

In this paper we present a new approach for assessment, analysis and management of operational risk for any economic subject.

II. MAIN DEFINITIONS

Based on the generally accepted definitions of operational risk, we emphasize the eventual nature of this phenomenon - operational risk is identificated primarily in the events: mistakes of employees of the enterprise, failures of information systems, political events, terroristic attacks, natural disasters. Since events have a different nature, there are difficulties in the formalized description and modeling of operational risk. The problem of assessment and management of operational risk is one of the most difficult problems in risk management [2-4].

There are no standard, regulatory documents and clear methods for operational risk management at commercial enterprises, so, we suppose the actual problem to develop a new methodology based on a process management approach [5]. Also, the assessment and management of operational risk by analyzing business processes and identifying weak points in them is described in detail in [6], but in relation to a commercial bank.

In the process approach, whole activity of bank with used resources is considered as a set of processes. Each process consists of a sequence of actions. Often, the output of one process directly forms the input of the next one. The advantage of the process approach is the continuity of management, which is provided at the junction of individual processes within their system, as

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4041
© 2019 Vasily Karasev, Ekaterina Karaseva.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

well as with various combinations of processes and their interactions.

The distribution of business process management information systems (BPM), workflow management systems (WFM), ERP II systems and case handling systems provides possibility to record individual events and keep event logs during the operation of the enterprise (bank). These datasets (or protocols) are very useful for analysis and making management decisions and they are applied in other areas, for example, in transport - many cars and public transport are equipped by information system with event log.

There is a task to "extract" processes from the event logs of the information systems in organizations and enterprises. This task is especially relevant in cases where a formal description or model of the system is missing or has low compliance with the actual behavior of the system, that is, in 98% cases. To solve this problem, Process Mining technologies [7] are being actively developed, which allow to extract a process model from the event log, check its consistency on the statistics of subsequent events, and improve the model if there is a discrepancy between the model and process data.

To construct models of business processes, there are several sets of graphic symbols (notations). At the same time, the description and analysis of business processes in bank based on notations only have some disadvantages:

- high-level processes are described simply and the value of such work is the highest, but the description of detailed processes at a lower level requires a heavy work with a low value. Reduction the level down (almost up to individual employees) leads to excessive complexity of the scheme, loss of visibility and control, while expenses at lower levels are often negligible. That is, 5 - 7 levels of processes is optimal;

- the fragmentation in description does not allow to get a complete picture of the bank as a whole system, to display its functional hierarchy, the structure of resources, to distinguish time stages strictly. Bank model in a process approach consists of a large number of local schemes reflecting various aspects of bank's activity from different levels and viewpoints, but interconnected through an intuitive image of the whole in the developer's mind only [8];

- absence of effective methods for analysis and optimization of business processes;

- the inability to quantitatively calculate the value of operational risk and the amount of reserved economic capital.

In [8], a new subject-event approach to the description of business systems is described. The enterprise (bank), as a system (or integrity, distributed in time), should be described as a stream of events: if we fix an exhaustive set of all events occurring in enterprise, then we get a full description of enterprise as a complete system. The basics of such eventual approach to the description of phenomena and processes were stated in the works of the famous philosophers L. Wittgenstein [9] and B. Russell [10].

The basis of our proposed *process-event approach* is the concept of separate chains of significant elementary events from the business processes of the bank. To determine the numerical value of the operational risk in the bank, we will use the logical-probabilistic method applied by the author to solve the problem of assessment, analysis and management of operational risk in accordance with the Basel II standard [11, 12].

We introduce several basic definitions.

Process - an orderly interconnected sequence of events and actions (functions, procedures, operations) performed by employees and departments of the organization to obtain the desired final result (goal achievement, problem solving, program implementation, service provision).

Event is the successful completion of action (function, procedure, operation), or successful state (permissible values of process parameters at a certain time period). This event is random. Since any function, procedure or operation may be incomplete during allotted time (due to equipment failure, power failure, personnel errors and other factors), this event may occur or may not occur with a certain probability.

An **elementary event** is a random event, which is impossible or impractical to split further into events (initiating events in relation to the primary event). An elementary event corresponds to the simplest indivisible action (operation) of the process.

How to select elementary events? This question is not simple and very important. In work [13], for example, a general principle for distinguishing events was formulated: an event is considered as elementary if at least one of three conditions is observed:

1. The event is executed or registered by one of the subjects of the system;

2. The event is executed with one of the objects of the system;

3. The fact of the presence (or absence) of an event is a condition for this event to be executed by another subject or the same subject but with another object.

Thus, the event is determined by specifying the object (s) with which it occurs, and the subject who performed or registered it. The subject can be a specific person, role, team, as well as software, sensors, etc. Events that affect the behavior of the system but are not associated with any of its subjects, should be attributed to the absolute subject. Because any event in the system is determined as events only under the conditions of its connection with a specific subject, the author describes the approach not as event-driven, but as subject-event.

A **process event chain** is a sequence of random events related to the successful completion of an action (functions, procedures, operations). In our interpretation, the number and sequence of events in the chain corresponds to the number and sequence of actions in the business process.

Thus, a business process can be considered as a sequence of non-random actions, but the result of each action is a random event.

Business processes can be executed both sequentially (when the output of one business process is the input for another), and in parallel.

Set of processes in bank can be defined by set of process event chains, so the state of the bank at time t

can be described by a set of current unrealized events St from parallel processes, taking into account the logical connections between them.

Events in some processes will have a high probability of occurrence, i.e. be almost reliable (for example, the event "compilation of annual reporting documents"), while in others the probability of their appearance will be significantly lower, for example, in business processes of management of investment projects. This is taken into account when probabilities of events are determined.

An **active process** is a process, where there is at least one operation, procedure, work, or function is performed at a given time.

Operational risk - the probability of non-appearance of elementary event in the chain. This is a slightly varying value, within the boundaries of a given time interval (before updating statistical data, it is advisable to consider this value as constant).

The **operational risk of a bank** is the integrated value of the operational risk of the entire bank, i.e. the logical sum of operational risk values at all parallel processes. This value is changed in dynamics during the entire time of bank's functioning.

III. THE METHOD

The totality of all the processes EP_i of a bank as the events chains forms a non-empty finite set B, $EP_i \in B$, $i = 1, ..., k$, where k is the total number of distinguished processes in the bank. Some processes from this set are performed in parallel (in time), so in every time t on the cut of the time diagram we have set of events S_t, consisting of currently unrealized at the time t, events chains EP_i, $i = 1, ..., k$.

Let us assign logical variables X_m, $m = 1, ..., n$ to events from the set S_t, where n is the number of bank's processes that are active at the time t. The logical variable X_m can take the value 1 (event occurred) with probability $P(X_m = 1)$, or the value 0 (event not occured) with probability $Q(X_m = 0) = 1 - P(X_m = 1)$. At time t, any event X_m, or several events at once, or all events from the set S_t may not occur.

The logical function of the integrated operational risk is written in the following form:

$$X_{OR} = \overline{X}_1 \vee \overline{X}_2 \vee ... \vee \overline{X}_n , \qquad (1)$$

Then the total integrated operational risk $P(X_{OR})$ can be calculated from the expression (probability function) [14]:

$$P(X_{OR}) = 1 - P(X_1)P(X_2) \ldots P(X_n) \ . \qquad (2)$$

The value of $P(X_{OR})$ is the probability of non-realization of any event in the set S_t. Now, if we know the value $P(X_{OR})$, we can calculate the size of the reserved economic capital from the expression:

$$RC = P(X_{OR})\, L , \qquad (3)$$

where L is the average loss due to operational risk events for the previous three years[1].

During the working day (trading day) the value of $P(X_{OR})$ is constantly changing. In this case, it is impractical to calculate this at each time moment.

In the traditional process approach, detailed diagrams of business processes (often in different notations) are compiled and one-time simulation is performed using special software, or diagrams are analyzed "manually". According to the results of simulation and analysis, weak and inefficient connections in the organization of processes are identified and measures are being taken to eliminate deficiencies. After some time, the simulation is performed again, as a rule, this is preceded by a large changes in the bank's activity due to the introduction of new technologies, improvement of the organizational structure, changes in the economic and political situation.

In the process-event approach, we primarily "extract" processes and form event chains. At the same time, we can also use business process diagrams, which were constructed earlier. Then, we cut off the fundamentally impossible sets of S_t (a combination of incompatible events) in chains of parallel and sequential business processes. Further, all possible S_t sets are generated and the $P(X_{OR})$ values are calculated in several most critical S_t sets (where $P(X_{OR})$ is the largest).

The value of $P(X_{OR})$ can be used as a current indicator of the reliability of bank, since this is essentially a convolution of other risks due to the fact that all risks are included in probabilities of elementary events in processes.

By analogy with the method, described in [12], we introduce the concepts of the upper and lower limits of operational risk, respectively, the upper and lower limits of capital reservation. The upper limit $P(X_{OR})_{up}$ is determined by the maximum value of operational risk obtained on the S_t sets, the lower limit is determined by the minimum value $P(X_{OR})_{down}$.

The calculation of the upper and lower limits of capital reservations is performed according to the formulas:

$$RC_{up} = P(X_{OR})_{up}\, L , \qquad RC_{down} = P(X_{OR})_{down}\, L , \qquad (4)$$

The choice of the reservation value is determined by the current economic and political situation, values of main market indices, the experience of the risk manager and the completeness of the information available to him (her).

IV. DISCUSSION

The process-event approach is further evolution of the traditional process approach to enterprise management. The implementation of event description for processes

1 This period is determined by requirements in Basel II Capital Accord for banks. In fact, the period can be determined by the risk manager, depending on the frequency of fixed cases of operational risk.

allow us to solve, in practice, the complex task of the quantitative assessment of operational risk which is too difficult to formalize.

At the current stage of development of the new technique, several advantages of the process-event approach have already become apparent:

1. To describe processes of bank's functioning, we need less variables than in the traditional process approach, using notations.

2. Modeling business processes through chains of events allows us to calculate the operational risk value on realization of events at time t, identify the most dangerous combinations of events and calculate the amount of reserved economic capital.

3. The process-event approach allows us to calculate the operational risk not only as a probability of unfavourable event among set of internal events of processes, but also taking into account external events (changes in legislation, natural disaster, fraud, armed robbery, terroristic acts etc.). We can perform this by adding corresponding logical variables to formula (1). However, it should be assumed that when we determine probabilities $P(X_{i,m})$, $m = 1, ..., n$ according to the event protocol (statistical data), the influence of external factors is already taken into account in the values of these probabilities.

4. The process-event approach uses simple and transparent formulas to calculate the operational risk and the amount of economic capital, reserved to cover possible losses. But this advantage is takes place in application the process-event approach to solution the problem of quantitative assessment of operational risk only. In other tasks (for example, optimization of organizational structure and bank management procedures) this advantage is not obvious.

The process-event approach also has disadvantages:

1. The absence of clear recommendations about the selection of events. We recommend to select elementary events that have a fundamental impact on the logics of the process and can be observed and recorded.

2. Difficulties may arise in determination probabilities of events $P(X_{i,m})$. In the absence of statistics, expert methods have to be applied, which reduce the reliability of results.

3. The process-event approach requires a certain level of knowledge and competence of the manager (specialist). The manager needs some preliminary practical experience in the decomposition of processes and the identification of elementary events.

V. CONCLUSION

The idea of event-process approach is new, therefore, there are many questions that require more detailed study.

Existing operational risk management technologies often overlook the main fact: operational risk is non-financial risk and realize in separate events.

Authors took this feature into account and proposed a process-event approach for quantitative assessment of operational risk, introduced basic definitions and provided logical and probabilistic models of operational risk in banks.

REFERENCES

[1] International Convergence of Capital Measurement and Capital Standards: A Revised Framework, Bank for International Settlements Press & Communications CH-4002 Basel, Switzerland, 2004, 251 p. [Online]. Available: https://www.bis.org/publ/bcbs107.pdf [Accessed: Feb. 20, 2019].

[2] Kaspereit T., Lopatta K., Pakhchanyan S., Prokop J. Systemic operational risk / Journal of Risk Finance. – Vol. 18, Issue 3. – 2017. – pp. 252-262.

[3] McKim V.L. Operational risk assessment / Journal of Business Continuity & Emergency Planning. – Vol. 10, Issue 4. – Summer, 2017. – pp. 339-352.

[4] Panjer H.H. Operational Risk: Modeling Analytics. – Wiley, 2006. – 448 p.

[5] Reengineering the Corporation: A Manifesto for Business Revolution, Michael Hammer and James Champy, Nicholas Brealey Publishing. London,1993 (Harper Collins in U.S.A), 223 p.

[6] Сазыкин Б.В. Управление операционным риском в коммерческом банке. М.: Вершина, 2008. – 272 с.

[7] Wil M.P. van der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes. – Springer-Verlag Berlin Heidelberg, 2011. – 352 p.

[8] Болдачев А., Шумаков А. О принципиально новом, субъектно-событийном подходе к моделированию бизнес-систем. [Online]. Available: http://www.rsci.ru/rt/researches/236426.php, [Accessed: Jul. 15, 2018]

[9] Wittgenstein L. Tractatus Logico-Philosophicus. New York: Harcourt, Brace, 1922, 75 p.

[10] Russel B. The Philosophy of Logical Atomism. London and New York, Routledge Classics, 2010, 162 p. [Online]. Available: https://sites.ualberta.ca/~francisp/NewPhil448/RussellPhilLogicalAtomismPears.pdf [Accessed: Feb. 25, 2019].

[11] Author, 2012.

[12] Author, 2016.

[13] Болдачев А. Субъектно-событийный подход к моделированию сложных систем. [Online]. Available: https://habrahabr.ru/post/256509/, [Accessed: Jul. 15, 2018].

[14] Solozhentsev E.D. The Management of Socioeconomic Safety / Solozhentsev E.D. –Cambridge Scholars Publishing, 2017. – 255 p.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 74-78

Analysis of Research Trends in the Fieldof Mechanical Engineering

Adrian Knapczyk
Department of Mechanical Engineering and AgrophysicsFaculty o Production and Power Engineering,
University of Agriculture
Krakow, Poland
adrian.knapczyk@urk.edu.pl

Sławomir Francik
Department of Mechanical Engineering and Agrophysics
Faculty o Production and Power Engineering,
University of Agriculture
Krakow, Poland
slawomir.francik@urk.edu.pl

***Abstract*—The work includes a bibliometric analysis of the main topics and research trends within the discipline of Mechanical Engineering. On the basis of analysis of data from the Scopus database, the Applied Mechanics Reviews magazine was chosen, which is the only one conducting scientific activity exclusively in the field of Mechanical Engineering. In the years of analysis (2002-2017), it was pointed out that the main research topics are optimization issues, material engineering and mechanics.**

Keywords—mechanical engineering, scientometrics, research trends, bibliometric analysis, literature review.

I. Introduction

Mechanical engineering is a complex and very broad field of scientific research that uses the achievements of other fields (mechanics, design and many others). The division and domains involved in mechanical engineering is well known [1]develop, manufacture, operate, and use mechanical artifacts.\"--Publisher description. Fundamentals of Mechanical Engineering -- Introduction to Mathematics for Mechanical Engineering -- Complex Analysis -- Differential Equations -- Laplace Transformation -- Fourier Analysis -- Linear Algebra -- Mechanics -- Statics of Rigid Bodies -- Dynamics -- Applications in Mechanical Engineering -- Materials Science and Engineering -- Atomic Structure and Microstructure -- Microstructure Characterization -- Mechanical Properties -- Physical Properties -- Nondestructive Inspection (NDI. However, it is necessary to indicate current themes and research trends. Such knowledge is needed to plan future research projects. Keeping track of current trends increases the effectiveness of research. It is necessary to define new global research trends that have evolved in this discipline to identify the main topics of research, discoveries and the global knowledge network. Bibliometric analysis is a statistical analysis, which can serve setting topics and research trends [2] etc. This method is characterised by a possibility of conducting extended quantitative analyses in an objectivised way. Bibliometric analyses are based on available, coherent, objectivised data. This tool is widely used especially in setting trends and research topics in dedicated areas [3].

The aim of this work was to identify trends and main research topics in the field of mechanical engineering.

II. Materials and Methods

The research was carried out in six stages, using elements of bibliometric techniques:

1.a. Creation of a set of journals based on searching for indexed items in the Scopus database for search: Subject area: Engineering => Mechanical Engineering.

1.b. Separating 10% of the top journals, do the subject area: Mechanical Engineering.

1.c. Separating journals, which in the subject area have only Mechanical Engineering.

1.d. Creation of set of publications based on separating journals. The search was carried out from 2002 to 2017 documents in English.

1.e. Construction and analysis of term maps (VOSviewer software) in during the period considered.

1.f. Quantitative analysis of the set of documents created in the aspect of: number of publications and number of citations, major countries.

VOS viewer is a free program that is used by researchers for bibliometric analysis, including analysis of research trends, but also to visualize selected areas of knowledge [4]–[7]. In this work, the program was used to create a map of terms: years of publication, intensity of quoting.

The key words with the highest occurrance frequency are displayed in a bigger cicles. The words with the lower frequency of occurance, on the other hand, have smaller cicles and font.

III. Results and discussion

There are 585 sources in the Scopus database in the

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4170
© 2019 Adrian Knapczyk, Sławomir Francik.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

field of Mechanical Engineering, including 508 journals. The next step was the separation of 10% of scientific journals with the highest CiteScore index in 2017, which included 72 journals. The ranking of the top 10 journals with selected parameters has been presented in Table 1. It can be noticed that there are also journals in this group that also deal with the subject matter of materials engineering and energy.

Journals analysis was carried out in terms of areas of activity. Finally, the journal Applied Mechanics Reviews has been selected for further analysis, which is the only scientific activity conducted exclusively in the field of Mechanical Engineering.

Applied Mechanics Reviews (AMR) is an international journal that raises the subject of applied mechanics and engineering (eg fluid mechanics, dynamics and vibration, etc.) [8].

Applied Mechanics Reviews in the analyzed period (2002-2017) publishes an average of 20 publications per year. In 2012, the magazine did not have publications indexed in the Scopus database. The main research institutions publishing are: The Royal Institute of Technology KTH (10 documents), Texas A and M University (10), University of Michigan (8 documents), University of Illinois at Urbana-Champaign (7 documents) and Xi'an Jiaotong University (7 documents). Research centers come from such countries as: USA (137 documents), China (34 documents), UK (25 documents), Italy (23 documents) and Canada (20 documents). Most publications were published by the following authors: Dankowicz H. (6 documents), Rega G. (4 documents), Alfonsi G. (3 documents), Hanifi A. (3 documents), Karamanos S. A. (3 documents).

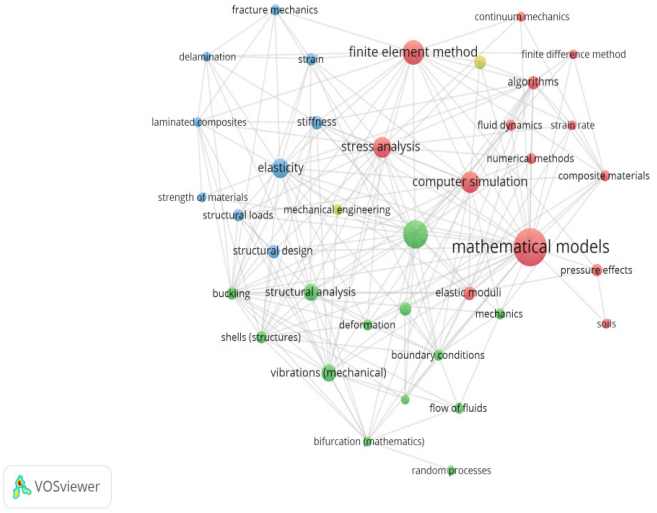
As part of the work, bibliometric data were collected in the analyzed period and a quantitative analysis was carried out for all key words (author and additional), and then they were summarized in separate periods: 2002-2006, 2007-2011, 2013-2017 (Fig.1). In the first period (Fig. 1a) one cluster of terms can be distinguish: "mathematical models", "problem solving", "structural analysis", "finite element method", "elastic module". This proves that scientists are very interested in optimization issues.

In the second period (Fig. 1b), you can see a much larger fragmentation of the clusters of keywords. The first cluster of words is: "elasticity", "mechanics", "materials properties", "programming theory", "paltes (structural components)", "structural analysis". The second cluster includes: "finite element method", "optimization", "dynamics", "models", "laminates". The third cluster of key words is: "failure analysis", "mathematical models", "fiber reinforced plastics", "composite materials". This proves that scientists are interested in issues related to material engineering and, as in the previous period, optimization issues.

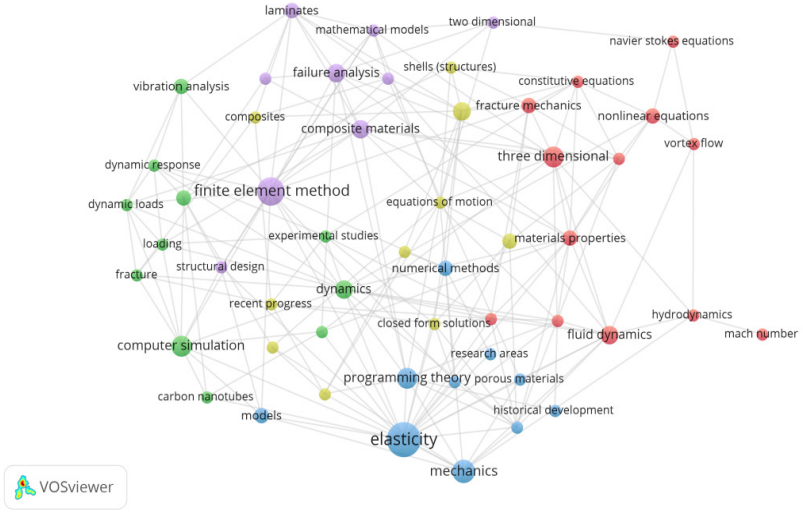
TABLE I. 10 SCIENTIFIC JOURNALS BY CITESCORE INDEX IN 2017 IN THE AREA OF MECHANICAL ENGINEERING

Source title	Rank 2017	CiteScore 2017	Publisher
Materials Science and Engineering: R: Reports	1/554	31.32	Elsevier
Nature Materials	2/554	25.47	Springer Nature
Advanced Materials	3/554	21.1	Wiley-Blackwell
Nano Letters	4/554	13.07	American chemical Society
International Materials Reviews	5/554	12.81	Taylor & Francis
Applied Energy	6/554	8.44	Elsevier
Applied Mechanics Reviews	7/554	7.62	ASME
Progress in Aerospace Sciences	8/554	7.25	Elsevier
Materials Today	9/554	6.8	Elsevier
Desalination	10/554	6.41	Elsevier

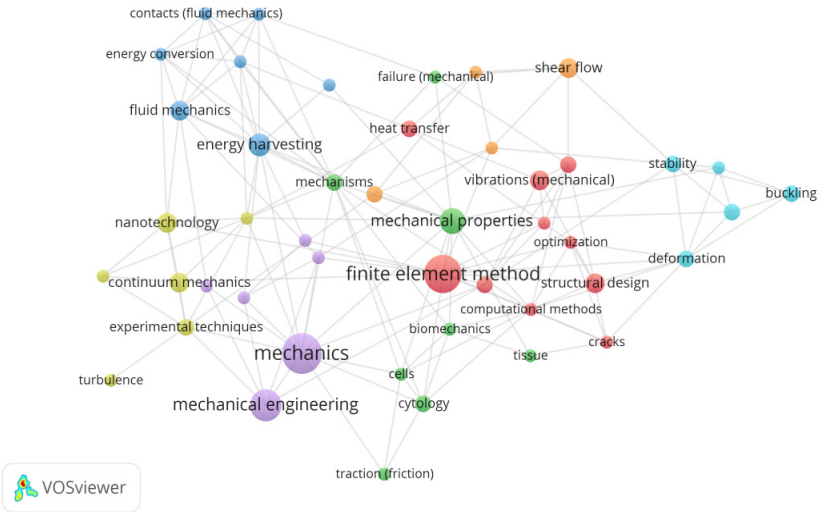
In the third period (Fig. 1c), as in the second period, there is a large fragmentation of clusters of keywords. The first group of entries can be distinguished: "mechanics", "driling platforms", "mechanical engineering", "linear systems", "future research directions". The second group includes only the term: "finite element method", while the last group includes: "design", "vibration analysis", "vibrations (mechanical)", engineering applications "and" shear flow ". It can be seen that the main research topics are: optimization, material engineering, design issues and mechanics (vibration analysis). In all separate periods the main groups of issues are material engineering (structure analysis, material properties research, etc.) and optimization problems (finite element method, computer simulations and others).



a)



b)



c)

Fig. 1. Set of term maps for the analysed periods: a) years 2002-2006, b) years 2007-2011, c) years 2013-2017.

Table 2 presents the most frequently quoted publications throughout the period of publishing the journal. Byun and Schere [9] in their publication, computational algorithms and functionality of the Community Multiscale Air Quality (CMAQ) system. This system is used to model the atmosphere of air including chemical and physical processes. Bert and Malik [10] use the differential quadrature method in computational mechanics. Qian et. al. [11] in their publication investigated mechanical properties (Young's modulus, bending test, tensile strength and compression) and presented examples of applications for carbon nanotubes. Raupach et. al. [12] experimental and theoretical knowledge of rough-wall turbulent boundary layers, drawing from both laboratory.

TABLE II. The most cited publications in journal Applied Mechanics Reviews

Author	Year	Title	Citation
[9]	2006	Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system	1153
[10]	1997	Differential quadrature method in computational mechanics: A review	950
[11]	2002	Mechanics of carbon nanotubes	853
[12]	1991	Rough-wall turbulent boundary layers	767

Examples of research topics in 2017:

- overview of the state of the art on the use of mechanical instabilities in solids [13],

- a review of experimental methods for determining the contact surface of nanometric particles [14], [15],

- a review of research on the use quadrature element method (QEM) in science and engineering [16],

- analysis of physical processes in the operation of ventilation and scaling capabilities, and the impact of these processes on the design of installations [17],

- a review of research on the analysis of the mechanical properties of lead-free solders [18],

- research review on contact modeling and the impact of this contact on material properties [19],

- review of research on mechanical properties and the applicability of crystalline nanowires [20],

- Morton Effect (ME) - symptoms, causes, theory of predictions and solutions [21],

- review of theoretical models, numerical and experimental tests [22];

- review of research on various types of nanocarbon polymers [23],

- review of modeling research and experimental measurements of dynamic forces occurring during the walking process [24],

- review of research on applied theory of continuous mechanics with non-local elastic response [25],

- a review of research on friction in the wheel-rail system [26],

- research review and experimental research on modeling of selected parameters during flat rolling [27].

IV. Conclusions

1. The analysis of journals in the Scopus database has been demonstrated by the Applied Mechanics Reviews magazine, which as the only conducts scientific activities exclusively in the field of Mechanical Engineering.

2. In the years 2002-2017, in the journal Applied Mechanics Reviews, they mainly concerned optimization, selected problems of mechanical engineering, mechanics and .advanced material engineering

3. The paper shows that the use of VOSviewer program for bibliometric analysis focused on identifying.

Acknowledgement

This research was financed by the Ministry of Science and Higher Education of the Republic of Poland (statutory activities DS-3600/WIPiE/2018, Faculty of Production and Power Engineering, University of Agriculture in Krakow).

References

[1] K.-H. Grote and E. K. Antonsson, *Springer handbook of mechanical engineering*. Springer, 2009.

[2] S. Francik *et al.*, "Bibliometric Analysis of Multiple Critteria Decision Making in Agriculture," *Tech. Sci.*, vol. 20, no. 1, pp. 17–30, 2017.

[3] A. Knapczyk, S. Francik, N. Pedryc, and T. Hebda, "Bibliometric analysis of research trends in engineering for rural development," in *17th International Scientific Conference Engineering for Rural Development*, 2018, pp. 700–707.

[4] F. T. Gizzi, "Worldwide trends in research on the San Andreas Fault System," *Arab. J. Geosci.*, vol. 8, no. 12, pp. 10893–10909, Dec. 2015.

[5] O. Tabatabaei-Malazy, R. Atlasi, B. Larijani, and M. Abdollahi, "Trends in publication on evidence-based antioxidative herbal medicines in management of diabetic nephropathy," *J. Diabetes Metab. Disord.*, vol. 15, no. 1, p. 1, Dec. 2015.

[6] S. H. Zyoud, W. S. Waring, W. M. Sweileh, and S. W. Al-Jabi, "Global Research Trends in Lithium Toxicity from 1913 to 2015: A Bibliometric Analysis," *Basic Clin. Pharmacol. Toxicol.*, vol. 121, no. 1, pp. 67–73, Jul. 2017.

[7] A. Watad *et al.*, "Is autoimmunology a discipline of its own? A big data-based bibliometric and scientometric analyses," *Autoimmunity*, vol. 50, no. 4, pp. 269–274, May 2017.

[8] American Society of Mechanical Engineers., "Applied mechanics reviews.," [Online]. Available: http://appliedmechanicsreviews.asmedigitalcollection.asme.org/journal.aspx. [Accessed: 04-Mar-2019].

[9] D. Byun and K. L. Schere, "Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System," *Appl. Mech. Rev.*, vol. 59, no. 2, p. 51, 2006.

[10] C. W. Bert and M. Malik, "Differential Quadrature Method in Computational Mechanics: A Review," *Appl. Mech. Rev.*, vol. 50, no. 9, p. 552, 1997.

[11] D. Qian, G. J. Wagner, W. K. Liu, M.-F. Yu, and R. S. Ruoff, "Mechanics of carbon nanotubes," *Appl. Mech. Rev.*, vol. 55, no. 6, p. 495, 2002.

[12] M. R. Raupach, R. A. Antonia, and S. Rajagopalan, "Rough-

Wall Turbulent Boundary Layers," *Appl. Mech. Rev.*, vol. 44, no. 1, p. 1, 1991.

[13] D. M. Kochmann and K. Bertoldi, "Exploiting Microstructural Instabilities in Solids and Structures: From Metamaterials to Structural Transitions," *Appl. Mech. Rev.*, vol. 69, no. 5, p. 050801, Oct. 2017.

[14] T. D. B. Jacobs and A. Martini, "Measuring and Understanding Contact Area at the Nanoscale: A Review," *Appl. Mech. Rev.*, vol. 69, no. 6, p. 060802, Nov. 2017.

[15] M. Ciavarella and A. Papangelo, "Discussion of 'Measuring and Understanding Contact Area at the Nanoscale: A Review' (Jacobs, T. D. B., and Ashlie Martini, A., 2017, ASME Appl. Mech. Rev., 69 (6), p. 060802)," *Appl. Mech. Rev.*, vol. 69, no. 6, p. 065502, Nov. 2017.

[16] X. Wang, Z. Yuan, and C. Jin, "Weak Form Quadrature Element Method and Its Applications in Science and Engineering: A State-of-the-Art Review," *Appl. Mech. Rev.*, vol. 69, no. 3, p. 030801, May 2017.

[17] Y. L. Young, C. M. Harwood, F. Miguel Montero, J. C. Ward, and S. L. Ceccio, "Ventilation of Lifting Bodies: Review of the Physics and Discussion of Scaling Effects," *Appl. Mech. Rev.*, vol. 69, no. 1, p. 010801, Jan. 2017.

[18] Y. Yao, X. Long, and L. M. Keer, "A Review of Recent Research on the Mechanical Behavior of Lead-Free Solders," *Appl. Mech. Rev.*, vol. 69, no. 4, p. 040802, Aug. 2017.

[19] H. Ghaednia, X. Wang, S. Saha, Y. Xu, A. Sharma, and R. L. Jackson, "A Review of Elastic–Plastic Contact Mechanics," *Appl. Mech. Rev.*, vol. 69, no. 6, p. 060804, Nov. 2017.

[20] Y. Zhu, "Mechanics of Crystalline Nanowires: An Experimental Perspective," *Appl. Mech. Rev.*, vol. 69, no. 1, p. 010802, Jan. 2017.

[21] X. Tong, A. Palazzolo, and J. Suh, "A Review of the Rotordynamic Thermally Induced Synchronous Instability (Morton) Effect," *Appl. Mech. Rev.*, vol. 69, no. 6, p. 060801, Oct. 2017.

[22] J.-J. Shu, J. Bin Melvin Teo, and W. Kong Chan, "Fluid Velocity Slip and Temperature Jump at a Solid Surface," *Appl. Mech. Rev.*, vol. 69, no. 2, p. 020801, Mar. 2017.

[23] S. Roy, J. Ryan, S. Webster, and D. Nepal, "A Review of In Situ Mechanical Characterization of Polymer Nanocomposites: Prospect and Challenges," *Appl. Mech. Rev.*, vol. 69, no. 5, p. 050802, Nov. 2017.

[24] A. Younis, O. Avci, M. Hussein, B. Davis, and P. Reynolds, "Dynamic Forces Induced by a Single Pedestrian: A Literature Review," *Appl. Mech. Rev.*, vol. 69, no. 2, p. 020802, Apr. 2017.

[25] A. R. Srinivasa and J. N. Reddy, "An Overview of Theories of Continuum Mechanics With Nonlocal Elastic Response and a General Framework for Conservative and Dissipative Systems," *Appl. Mech. Rev.*, vol. 69, no. 3, p. 030802, Jun. 2017.

[26] U. Olofsson and Y. Lyu, "Open System Tribology in the Wheel–Rail Contact—A Literature Review," *Appl. Mech. Rev.*, vol. 69, no. 6, p. 060803, Nov. 2017.

[27] A. Steinboeck, A. Ettl, and A. Kugi, "Dynamical Models of the Camber and the Lateral Position in Flat Rolling," *Appl. Mech. Rev.*, vol. 69, no. 4, p. 040801, Aug. 2017.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 79-83

Comparison of Algorithms for Construction Detection using Airborne Laser Scanning and nDSM Classification

Sergejs Kodors
Rezekne Academy of Technologies
Rezekne, Latvia
sergejs.kodors@rta.lv

***Abstract* — Traditional approach to classify the point cloud of airborne laser scanning is based on the processing of a normalized digital surface model (nDSM), when ground facilities are detected and classified. The main feature to detect a ground facility is height difference between adjacent points. The simplest method to extract a ground facility is region-growing algorithm, which applies threshold to identify the connection between two points. Region growing algorithm is working with the constant value of height difference. Therefore, it is not applicable due to diverse conditions of earth surface, when height difference must be defined for each region separately. As result, researchers propose hierarchical, statistical and cluster methods to solve this problem. The study goal is to compare four algorithms to generate nDSM: region growing, progressive morphological filter, adaptive TIN surfaces and graph-cut. The experiment is divided into two stages: 1) to calculate the number of detected and lost buildings in nDSM; 2) to measure the classification accuracy of extracted shapes. The experiment results have showed that progressive morphological filter and graph-cut provides the minimal loss of buildings (only 1%). The most effective algorithm for ground facility detection is the graph-cut (total accuracy 0.95, Cohen's Kappa 0.89, F_1 score 0.93).**

Keywords — buildings, LiDAR, nDSM, remote sensing.

I. Introduction

The application of a normalized digital surface model (nDSM) is common approach to detect ground facilities in 3D point cloud obtained by airborne laser scanning (ALS). The generation of nDSM is related with the construction of a digital elevation model (DEM), when the LiDAR points of ground facilities are extracted from all dataset (see Eq.1):

$$\mathrm{nDSM}=\mathrm{DSM}-\mathrm{DEM}, \quad (1)$$

where *DSM* is a digital surface model or all points of dataset, *DEM* – the points, which belong to a ground, and *nDSM* – the points of surface facilities.

The construction of DEM or nDSM is actually the filtering of DSM, when the extra points are removed out from it (see Eq.2-3):

$$\mathrm{DEM}=\mathrm{DSM}-f(\mathrm{DSM}) \quad (2)$$

$$\mathrm{nDSM}=\mathrm{DSM}-f(\mathrm{DSM}) \quad (3)$$

Usually the filtering function *f*(*DSM*) is based on the height difference between adjacent points. For example, the region growing algorithm segments whole image into regions, where the pixels have difference of attributes smaller than threshold. If the processed area is sufficiently large, the largest segment should belong to the category "ground", that can be applied to obtain DEM (see Eq.2) and then – nDSM (see Eq.1).

The goal of this study is to compare four algorithms to generate nDSM: region growing, progressive morphological filter, adaptive TIN surfaces and graph-cut.

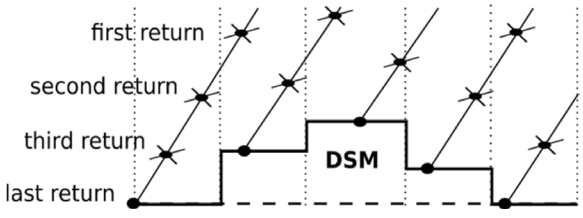
Each algorithm applies different principles to detect ground facilities using height difference, from a constant threshold to variable. The study was carried out in two parts. The detected number of ground facilities is evaluated in the first part, the correctness of extracted building shapes – in the second part.

II. LiDAR Data Preprocessing

LiDAR data is filtered and transformed before a nDSM generation. A laser pulse provides many returns in the case of vegetation, considering this fact a point cloud is filtered by the last return.

This study compares grid-based methods, when a point cloud is projected into 2D grid and segmentation algorithms are processing pixels.

In this study, the point cloud is transformed using the lowest points, when the height of the lowest point is assigned to the pixel to obtain 2D DSM (see Fig.1).



Schema of DSM generation from LiDAR data

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4032
© 2019 Sergejs Kodors.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

III. SEGMENTATION ALGORITHMS

This study compares four algorithms to generate nDSM and to detect man-made constructions. Each algorithm is selected considering its principle of height difference definition:

- region growing algorithm applies a constant threshold (see Fig.2a);
- adaptive morphological filter is a hierarchical method, which applies the *min* and *max* height of points in a processing widow (see Fig.2b);
- adaptive TIN surfaces – a hierarchical method, which considers the slope of earth and applies statistics to filter extreme points (see Fig.2c);
- graph-cut analyzes the local similarity of adjacent pixels considering the global information about clusters (see Fig.2d).

A. Region Growing

It is classical image segmentation algorithm, which uses seeds for region growing. Each iteration of the main loop provides the addition of adjacent pixels to already existing region [1], if pixels satisfy a predefined condition (threshold).

```
Function RegionGrowing( dsm )
  segments
  While( NextSeedPoint( dsm ) )
    region  NextSeedPoint( dsm )
    Do
      contour  GetAdjacentPoints( region )
      pixels  FitnessFunc(region, contour)
      region  region  pixels
    While( |pixels| > 0 )
      segments  segments  region
  End while
  Return segments
End function
```

In the case of DSM processing, next seed is the point with the minimal height among all unlabeled pixels. Adjacent points can be selected using pixel connectivity with 4 or 8 neighbors. The selected adjacent points are compared with the closest segment point, if the similarity of points is greater than a predefined threshold, they must be included into the segment. Traditionally, the similarity of two points is defined measuring the height difference between them.

B. Progressive Morphological Filter

Kilian et al. have proved, that mathematical morphology is applicable to remove buildings and trees from LiDAR data [2], however, the proposed method had difficulties with all non-ground objects of various sizes due to the fixed size of a filtering window. Therefore, Zhang et al. proposed progressive model, which increases the window sizes of morphological filters gradually [3].

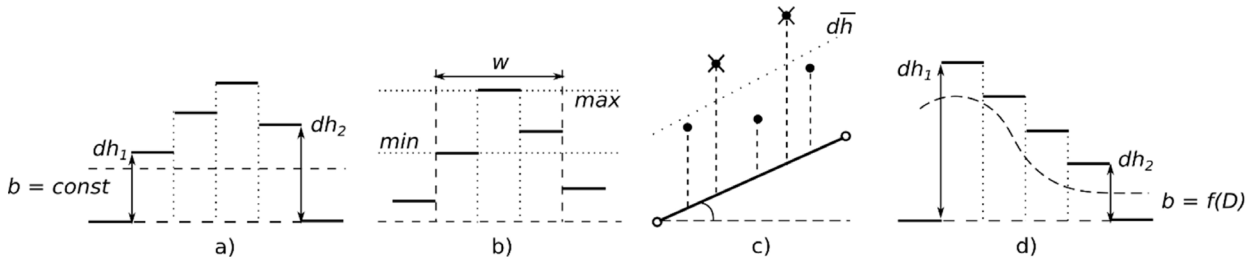
The mathematical morphology composes operations based on the set theory. Two fundamental operations, the "dilation" (see Eq.4) and the "erosion" (see Eq.5), are applied to enlarge (dilate) or to reduce (erode) the size of objects in an image [3]. The combination of an erosion and a dilation generates "opening" and "closing" operations that are applied to filter LIDAR data.

$$d_p = \max(w \in W_p) \tag{4}$$

$$e_p = \min(w \in W_p) \tag{5}$$

where p – a pixel, W_p – the window around a pixel p.

The opening operation is achieved by performing the erosion of image followed by the dilation, while the closing operation is accomplished by carrying out the dilation firstly, and then – the erosion [3].



Schemas of height usage: a) region growing; b) progressive morphological filter; c) adaptive TIN surfaces; d) graph-cut

```
// s  - slope of earth elevation,
// dx       - cell size (in geospace),
// b  - half of window size,
// n  - number of iterations,
// dh0 - initial height difference
// (threshold)
// mh       - maximal difference of
height

Function PrMFiltr(dsm, s, dx, b, n,
dh0, mh)
  dh dh0
  labels  0 // |labels| = |dsm|
  windows
  For each w in windows
    dsm′  OpeningFunc( dsm, w )
    //p – pixel, p′ ϵ dsm′
    For each p in dsm
      If (p – p′) > dh then
        labels[p]  k+1
      End if
    End for
    dsm  dsm′
    dh
    If dh > mh then
      dh  mh
    End if
  End for
  //pixels equal to zerro are ground
points
  Return labels
End function
```

C. Adaptive TIN surfaces

This method is based on constructing of triangular irregular network (TIN), when statistics on the height distance from point to triangle facet is obtained and points are filtered using the median in each iteration [4].

This study applies simplified model without statistics on angle and without mirror points. Only the vertical distance to point is calculated (see Eq.7) using the equation of plane (see Eq.6).

$$Ax+By+Cz=D \tag{6}$$

$$z' = z \quad \frac{D-Ax-By}{C}, \tag{7}$$

where $p = (x, y, z)$ are the coordinates of points and $p' = (x, y, z')$ is the point's projection on triangle facet.

```
Function AdaptiveTINsurfaces( points )
  dx  64
  While dx > 0.5
    dsm  GetDSMbyMinPoints( points,
dx )
    tin  GenerateTIN( dsm )
    distances  DistPntTriFacet(tin,
points)
    median  GetMedian( distances )
    stdDev  GetStdDev( distances )
    points
      dx dx / 2
  End while
  Return points
End function
```

D. Graph-cut

Min-cut/max-flow algorithms are graph methods to complete image segmentation. It was proved, that graph-cut algorithms can complete the semantic segmentation of remote sensing data, if seed points belong to classes and the distance between adjacent points is defined using metric based on class features [5]. This method was applied to detect buildings using LiDAR data [6]. One of graph-cut methods is Dinic's algorithm.

```
Function DinicsAlgorithm( dsm )
  oSeeds  FindObjectSeeds( dsm )
  bSeeds  FindBackgroundSeeds( dsm )
  graph  CreateGraph( dsm, oSeeds,
bSeeds )
  source  GetSource( graph )
  sink  GetSink( graph )
  path  FindShortPath(source, sink,
graph )
  While path  Ø
   minv  GetMinimalCapacity( path )
   graph  LaunchFlow( graph, path,
minv )
   path  FindShortPath(source, sink,
graph)
  End while
  segments  CutGraph( graph )
  Return segments
End function
```

In the case of DSM processing, objects are surface facilities and background – ground. The seeds are the pixels with strong height difference, the higher point is assigned to object, the lower – to background.

IV. Materials and Methods

The experiment is carried out in two stages:

1) the calculation of detected buildings;

2) the classification of all nDSM objects using the random forest algorithm based on the evaluation of geometric features.

The detected building is identified using the intersection of manually created layer and generated nDSM, but the shape of detected building can be distorted enough to be unclassified. Other problem are noise shapes, which are similar to buildings. Therefore, the shapes of nDSM objects are classified using the random forest algorithm, which classifies shapes using their geometric features. The classification method is described in the source [7].

The experiment was completed using the LiDAR data of urban area. The area of region is 1 km^2. It has relatively flat surface and contains forested areas and 251 buildings.

V. Results and Discussions

The example of the detected buildings is depicted in Fig.3. The number of detected buildings by each algorithm is provided in Table I.

Tables II-V are confusion matrices to evaluate the shapes of detected objects, where the columns are expected outputs (*E*), the rows – classification results (*R*). The number of classified buildings in Tables II-V is different comparing with Table I, because Table I depicts the number of buildings, but Tables II-V depict the number of building parts. Many parts can belong to one building, if algorithm distorted it (see Fig.3c-d). On the other hand, algorithm can group many buildings into one shape (see Fig.3b).

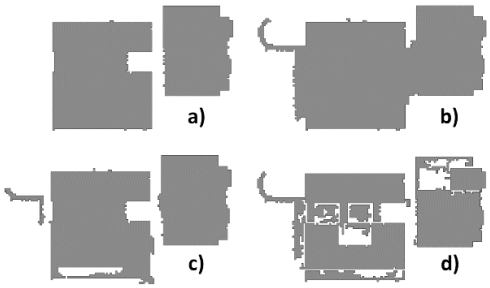


Fig. 1. Fig.3. Detected buildings by different algorithms: a) region growing; b) progressive morphological filter; c) adaptive TIN surfaces; d) graph-cut

TABLE I. Number of Detected Objects

Algorithm	Buildings	Noises
Expected Value	251 (100%)	0
Region Growing	159 (63%)	419
Adaptive TIN Surfaces	245 (98%)	598
Graph-cut	247 (99%)	321
Progressive Morphological Filter	248 (99%)	638

TABLE II. Confusion Matrix of Region Growing

R. / E.	Buildings	Noises
Buildings	**0.228** (128)	**0.127** (71)
Noises	**0.025** (14)	**0.620** (348)
Total: 561; Accuracy: 0.849; Kappa: 0.646		

TABLE III. Confusion Matrix of Adaptive TIN Surfaces

R. / E.	Buildings	Noises
Buildings	**0.198** (162)	**0.092** (75)
Noises	**0.073** (60)	**0.638** (523)
Total:820; Accuracy: 0.835; Kappa: 0.592		

TABLE IV. Confusion Matrix of Progressive Morphological Filter

R. / E.	Buildings	Noises
Buildings	**0.212** (183)	**0.124** (107)
Noises	**0.049** (42)	**0.615** (531)
Total: 863; Accuracy: 0.827; Kappa: 0.590		

TABLE V. Confusion Matrix of Graph-Cut

R. / E.	Buildings	Noises
Buildings	**0.389** (222)	**0.007** (4)
Noises	**0.047** (27)	**0.556** (317)
Total: 570; Accuracy: 0.946; Kappa: 0.888		

Considering the data of Table I, graph-cut, progressive morphological filter and adaptive TIN surfaces have the best results (98% – 99%), in turn, the region growing has the greatest loss of detected buildings – 37%. And graph-cut provides the smallest number of noises.

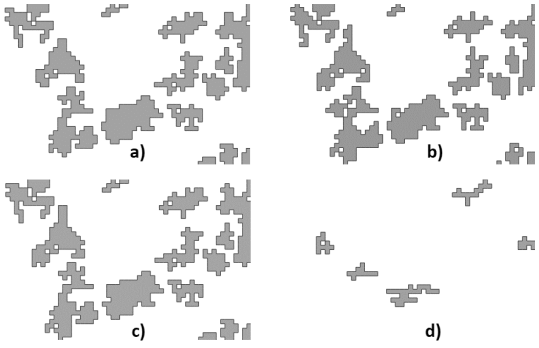
The analysis of confusion matrices (see Table II-V) provides next results:

- graph-cut obtained the best classification results (see Table V);
- despite the fact, that region growing algorithm detects the smallest number of buildings, it has sufficiently good classification results (see Table I-II);
- the main problem of region growing algorithm, adaptive morphological filter and adaptive TIN surfaces; are noises classified as buildings (see Tables II-IV).

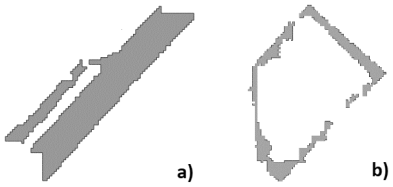
Considering the large amount of classified noises as buildings, the visual analysis of noises was completed. Fig.4 depicts the noises of similar region, where region growing, progressive morphological filter and adaptive TIN surfaces provide visually similar noises, compact and with larger area unlike graph-cut noises. According to the research [7], the compactness is the main feature in the classification process just like area. Therefore, these methods obtain the smaller classification accuracy as opposed to graph-cut.

Other problem of progressive morphological filter and adaptive TIN surfaces is exactly the usage of window, which approximates results for all pixels in the window (see Fig.5).

Nowadays, the popular solution for semantic segmentation is deep learning. For example, deep learning scholars propose next results for building detection using airborne LiDAR data: overall kappa – 0.89 [8] and F_1 score 0.93 [9] and 0.95 [10]. So, the accuracy of graph-cut method, which provided kappa 0.888 and F_1 score 0.933 (recalculating confusion matrix), is comparable with deep learning solutions.



Noises provided by algorithms: a) region growing; b) progressive morphological filter; c) adaptive TIN surfaces; d) graph-cut



Examples of window cuts: a) progressive morphological filter; b) adaptive TIN surfaces

VI. Conclusions

The completed experiments have showed, that graph-cut method detects the largest amount of buildings (99%) and provides the best classification accuracy (overall accuracy 0.95, kappa 0.89 and F_1 score 0.93).

The main problem of other methods is the high number of noises, which have large area and compact shapes similar to buildings.

The usage of methods with dynamic threshold better detects buildings, however, methods must be cluster based for more correct shape extraction.

ACKNOWLEDGMENTS

Author expresses his gratitude to the State Land Service of Latvia and to Latvian Geospatial Information Agency for providing samples for the research purposes.

References

[1] S. Szénási, "Distributed Region Growing Algorithm for Medical Image Segmentation," *International Journal of Circuits, Systems And Signal Processing*, vol. 8, pp. 173-181, 2014.

[2] J. Kilian, N. Haala, and M. Englich, "Capture and evaluation of airborne laser scanner data," *International Archives of Photogrammetry and Remote Sensing*, vol. 31, pp. 383–388, 1996.

[3] K. Zhang, S.-C. Chen, "A Progressive Morphological Filter for Removing Nonground Measurements From Airborne LIDAR Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 4, pp. 872-882, 2003.

[4] P. Axelsson, "DEM Generation from Laser Scanner Data using Adaptive TIN Models," International Archives of Photogrammetry and Remote Sensing, vol. 33, pp. 110-117, 2000.

[5] S. Kodors, "Land Cover Recognition using Min-Cut/Max-Flow Segmentation and Orthoimages," *Proceedings of the 10th International Scientific and Practical Conference*, vol. 3, pp. 127-133, 2015.

[6] S. Kodors, A. Ratkevics, A. Rausis, J. Buls, "Building Recognition Using LiDAR and Energy Minimization Approach," *Procedia Computer Science*, vol. 43, pp. 109-117, 2015.

[7] S. Kodors, "Geometric Feature Selection of Building Shape for Urban Classification," Proceedings of the 11th International Scientific and Practical Conference, vol. 2, pp. 78-83, 2017.

[8] Z. Sun, X. Zhao, M. Wu, C. Wang, "Extracting Urban Impervious Surface from WorldView-2 and Airborne LiDAR Data Using 3D Convolutional Neural Networks," *Journal of the Indian Society of Remote Sensing*, vol. 47, pp. 1–12, 2018. https://doi.org/10.1007/s12524-018-0917-5

[9] Z. Yang, W. Jiang, B. Xu, Q. Zhu, S. Jiang, W. Huang, "A Convolutional Neural Network-Based 3D Semantic Labeling Method for ALS Point Clouds," *Remote Sensing*, vol. 9, no. 9, p. 936, 2017. https://doi.org/10.3390/rs9090936

[10] Y. Sun, X. Zhang, X. Zhao, Q. Xin. "Extracting Building Boundaries from High Resolution Optical Images and LiDAR Data by Integrating the Convolutional Neural Network and the Active Contour Model," *Remote Sensing*, vol. 10, no 9, p. 1459, 2018. https://doi.org/10.3390/rs10091459

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 84-88

A Method Football Team Model Optimization and Application of the *Optimization Control*

Nguyen Hoang Mai
Electrical Engineering
Danang University of the Science and Technology
Danang, Vietnam
nhmai@dut.udn.vn

Tran Van Dung
Electrical Engineering
Daklak Technical College
Daklak, Vietnam
tranvandungcdn@gmail.com

***Abstract*—The development of the AI, IoT, and Big Data have to become strongly apply to discrete event strings systems. That are modern developments of the world. Therefore, we have to have an advanced method to develop adaptive applications, especially with MIMO discrete event systems. There is a limit while using a continuous calculation to control systems because the big calculation is an obstacle. So we have to find an optimization method to reduce the number of parameters in the calculation at any time. We could do it by choice the main parameters and except auxiliary parameters. In this paper, we introduce a Football Team Optimization (FTO) method, which is a new method to do optimization problem while control with many parameters system. The application and analysis to compare any method as PSO, traditional PID, which takes out the difference of this algorithm.**

***Keywords*—Football team model, traditional PID, discrete event system, robot team, self-organize**.

I. INTRODUCTION

We have to meet many applications in the practice to control with a selection of the parameters for one target. In the MIMO systems, as swarm optimization of robot team, choice sensors system in the machines, control planes go up and go down to runway and etc… we have one target for control variables. We call that is a Football Team Model, which is advanced control by self-organized. We could describe as figure 1.

Figure 1. The football team model on the stadium.

While the team is in the stadium, members of the team shall run to the ball and touch it into the goal of the opposition.

So only member contact to the ball at a time (we call this is the main member) and other members find good positions to wait for opportunities from the main member or guide leader. This gives the benefit of loss energy, motion time, save a total of the action time of the system.

Figure 2. Model of robotics team of military

In the technical military at figure 2, soldiers or robots will have to meet cases, which choice some objects to supervise with mapping point to point. That is time subject have to select objects for them. They need use optimization algorithm to take out final.

In the sensor systems, which use the sensor to measure process or parameter. We could use priority to main sensors, which change role by time. So we need to use optimal selection to the choice main sensor. Figure 3.

In the control systems with many parameters change by time, for example of adaptive control [1], [3],[4], we change the value of parameters to do adaptation into the process. We have priority to select forward and feedback signals of the MIMO systems [1],[2],[3],[6].

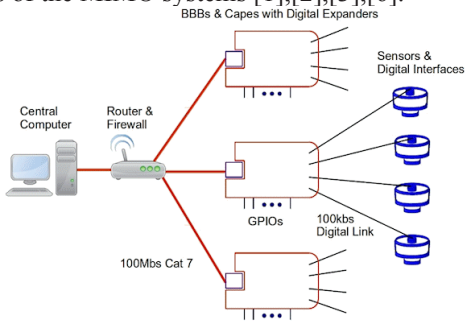


Figure 3. Model of sensors system.

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4195
© 2019 Nguyen Hoang Mai, Tran Van Dung.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

Many controllers always calibrate parameters, for example, hybrid PID controller in figure 4, the factors K_P, K_I, K_D change value to adopt state of the system. So if you have a matrix of parameters, you will have the difficult choice main factor to calibrate for accuracy is best. And also many other applications in the practice.

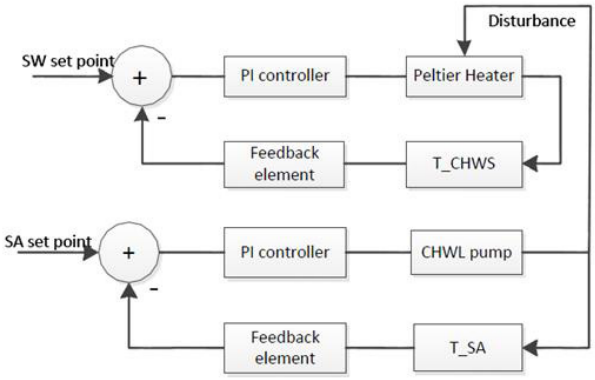


Figure 4. PID system

The main problem of FTO is a choice correction of the parameter. If you wrongly choice, you will spend much time to calculate values you need and the system can't respond well, it is very important for nonlinear systems. The parameters of nonlinear systems change chaos and maybe not a converging, so we can estimate value to decide choice any best parameter by gradient and acceleration of signal.

II. PROBLEMS

There is a system MIMO as figure 5. In that, $\mathbf{x} = [x_1, x_2, \ldots, x_n]'$ is a state vector with assume continuous signal and exists high order of derivations, the **x** includes input signals $\mathbf{y_0}$, which is a subset of **x**, $\mathbf{y} = [y_1, y_2, \ldots, y_m]'$ is an output vector, $\mathbf{d} = [d_1, d_2, \ldots, d_r]'$ is a disturbance vector. We call target function **f(e)**, with is an error of the output signals y_i, i=1..m.

$$e_i = y_{i0} - y_i; i = 1..m \tag{1}$$

With y_{i0} is a set point of output y_i.

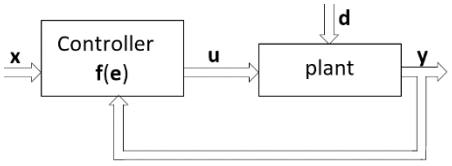


Figure 5. Model of the system.

The **f(e)** does the problem to reduce error to minimum values.

From that, we have:

$$\begin{aligned} &\mathbf{x} = \left[x_1, x_2, \ldots, x_n\right]' \\ &\mathbf{x}^{(1)} = \left[\dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n\right] \\ &\ldots \\ &\mathbf{x}^{(l)} = \left[x_1^{(l)}, x_2^{(l)}, \ldots, x_n^{(l)}\right] \end{aligned} \tag{2}$$

We define "distance of control" is a state variable x. That is a difference of set parameter value and output value. Because output signal depends on some input signals (MISO), so we can see that's crossing of the input signals with output signals as figure 6 to calibrate parameters of controller [8],[9].

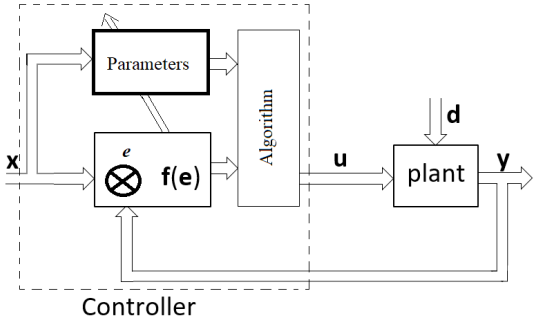


Figure 6. Model of the system with calibrates parameters.

In figure 6, the block **f(e)** decides **e**vector and choices parameter area for calibration. After reduces area of searching, **f(e)** does any parameter to directly calibrate.

III. CONTROL ALGORITHM

We go back to our football team as figure 1. A system controls one target at a time with the choice main parameter. Start at $\mathbf{x} \in R^{nx1}$, we choose the main parameter as follow with absolute values:

$$\tilde{x}_k = Min\left(x_1, x_2, \ldots, x_n\right) \tag{3}$$

$$\dot{\tilde{x}}_i = Max\left(\dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n\right) \tag{4}$$

$$\ddot{\tilde{x}}_j = Max\left(\ddot{x}_1, \ddot{x}_2, \ldots, \ddot{x}_n\right) \tag{5}$$

…

We could choice higher order so it isn't well because have to difficultly calculate. After defining the main parameter, we have tree variables with sequence *k, i, j*. In general, $k \neq i \neq j$, so we continue to filter for one time to decide tree couples of variables $\left(\dot{x}_k, \ddot{x}_k\right), \left(x_i, \ddot{x}_i\right), \left(\dot{x}_j, x_j\right)$. If we choose a higher order then we will have set as follow:

$$\begin{cases} \tilde{x}_k = Min\left(\mathbf{x}\right) with\, k \in 1..n \\ \dot{\tilde{x}}_i = Max\left(\dot{\mathbf{x}}\right) with\, i \in 1..n \\ \ddot{\tilde{x}}_j = Max\left(\ddot{\mathbf{x}}\right) with\, j \in 1..n \\ \ldots \\ \tilde{x}_q^{(p)} = Max\left(\mathbf{x}^{(p)}\right) with\, q \in 1..n; \end{cases} \tag{6}$$

And set filter:

$$\begin{cases} \left(\dot{x}_k, \ddot{x}_k, \ldots, x_k^{(p)}\right) \\ \left(x_i, \ddot{x}_i, \ldots, x_i^{(p)}\right) \\ \left(x_j, \dot{x}_j, \ldots, x_j^{(p)}\right) \\ \ldots \\ \left(x_q, \dot{x}_q, \ldots, x_q^{(p-1)}\right) \end{cases} \tag{7}$$

We define a set of time variable:

$$t_k = \frac{x_k}{\dot{x}_k}, t_k^{(1)} = \frac{\dot{x}_k}{\ddot{x}_k}, ..., t_k^{(p)} = \frac{x_k^{(p-1)}}{x_k^{(p)}} \quad (8)$$

$$t_i = \frac{x_i}{\dot{x}_i}, t_i^{(1)} = \frac{\dot{x}_i}{\ddot{x}_i}, ..., t_i^{(p)} = \frac{x_i^{(p-1)}}{x_i^{(p)}} \quad (9)$$

$$t_j = \frac{x_j}{\dot{x}_j}, t_j^{(1)} = \frac{\dot{x}_j}{\ddot{x}_j}, ..., t_j^{(p)} = \frac{x_j^{(p-1)}}{x_j^{(p)}} \quad (10)$$

...

$$t_q = \frac{x_q}{\dot{x}_q}, t_q^{(1)} = \frac{\dot{x}_q}{\ddot{x}_q}, ..., t_q^{(p)} = \frac{x_q^{(p-1)}}{x_q^{(p)}} \quad (11)$$

After we have a set of times, we choice main variable with minimum time in the absolute domain:

$$t_l = Min\left(t_k, t_i, t_j, ..., t_q\right) \quad (12)$$

In that, we have the main parameter is an x_l. If we have any time variables as same at the time then we continue choice:

$$t_l^{(1)} = Min\left(t_k^{(1)}, t_i^{(1)}, t_j^{(1)}, ..., t_q^{(1)}\right) \quad (13)$$

We choice x_l with $t_l^{(1)}$. If we have any time variables as same then we continue select to the end of the set variable. After we calculate:

$$t_l = Min\left[\left(\prod_{d=0}^{p-1} t_k^{(d)}\right), \left(\prod_{d=0}^{p-1} t_i^{(d)}\right), ..., \left(\prod_{d=0}^{p-1} t_q^{(d)}\right)\right] \quad (14)$$

And the main parameter is an x_l. So we have to choose the correct parameter to do priority control.

From that, we have a theory:

Theory: If there is a MIMO as (2), we can choice control variable with the law from (6) to (14) to ensure the control time is minimum.

Proof: we use loss function as follow:

$$f_k(e) = x_k - e_k = x_k - \dot{x}_k t_k = x_k - \ddot{x}_k t_k^{(1)} t_k =$$

$$x_k - x_k^{(3)} t_k^{(2)} t_k^{(1)} t_k = ... = x_k - x_k^{(p-1)} \prod_{d=0}^{p-1} t_k^{(d)}$$

From there, we see that if $e_k \rightarrow$ min then $\dot{x}_k t_k \rightarrow$ min. So $\dot{x}_k$ is a speed to goes to zero position, we want speed bigger than more. Therefore, the time t_k will be minimum. If we have many $\dot{x}_k$ as same then we choose $\ddot{x}_k \rightarrow \max$, and etc. Finally, we choice time as (14) then insure error is minimum.

The calculation process uses a Min-Max function to compare and refine the main parameter with minimum time. The parameter x_l shall be optimal parameter because it has minimum time to go to the control variable.

Some system could use this method to do priority control as ANN, Gen Network, AI, PID-MIMO… It has a good benefit for multi-direction system control.

IV. APPLICATION

We apply about result into some examples to illustrate application.

The first, we use tree order SISO with transfer function:

$$G(s) = \frac{1}{s^3 + s^2 + s + 1} \quad (15)$$

Use PID controller, we find stabilization domain of K_P, K_I, and K_D and simulation result as figure 7.

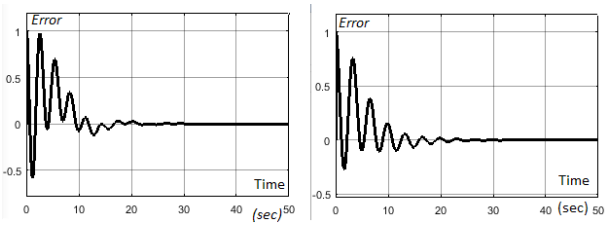


Figure 7. Simulation of the system (15)

From that, we can see the response of the system has stabilization and could control. However, the response is not "nice" for inter oscillation and we have to improve the controller.

Start $K_P = 1.5$; $K_I = 2.6$; $K_D = 0.5$, decided by Nichols – Ziegler method. Apply (6) and (12), we have to see K_P is an important factor at the start point. So we will measure error and calibrate K_P to reduce error. Use testing calculation, we choice K_P from zero time to established time. Next, we calibrate K_I for static error. We have resulted in figure 8.

By figure 8, we can see that the change of K_P has changed the response of the system. The transient time reduces to 15 seconds, so the number of oscillation cycle increase. Therefore, we could calibrate K_D to improve that.

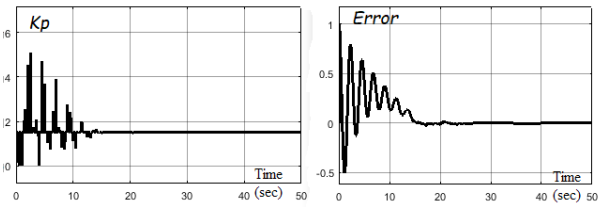


Figure 8. After calibrating K_P

We see an example more for the problem, that is an application of FTO into MIMO system with robot 2DoF, which uses DC motor as figure 9.

For this model as ideal links, we calculate inertial torque and Coriolis force, we have a dynamic equation as follow:

$$T_1 = \left(\frac{1}{3} m_1 l_1^2 + m_2 l_2^2 + \frac{1}{3} m_2 l_2^2 + m_2 l_1 l_2 \cos\theta_2\right) \ddot{\theta}_1 + \left(\frac{1}{3} m_2 l_2^2 + \frac{1}{2} m_2 l_1 l_2 \cos\theta_2\right) \ddot{\theta}_2 - m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin\theta_2 - m_2 l_1 l_2 \dot{\theta}_2^2 \sin\theta_2 + \left(\frac{1}{2} m_1 + m_2\right) g l_1 \cos\theta_1 + \frac{1}{2} m_2 g l_2 \cos\left(\theta_1 + \theta_2\right) \quad (16)$$

$$T_2 = \left(\frac{1}{3}m_2 l_2^2 + \frac{1}{2}m_2 l_1 l_2 \cos\theta_2\right)\ddot{\theta}_1 + \frac{1}{3}m_2 l_2^2 \ddot{\theta}_2 + m_2 l_1 l_2 \dot{\theta}_1^2 \sin\theta_2 + \frac{1}{2}m_2 g l_2 \cos(\theta_1 + \theta_2) \quad (17)$$

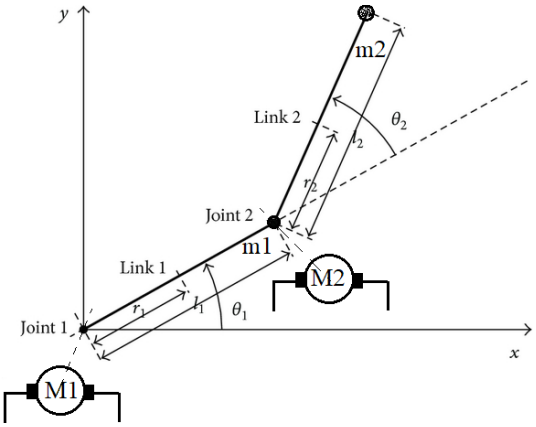


Figure 9. Model of robot 2DoF.

We use the PID controller with K_p, T_I, T_D change value by nonlinear properties of the system. The equations (15), (16) could write as a matrix:

$$\mathbf{T} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix}\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} + \begin{bmatrix} g_1(\theta) \\ g_2(\theta) \end{bmatrix} = \mathbf{M}(\theta)\ddot{\boldsymbol{\theta}} + \mathbf{N}(\theta,\dot{\theta})\dot{\boldsymbol{\theta}} + \mathbf{G}(\theta) \quad (18)$$

With:

$$m_{11} = \frac{1}{3}m_1 l_1^2 + m_2 l_2^2 + \frac{1}{3}m_2 l_2^2 + m_2 l_1 l_2 \cos\theta_2$$

$$m_{12} = \frac{1}{3}m_2 l_2^2 + \frac{1}{2}m_2 l_1 l_2 \cos\theta_2$$

$$m_{21} = \frac{1}{3}m_2 l_2^2 + \frac{1}{2}m_2 l_1 l_2 \cos\theta_2$$

$$m_{22} = \frac{1}{3}m_2 l_2^2 \; ; \; n_{11} = m_2 l_1 l_2 \dot{\theta}_2 \sin\theta_2$$

$$n_{12} = -m_2 l_1 l_2 \dot{\theta}_2 \sin\theta_2 \; ; n_{21} = m_2 l_1 l_2 \dot{\theta}_1 \sin\theta_2$$

$$n_{22} = 0 \; ;$$

$$g_1(\theta) = \left(\frac{1}{2}m_1 + m_2\right) g l_1 \cos\theta_1 + \frac{1}{2}m_2 g l_2 \cos(\theta_1 + \theta_2)$$

$$g_2(\theta) = \frac{1}{2}m_2 g l_2 \cos(\theta_1 + \theta_2)$$

Define the error of joint variables as follow:

$$\begin{cases} \theta_{e1} = \theta_{1sv} - \theta_1 \\ \theta_{e2} = \theta_{2sv} - \theta_2 \end{cases} \Rightarrow \boldsymbol{\theta}_e = \begin{bmatrix} \theta_{e1} \\ \theta_{e2} \end{bmatrix} \quad (19)$$

We will measure derivative of error to decide the main variable as (6) to (14). After the found error, we do a calibration of joint's parameters, and the process does continue…

We define "parameter's distance" is a difference of parameter in the online process. It depends on the measurement unit of the variable. From (17) we have:

$$\ddot{\boldsymbol{\theta}} = \mathbf{M}^{-1}(\theta)\left[\mathbf{T} - \mathbf{N}(\theta,\dot{\theta})\dot{\theta} - \mathbf{G}(\theta)\right] = \ddot{\boldsymbol{\theta}}_{sv} - \ddot{\boldsymbol{\theta}}_e \quad (20)$$

V. SIMULATION RESULTS

We build a simulation model as figure 10. Use an observer of controller parameters from estimation values at start time, we continue measure parameters value and their derivative values. Use from (6) to (14) to calculate and choice priority parameter at the time.

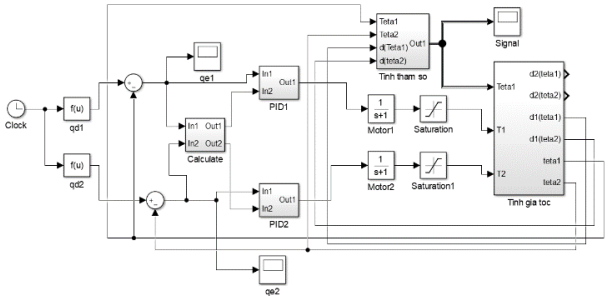


Figure 10. Model of robot 2 DoF to simulate.

We start action with estimation of PID parameters [10],[11]. Although robot manipulator is a nonlinear system, so we could use PID controller to control with strain domain. After design controller and do it, we receive an error of the trajectories of joint 1st is 15% and joint 2nd is 10% (average value). This is a big error in practice if use big trajectory more as figure 11.

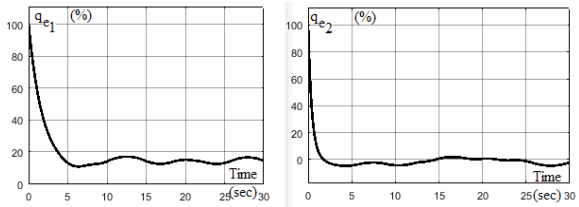


Figure 11. Simulation results with fix PID

Now we use advance algorithm above, from (6) to (14). We choice K_p is a starting strategy. We measure the error of trajectory and calculate time by derivation of the error. The first result is minimum time with K_p and next K_I. So we only use K_p to explain. The simulation results with new algorithm give better error than the traditional method of PID. We can see the error of the 1st joint is 2% and 2nd is 1% (average value) as figure 12. That is a big weight of the advanced algorithm in this paper.

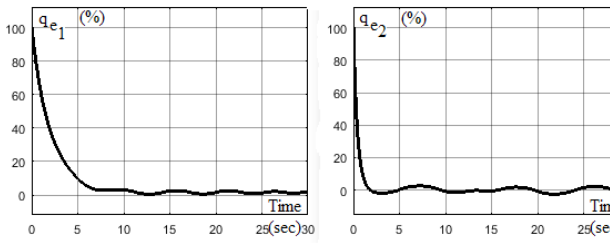


Figure 12. Simulation results with FTO

We want to do note an important issue to systems with self – calibration (adaptive regulators) controller. If choice true parameter then could do optimization time to find a balance point. In this case, we choice K_p factor by experiment because this is important of linear systems. The choice parameters can be affected to other parameters, so we need measure and change calibration between parameters.

We could take out many examples use FTO in practice as a team robot, team machine, distribution systems, etc…

VI. DISCUSSION

The theory and simulation results give a new acknowledge of control problems in the nonlinear MIMO system. There is the difference to PSO method [1],[2],[3],[4],[5] because there is a collection in PSO, so only there is an election in FTO.

The FTO method is best used in the distribution discrete system with discrete event string. In the some of analog event string, we also use FTO, so there are difficulty calculations will be apparent. Especially with un-focus parameters (chaos), we can't decide a time to finish process or accept the result with the big error.

The FTO will become a good application if we combine to AI or ANN to process discrete events in factories, medicine, military, game, transport control, and so on…

VII. CONCLUSION

In this paper, we have introduced a new method (FTO) to elect the main parameter to control in the nonlinear MIMO system. This method helps to reduce the calculation steps to find optimal elements of the process, especially in the self-organized system. When systems go outside your eye, then self-organize is an important very much, as robot team in space, distribution system with share optimization of parameters…

REFERENCE

[1] W. Deng, H. M. Zhao, J.J. Liu, X.L. Yan, Y.Y. Li, L.F. Yin, C.H. Ding, "An improved CACO algorithm based on the adaptive method and multi-variant strategies", Soft Computing, vol. 19 no. 3, (2015), pp. 701- 713.

[2] D. N. Wilke. *Analysis of the particle swarm optimization algorithm*, Master's Dissertation, University of Pretoria, 2005.

[3] Natika W. Newton, *Understanding and Self-Organization*, Front Syst Neurosci. 2017, Mar 2. doi: 10.3389/ fnsys. 2017.00008

[4] D. N. Wilke, S. Kok, and A. A. Groenwold, *Comparison of linear and classical velocity update rules in particle swarm optimization: notes on diversity*, International Journal for Numerical Methods in Engineering, Vol. 70, No. 8, pp. 962–984, 2007.

[5] P. Engelbrecht. *Fundamentals of Computational Swarm Intelligence.* Wiley, 2005.

[6] Eberhart RC, Shi Y (2000) Comparing inertia weights and constriction factors in particle swarm optimization. In: Proceedings of IEEE congress evolutionary computation, San Diego, CA, pp 84–88 10.

[7] Kennedy J (2006) Swarm intelligence. In: Handbook of nature-inspired and innovative computing. Springer, New York, pp 187–219 13.

[8] Talbi EG (2009) Metaheuristics: from design to implementation. Wiley, UK 14. Shi Y, Eberhart RC (1998) Parameter selection in particle swarm optimization. In: The proceedings of evolutionary programming VII (EP98), pp 591–600 15.

[9] W. Deng, R. Chen, B. He, Y.Q. Liu, L.F. Yin, J.H. Guo, "A novel two-stage hybrid swarm intelligence optimization algorithm and application", Soft Computing, vol. 16, no. 10, (2012), pp. 1707-1722

[10] Pedersen MEH (2010) Good parameters for particle swarm optimization. Hvass Laboratories Technical Report HL1001

[11] Bansal JC, Singh PK, Saraswat M, Verma A, Jadon SS, Abraham A (2011) Inertia weight strategies in particle swarm optimization. In: Third world congress on nature and biologically inspired computing (NaBIC 2011), IEEE, Salamanca, Spain, pp 640–647

[12] Kaveh A, Talatahari S (2010) A novel heuristic optimization method: charged system search. Acta Mech 213(3–4):267–289 42 2 Particle Swarm Optimization

[13] Qian X, Cao M, Su Z, Chen J (2012) A hybrid particle swarm optimization (PSO)-simplex algorithm for damage identification of delaminated beams. Math Probl Eng, Article ID 607418, p 11

[14] Chen WN, Zhang J, Chung HSH, Zhong WL, Wu WG, Shi Y (2010) A novel set-based particle swarm optimization method for discrete optimization problems. IEEE Trans Evol Comput 14 (2):278–300

[15] Gomes MH (2011) Truss optimization with dynamic constraints using a particle swarm algorithm. Expert Syst Appl 38:957–968

[16] Kaveh A, Zolghadr A (2011) Shape and size optimization of truss structures with frequency constraints using enhanced charged system search algorithm. Asian J Civil Eng 12:487–509

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 89-93

Potential Benefits of Web-based Idea Management System Based on Practical Evidence

Elīna Miķelsone
Managament DepartmentBA School of Business and Finance
Riga, Latvia
mikelsone.elina@gmail.com

Tatjana Volkova
Managament Department
BA School of Business and Finance
Riga, Latvia
tatjana.volkova@ba.lv

Elita Lielā
Managament Department
BA School of Business and Finance
Riga, Latvia
elita.liela@ba.lv

***Abstract*— During the last decade software developers have put increasing focus on developing information systems supporting decision making, collaboration, knowledge and idea management. Idea management systems (IMS) - manageable systematic tools to generate and evaluate ideas - are an important element within this trend. Web-based IMS are used by many well-known organizations, e.g. Volkswagen, Cisco, Microsoft, Lidl, GE Healthcare, Nestle, Procter and Gamble, Tchibo, Henkel, Roche, Fujitsu, Boeing, Goodyear, Xerox, Panasonic, Pentax, Siemens, Virgin trains etc. Although the research on web-based IMS and how these systems are applied practically is very limited. Therefore, the authors of this research aim to explore practical web-based IMS application and potential benefits. Research aim is based on theoretical and empirical approaches to deliver understanding on how web- based IMS materializes and what are potential benefits from its application to increase organizational effectiveness. Web-based IMS application will be researched by analysing separate use-cases and reprocessing the results using content analysis. 1st step – analysis of 100 users cases based on available information about application cases from homepages of web-based IMS developers. 2nd step –analysis of the results from 1st step. The analysis will be conducted by analysing information from protocols of use-cases. Protocols will be organized as category maps. Category maps are based on literature studies (finding deductive elements with available data), on data (inductive elements) and organized around Adaptive Structuration Theory elements. As the 3rd step use case descriptions and comparison will be applied. The results show that the web-based IMS are universally applicable in different organisation for fulfilling the various tasks. Benefits of web-based IMS application could be divided into 4 main groups towards benefits of- the idea management, innovation management, cooperation, and overall effectiveness of organization. Academic contribution: (1) most extensive research of web-based IMS based on 100 use cases; (2) empirical overview of potential benefits of web-based IMS application; (3) Development of IMS use case analysis protocol applicable in further researches. Practical contribution: (1) Empirical overview of web-based IMS application experience in various organizations that can stimulate the application of web-based IMS in other organizations and to deliver benefits to organization; (2) The benefits are summarized and can help in decision making whether to apply and implement a web- based IMS in and an organization.**

***Keywords*—Benefits, Idea Management Systems, Organisational Effectiveness.**

I. Introduction

IT application and information management in organizations have become as crucial as ever for matching 3 major trends– (1) in the age of innovation economy tools that provide means for acquiring, evaluation and development of knowledge and ideas are extremely important; (2) the growing role of ICT increases the importance of web-based tools that allow to support innovation process; (3) web-based IMS are becoming more important in the context of open innovation and co-innovation, giving the access to both internal and external sources of ideas and knowledge.

There is a main literature gap on providing answers to the questions how web-based IMS metalizes within organisations and what are the potential benefits from web-based IMS application? See the topicality and gap in Fig.1.

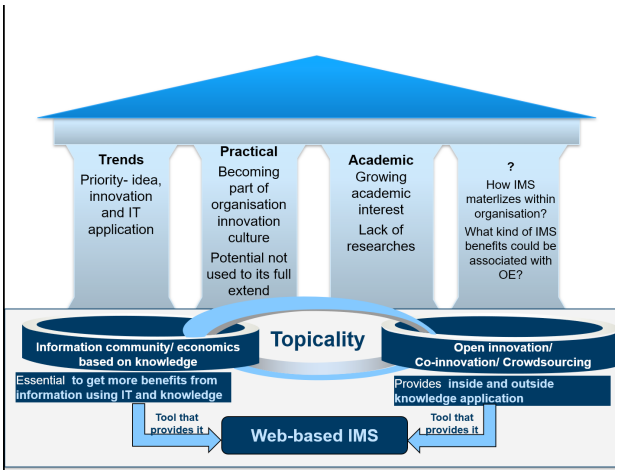


Fig. 1. Research topicality and gap.

Organisational effectiveness (OE) is a one of the main topics in management science [1], [2], and in recent

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4166
© 2019 Elīna Miķelsone, Tatjana Volkova, Elita Lielā.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

decades topicality of OE research has grown rapidly [3], [4]. Idea management (IM) as a separate research topic emerged and became very topical from 2002 focusing mainly on IM application and problems, but also about practical increase of web-based IMS application [5].

Research aim - based on theoretical and empirical approaches to deliver understanding on how web-based IMS materializes and what are potential benefits from web-based IMS application to increase organizational effectiveness.

To reach the research aim case studies were selected as the data selection method, but for data analysis – content analysis. In section II more information about methods and materials.

II. Materials and Methods

A. Definitions and basics

Idea management (IM) is a process of idea generation, evaluation, and continuation. Idea management system (IMS) is a tool, tool kit or complex system which provides systematic, manageable process of idea generation, evaluation and continuation [5], [6]. In this research authors concentrates on IMS research stream that research existing IMS not creation of new systems [7]- [9].

Adaptive structuration theory (AST) which argues that the effects of ICT on outcomes depend on the ICT structures and on the adaptive structures that forms when members apply ICT [10).

Organizational effectiveness (OE) – multidimensional measurement which could consist from financial / nonfinancial, internal/external, subjective and objective dimensions, which reflects achievements of the organization towards the goals set[11].

B. Methodology

Web-based IMS application researched by analysing separate case studies and reprocessing the results using content analysis (see in Table I).

TABLE I. Data Selection and Analysis

Section	Description
Data Selection	Case analysis. Justification - to analyze cases of web-based IMS application.
Data Analysis	Constant analysis. Justification- to compare application cases and explore benefits of web-based IMS application, that could be associated with OE.
Time Period	2010-2017
Method Application Steps	1.Creation of case study protocol. 2.Case study documents based on information available on 108 web-based IMS websites (selected in previous author's research) and individual communication with the purpose of obtaining additional information. Selection of 100 cases with most information available for our case study protocols. 3.Content analysis of the obtained materials. Creating and comparing case descriptions.

Data selection method: 100 case studies based on application cases published by web-based IMS developers and available information about the applications cases from homepages of web-based IMS developers. Data analysis method: content analysis of protocols from case studies. Protocols are organized as category maps (see in Table II).

TABLE II. Category Map

Organizational category	Basic category
Web-based IMS	Product name
	Application aim
Organizational system	Organization using the product
	Organization size
	Country
	Previous experience with IMS (yes/no)
	Moderation of IMS (automatic, manual, mixed)
	Geographical area
	Number of persons participating in IM
	Who owns the ideas
Other structural sources - tasks	Task
	Time period
	IMS application time (internal IM, external IM, mixed IM)
	Persons participating in IM
	Active/ passive IM
	Ideas created for improving product/ process/ organizational/ marketing
	Usable for solo or group idea generation sessions
	Rewards for best ideas (yes/no)
Direct results from IMS	Quantity of ideas
	Quality of ideas (how many ideas are developed)
	Participation (how many participants)
IMS application types	IMS application is in accordance to IMS type
End results	Did application achieve its aim?
	Are there any other results?
	Were new structures created?

The analysis was conducted by analysing information from protocols of case studies which are organized as category maps.

Category maps are based on literature studies (finding deductive elements with available data), on data (inductive elements) and organized around Adaptive Structuration Theory elements. The case study descriptions and comparison were applied to analyse.

Main research questions: (RQ1) How web-based IMS materialize within organisations? (RQ 2) What are the main web-based IMS benefits that could be linked with OE?

III. Results and Discussion

Research gives insight in practical application cases of web-based IMS and empirically highlights potential benefits from web-based IMS application. Research shows that web-based IMS are applied in wide variety of well-known organizations, e.g. Volkswagen, Cisco, Microsoft, Lidl, GE Healthcare, Nestle, Procter and Gamble, Tchibo, Henkel, Roche, Fujitsu, Boeing, Goodyear, Xerox,

Panasonic, Pentax, Siemens, Virgin trains.

Case studies are mostly from large organizations. Unfortunately, information about web-based IMS application in small organizations is sparse. 100 use cases are collected from both service providers and manufacturers from various industries and regions. 48 of companies reviewed didn't have any previous experience with web-based IMS application. Beneficial application cases by enterprises show that web-based IMS are easily adaptable for businesses.

Web-based IMS are used both locally and transnationally with positive results on involvement, idea quality and quantity. It is worth noting the fact that multinational companies use systems mostly for internal IM, but in very rare cases for international mixed or external ideas management. In all application cases web-based IMS are moderated both manually and automatically and all rights to ideas are owned by companies.

Web-based IMS are towards idea development from various spheres (marketing, product development, process improvements, organizational improvements). Most often ideas are generated for creating new products. This indicates the universality of the usability of the web-based IMS. Web-based IMS are flexible in terms application timeframe as there are companies that use IMS for multiple years and companies that apply IMS for separate occasions, sometimes just for few days. Whether the duration of the use experience is correlated with the results of the IMS, the authors will study in next paper.

Web-based IMS are mostly used for active internal or external IM, and mixed application cases are few. Employees, clients, general public and experts are most common groups participating in idea generation according to reviewed cases. It is rare for the web-based IMS to be adapted for specific expertise and knowledge of participating groups (e.g. Employees generate ideas and clients evaluate). Comparing the size of the network with the extent of involvement, it can be concluded that companies have the potential to increase the volume of creators / evaluators involvement, as the size of the network is much higher than the actual volume of involvement.

By examining the type of web-based IMS, it can be concluded from the pre-established classification, compared to practical use, that all systems are used according to the type of web-based IMS, e.g. The conclusion corresponds with idea found in AST – system structures have to be applied in accordance to its characteristics and aims to achieve better results.

The main results of the application of web-based IMS are the quality of ideas, quantity of ideas and involvement of creators / evaluators - provide an additional basis for studying the implications of results, as the results of the web-based IMS are different in case studies. For example, the involvement of large employees as creators of ideas / assessors results in a high quantity and quality of ideas. It verifies the practical significance of theoretical research questions. If we look at the formation of new structures, they are mostly related to the integration of the web-based IMS into the activities of a particular company (mostly in innovation and project management), which in turn determines the formation of new rules. Intellectual resources, such as patents, idea databases, etc., are created as a new structure. This conclusion verifies the research question put forward and justifies the choice to explore how the use of web-based IMS affects not only OE but also the formation of new structures as determined by AST.

The results of the study show that web-based IMS can be used to achieve different goals with diverse benefits that support theoretically derived conclusions and 100 results of web-based IMS research. Study proposes to categorize benefits of web-based IMS application such as an innovation management benefits, ideas management benefits, cooperation benefits and overall benefits of organization.

The use of web-based IMS for companies can results in IM benefits. For example, in two years, Citrix (USA) used IdeaScale web-based IMS to improve the IM process for both internal and external IM, involving 2,000 employees and customers, who generated a total of 1,800 ideas [12]. In turn, Amer Sports (Finland), using web-based IMS Qmarkets, with the aim of improving the IM process, also led to increased employee engagement in the innovation process, that is, out of 8,500 employees in IM years, 4250 were involved [13].

Innovation benefits are often mentioned as the main purpose for using web-based IMS. For example, in order to improve innovation management, Virgin Trains (UK) used web-based IMS Sideways 6, involving 7,000 employees who created ideas to improve innovation management. As a result, various innovations were introduced, employee engagement increased, and IM became easier [14]. The Avios (UK) used the same web-based IMS to promote creative ideas and created an environmental object in the Light the bulb campaign that lights up when someone submits an idea. 400 employees took part in presenting 300 ideas. The company concluded that process IM has been improved by engaging more employees [15] (see in Table III).

TABLE III. Benefits associated with OE

	In more than 50% of cases	In less than of 50% case
IM	Identify and develop new ideas. Idea storage. Structured and controlled IM. Improved IM processes. IM economize time IM without geographic and time barriers IM without barriers for involvement	-
Innovation management	Implementation of ideas Motivate innovation culture Higher creativity Speed up current innovation management process Increased innovation potential Provide ideas for new products Provide ideas for new processes Provide ideas for marketing Provide organizational improvements Support for open innovations	Greater freedom that creates better innovation habits. Improved innovation capacity. Decreased risks in Idea implementation process. Ability to identify employees that are ready to become main innovators
Cooperation	Co -creation opportunities Improvements in previous cooperation Improvements in external cooperation in IM processes Increase in participation Team work Motivation of the involved persons Networking Satisfaction with work Improved relationships in the company Strengthening of trust	Expand the number of involved persons in knowledge application. Fast and effective reaction to problems. Develop community and strengthen trust. Improve team productivity.
Organization management	Effective decision making Improvements in information management Management effectiveness Growth of the company Improved quality, client satisfaction, financial indicators. Aims achieved Defining aims Introduction of new products in the market Larger market share Ability to react to changes	Competitive advantage over competitors. Establishing new structures, IMS are integrated in organization's system, processes.

Cooperation as a goal and benefit in the use of web-based IMS appears very often, characterized by co-operative opportunities, improvement of internal co-operation, improvement of external co-operation, increase of involvement, teamwork, motivation of stakeholders, networking, satisfaction with work, strengthening of relationship in the company, strengthening of trust. Many companies only involve their employees in the use of web-based IMS, for example, the international company Autoliv - a manufacturer of security solutions for cars, has used IIS BrightIdea to create solutions for various innovation-related issues and to introduce an open innovation approach. In one year, 6000 people were involved, mostly engineers, creating 1584 ideas, 99% of which were introduced. As a result, the innovation team grew by 800% [16]. Employee Y increased the motivation of the employees by engaging them in the development of the company and promoting innovation as the goal of the company using web-based IMS Ezassi and reaching 96% of employees within one year. The company admits that 56% of ideas were used and developed [17]. Since 2005, Zebra Technologies has set the task of motivating employees through web-based IMS Imaginatik involving 2,500 employees. The use of web-based IMS was integrated into the company's innovation culture, for example, in order to motivate employees IM, the most active participants are given innovatory stripes (zebras of innovation), which confirm involvement in IM [18].

Earlier cases included general management effectiveness indicators related to decision making, financial results, etc., but there are several instances of the use of web-based IMS highlighting the direct and indirect impact of the use of web-based IMS on overall organization benefits indicators. For example, the use of Fujitsu Emeia web-based IMS Hype has resulted in an income of 25 million euros, integrating web-based IMS into project management methodology and involving 156,000 employees over 18 months [19]. In turn, the web-based IMS Idea Spotlight saving 100,000 pounds a year helped Waitrose (UK), which already involved 60000 employees for the first 6 months and gained 2100 ideas, of which 124 were introduced, but the company concluded that 50% of all ideas were with development potential [20]. Using the web-based IMS IdesScale for one year, Kane (USA), with the involvement of 1,000 employees, created ideas that resulted in 1% financial savings [21].

IV. Conclusions

A. Summary

Research shows that web-based IMS are applied in wide variety of organizations both of service providers and manufacturers from various industries and regions. Web-based IMS are used both locally and international with positive result. Web-based IMS are applied to tasks from various spheres (marketing, product development, process improvements, organizational improvements). Web-based IMS are mostly used for active internal or external IM, and mixed application cases are few. Employees, clients, general public and experts are most common groups participating in idea generation according to reviewed web-based IMS application cases. The conclusion corresponds with idea found in AST – system structures have to be applied in accordance to its characteristics - aims to achieve better results. Study proposes to categorize benefits of web-based IMS application - innovation management benefits, ideas management benefits, cooperation benefits and overall benefits of organization.

B. Contribution and Further Researches

Theoretical contribution: (1) paper fulfils an identified need to clarify web-based IMS concept applying theoretical and empirical approaches; (2) it is the widest web-based IMS empirical research based on 100 case studies; (3) indicated benefits that could be associated with OE; (4) web-based IMS case study analysis protocol is applicable in further researches. Practical contribution: (1) empirical and analytic summary of web-based IMS application cases in various organizations that can be studied and adapted by other organizations; (2) evaluation of benefits and limitations of web-based IMS is applicable in decision making process in organizations when approaching web-based IMS application. Further research directions: (1) How organisations adapt/customize systems to their

own use and implement it in their operations? (2) How different classes of web-based IMS impact organisation results? That also collides with van den Ende et. al. [22] call.

References

[1] P.S. Goodman and A.M. Saks, New Perspectives on Organizational Effectiveness, San Francisko: Jossey – Bass, 1977.

[2] S. Biswas, "Relationship between psychological climate and turnover intentions and its impact on organisational effectiveness: A study in Indian organisations," IIMB Management Review, vol.22, pp.102-110, 2010. DOI: 10.1016/j.iimb.2010.04.013.

[3] C. Mausolff and J. Spence, "Performance measurement and program effectiveness: a structural equation modeling approach," International Journal of Public Administration, vol.31 (6), pp.595-615, 2008. DOI: 10.1080/01900690701640929.

[4] J.D. Lecy, H.P. Scmitz and H. Swedlund, "Non-govermental and not-for-profit organizational effectiveness: a modern synthesis," Voluntantas: International Journal of Voluntary and Nonprofit Organizations, vol. 23 (2), pp. 434-457, 2012. DOI: 10.1007/s11266-011-9204-6.

[5] E. Mikelsone and E. Lielā, "Discussion on the Terms of Idea Management and Idea Management Systems," Journal of Regional Formation and Development Studies, vol. 3(17), pp.97-110, 2015.

[6] S. Gerlach and A.Brem, "Idea management revisited: A review of the literature and guide for implementation, " International Journal of Innovation Studies, vol 1 (2), pp. 144-161, 2017. DOI: 10.1016/j.ijis.2017.10.004.

[7] A. Westerski, T.Dalamagas and C.A. Iglesias, "Classifying and comparing community innovation in Idea Management Systems," Decision Support Systems, vol.54(1), pp.1316-1326, 2013. DOI: 10.1016/j.dss.2012.12.004.

[8] A.R. Sandriev and O.V. Pratchenko, "Idea management in the system of innovative management, " Mediterranean Journal of Social Sciences, vol.5(12), pp.155-158, 2014. DOI: 10.5901/mjss.2014.v5n12p155.

[9] M.Beretta, "Idea Selection in Web-Enabled Ideation Systems," The Journal of Product Innovation Management, vol.36(1), pp. 5-23, 2019. DOI: 10.1111/jpim.12439.

[10] M.S. Poole, "Adaptive Structuration Theory," in Encyclopedia of Management Theory, Ed. E.H. Hessler, Pace University: SAGE Publications, Inc., 2013, pp.22-25.

[11] E. Mikelsone, and E. Liela, "Idea Management and Organizational Effectiveness: Research Gap," Journal of Business Management, vol. 12(1), pp.4-24, 2016.

[12] IdeaScale, "Case Study: Citrix," 2018. [Online]. Available: https://ideascale.com/casestudy/citrix/ [Accessed: Dec. 17, 2018].

[13] Qmarket, "Case study: Amer Sports," 2018. [Online]. Available: https://www.qmarkets.net/blog/qmarkets-amer-sports-innovative-partnership-created-unprecedented-employee-engagement/ [Accessed: Dec. 17, 2018].

[14] Sideways 6, "Virgin Trains," 2018. [Online]. Available: https://fast.wistia.com/embed/medias/6705a0nxg7 [Accessed: Dec. 17, 2018].

[15] Sideways 6, "The Avios," 2018. [Online]. Available: https://fast.wistia.com/embed/medias/uc1pq9clit [Accessed: Dec. 17, 2018].

[16] BrightIdea, "Autoliv Case Study, " 2017. [Online]. Available: https://www.brightidea.com/customers/autoliv-saving-lives-through-innovation/ [Accessed: Dec. 17, 2018].

[17] e-Zassi, "Innovation Software Case Study: Creating a Culture of Innovation," 2018. [Online]. Available: https://ezassi.com/culture-of-innovation-using-innovation-software-case-study/ [Accessed: Dec. 17, 2018].

[18] Imaginatik, "Case studies," 2018. [Online]. Available: https://www.imaginatik.com/innovation-management-case-studies/ [Accessed: Dec. 17, 2018].

[19] Hype, "Fujitsu Case Study," 2018. [Online]. Available: https://i.hypeinnovation.com/clients/fujitsu [Accessed: Dec. 17, 2018].

[20] Wozoku, "Case Study Waitrose," 2018. [Online]. Available: https://www.wazoku.com/case-study-waitrose/ [Accessed: Dec. 17, 2018].

[21] IdeaScale, "Case Study: Kane," 2018. [Online]. Available: https://ideascale.com/casestudy/kane/ [Accessed: Dec. 17, 2018].

[22] J. van den Ende, L. Frederiksen and A. Prencipe, "The front end of innovation: Organizing search for ideas," Journal of Product Innovation Management, vol.32(4), pp.482-487, 2015. DOI: 10.1111/jpim.12213.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 94-97

Brain Connections Analysis Using Graph Theory Measures

Olesja Minejeva
Riga Technical University
Riga, Latvia
o.minejeva@inbox.lv

Zigurds Markovics
Riga Technical University
Riga, Latvia
Zigurds.Markovics@rtu.lv

Nauris Zdanovskis
Riga Stradins University
Riga, Latvia
nzdanovskis@gmail,com

***Abstract*—Brain is a part of the organism's complex structure that performs many functions, which are responsible for the main human abilities: to talk, to hear, to move, to see, etc. The brain consists of several areas that are not only directly connected with the different body systems, but also depend and may affect each other. Researchers and doctors are trying to summarize and visualize these relationships for an important purpose – to get the information about possible reactions of the body in case of various diseases, possibilities of recovery, risks, etc. important issues. Neurologists are looking for ways to "move" through the brain in virtual space for viewing the synapses between different areas. It might be useful to get a general idea of how brain regions are interrelated. The term "connectome", which is the complete structural description of the brain connections, or the map of connections, is used for the common perception of brain relationships. Connectome is a network of thousands of nerve fibres that transmits signals between the special regions responsible for functions such as vision, hearing, movement and memory, and combines these functions in a system that perceives, decides and acts as a whole. So, the relationships of brain neural regions can be represented as a graph with vertices corresponding to specific areas, but edges are links between these areas. This graph can be analysed using quantitative measures, like node degree, centrality, modularity etc. This article discusses the different network measures for the connections between brain's regions. The purpose is to determine the most important areas and the role of individual connections in the general functional brain model.**

***Keywords*— Brain network, connectome, functional connectivity, graph theory.**

I. Introduction

Everyone has a unique combination of genetics, environmental impact and life experience. These factors affect the detailed "structure" of the brain, as even twins can have different levels of neural links. By arranging these connections, the researchers try to understand what could be the connectomes of different people.

Connectome is a full description of the structural connections of the brain. These connections can be between different elements of the nervous system - from neurons to whole areas of the brain [1].

Connectome is enough complex and poorly to be understood. So far, only one connectome was developed to the end. It was worm's *Caenorhabditis elegans* connectome. The problem is that this worm has only 300 neurons connected to each other by 7000 links, but the human has 100 billion times more neurons and a million times more links. [2].

There are currently several Connectivity Atlases, which, like maps, display different locations or regions of human brain. Several scientific projects are devoted to this problem. One of them is the *Brainnetome Atlas* [3]. The main goal is to explore the cerebral hierarchy by highlighting the two main elements - nodes and connections between them. The most important thing is not only to identify the structural architecture of these nodes, but to combine it with the functional consistency of the individual regions, i.e. how regions can affect each other and how the presence or interception of links can affect the functionality of the affected regions.

It turns out, that the structure of the brain can be visualized in the form of a graph. The graph structure consists of vertices or nodes and edges or connections. In different complex systems, they can display different elements and links between them, such as people and their social relationships, web sites and hyperlinks, etc. [1], [4]. In the case of the brain, regions will play the role of vertices, and the edges will implement links between these regions.

II. Materials and Methods

Graph theory offers a variety of tools for working with complex system models that also include brain structure. So, it is possible to calculate a number of mathematical values for the graph you created and then integrate them into the complex network.

The graph type must be determined before any calculations are made. This article analyses the graph obtained from the *Brainnetome Atlas* project, where information about the brain regions and the existence of connections is provided, and shows is this region connected with this. So here it's about the so-called *undirected graph* [5]. In order to get and visualize this graph, we need data about

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4141
© 2019 Olesja Minejeva, Zigurds Markovics, Nauris Zdanovskis
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

the links that are gathered in one matrix, which is called *adjacency matrix*. This is a two-dimensional matrix consisting of rows and columns that reflect the names of regions (Fig. 1) [6] – [8].

Lobe	Gyrus	Left and Right Hemisphere
	SFG, Superior Frontal Gyrus	SFG_L(R)_7_1
		SFG_L(R)_7_2
		SFG_L(R)_7_3
		SFG_L(R)_7_4
		SFG_L(R)_7_5
		SFG_L(R)_7_6
		SFG_L(R)_7_7
	MFG, Middle Frontal Gyrus	MFG_L(R)_7_1
		MFG_L(R)_7_2
		MFG_L(R)_7_3
		MFG_L(R)_7_4
		MFG_L(R)_7_5
		MFG_L(R)_7_6
		MFG_L(R)_7_7
	IFG, Inferior Frontal Gyrus	IFG_L(R)_6_1
		IFG_L(R)_6_2
		IFG_L(R)_6_3
Frontal Lobe		IFG_L(R)_6_4

Names of brain regions from the adjacency matrix

Storage of information in the form of an adjacency matrix is usually associated with a term *density*. This parameter is equal to the ratio between the actual number of edges in the graph and the total number of possible edges [1], [5].

The names of 246 regions (SFG_L(R)_7_1, MFG_L(R)_7_7, IFG_L(R)_6_3 utt.) and two types of numbers – 0 and 1, are included in the adjacency matrix. If the edge between region *a* and region *b* exists then the corresponding matrix element is $A_{ab} = 1$, else $A_{ab} = 0$ [7]. The data is stored in *.csv* file format. A fragment of the adjacency matrix is shown in the Fig. 2. The number of edges connecting the vertex with others is called *node degree*. Thus, the number of "1" in the corresponding line of the adjacency matrix corresponds to the node degree of this row. Nodes with the highest degrees tend to be called *hubs* [1], [5]. On the basis of node degree we can calculate so-called *assortativity* [10], which describes the correlation between connected vertices. If this value is positive, it means that a high degree nodes prone to be connected to each other.

Another characteristic is an *average path length*. When we guess about the classic mathematical implementation, the path length is the number of edges to pass through to get from one vertex to the other. When it comes to brain topology, the path length can be used to assess the possibilities of transmitting information between the regions. Short path lengths make it quick to transfer information and reduce resource consumption during the transmitting process. This factor leads to the term „small world" [9].

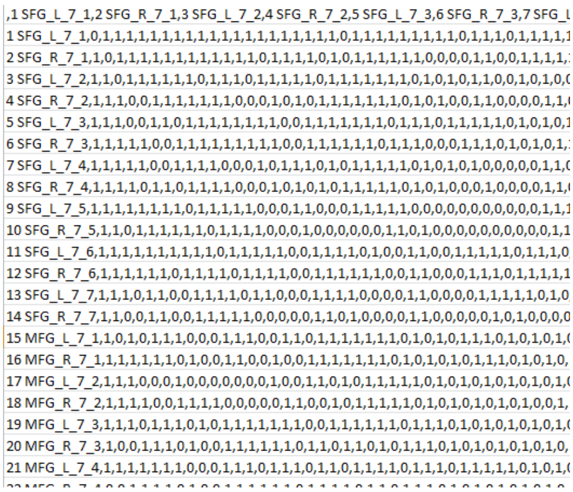

,1 SFG_L_7_1,2 SFG_R_7_1,3 SFG_L_7_2,4 SFG_R_7_2,5 SFG_L_7_3,6 SFG_R_7_3,7 SFG_L
1 SFG_L_7_1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,0,1,1,1,0,1,1,1,1
2 SFG_R_7_1,1,0,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,0,1,0,1,1,1,1,1,0,0,0,0,1,1,0,0,1,1,1,1,
3 SFG_L_7_2,1,1,0,1,1,1,1,1,1,0,1,1,1,0,1,1,1,1,1,0,1,1,1,1,1,1,1,0,1,0,1,0,1,1,0,0,1,0,1,0,
4 SFG_R_7_2,1,1,1,0,0,1,1,1,1,1,1,1,0,0,0,1,0,1,0,1,1,1,1,1,1,1,0,1,0,1,0,0,1,1,0,0,0,0,1,1,
5 SFG_L_7_3,1,1,1,0,0,1,1,0,1,1,1,1,1,1,1,1,0,0,1,1,1,1,1,1,1,1,0,1,1,1,0,1,1,1,1,1,1,0,1,0,1,0,1
6 SFG_R_7_3,1,1,1,1,1,0,0,1,1,1,1,1,1,1,1,1,0,0,1,1,1,1,1,1,0,1,1,1,0,0,0,1,1,1,0,1,0,1,0,1,
7 SFG_L_7_4,1,1,1,1,1,0,0,1,1,1,1,0,0,0,1,0,1,1,1,0,1,0,1,1,1,1,1,0,1,0,1,0,1,0,0,0,0,0,1,1,
8 SFG_R_7_4,1,1,1,1,0,1,1,0,1,1,1,1,0,0,0,1,0,1,0,1,0,1,1,1,1,1,0,1,0,1,0,0,0,1,0,0,0,0,1,1,
9 SFG_L_7_5,1,1,1,1,1,1,1,1,0,1,1,1,1,1,0,0,0,1,1,0,0,0,1,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,1,1,1
10 SFG_R_7_5,1,1,0,1,1,1,1,1,1,0,1,1,1,1,0,0,0,1,0,0,0,0,0,0,1,1,0,1,0,0,0,0,0,0,0,0,0,0,1,1
11 SFG_L_7_6,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,0,0,1,1,1,1,0,1,0,0,1,1,0,0,1,1,1,1,1,0,1,1,1,0
12 SFG_R_7_6,1,1,1,1,1,1,0,1,1,1,1,0,1,1,1,1,0,0,1,1,1,1,1,1,0,0,1,1,0,0,0,1,1,1,0,1,1,1,1,1
13 SFG_L_7_7,1,1,1,0,1,1,0,0,1,1,1,1,0,1,1,0,0,0,1,1,1,1,0,0,0,0,1,1,0,0,0,0,1,1,1,1,1,0,1,0
14 SFG_R_7_7,1,1,0,0,1,1,0,0,1,1,1,1,1,0,0,0,0,0,1,1,0,1,0,0,0,0,1,1,0,0,0,0,0,1,0,1,0,0,0,0
15 MFG_L_7_1,1,0,1,0,1,1,1,0,0,0,1,1,1,0,0,1,1,0,1,1,1,1,1,1,1,0,1,0,1,0,1,1,1,0,1,0,1,0,1,
16 MFG_R_7_1,1,1,1,1,1,1,1,0,1,0,0,1,1,0,0,1,0,0,1,1,1,1,1,1,1,0,1,0,1,0,1,0,1,1,1,0,1,0,1,0,
17 MFG_L_7_2,1,1,1,0,0,0,1,0,0,0,0,0,0,0,1,0,0,1,1,0,1,0,1,1,1,1,1,0,1,0,1,0,1,0,1,0,1,0,1,
18 MFG_R_7_2,1,1,1,1,0,0,1,1,1,1,0,0,0,0,0,1,1,0,0,1,0,1,1,1,1,1,0,1,0,1,0,1,0,1,0,1,0,0,1,
19 MFG_L_7_3,1,1,1,0,1,1,1,0,1,0,1,1,1,1,1,1,1,1,0,0,1,1,1,1,1,1,1,0,1,1,1,0,1,0,1,0,1,0,1,0,1,
20 MFG_R_7_3,1,0,0,1,1,1,0,1,0,0,1,1,1,1,1,1,0,1,1,0,1,1,0,1,0,1,1,1,0,1,0,1,0,1,0,1,0,1,0,
21 MFG_L_7_4,1,1,1,1,1,1,1,1,0,0,0,1,1,1,0,1,1,1,0,1,1,0,1,1,1,1,0,1,1,1,0,1,1,1,1,1,1,0,1,0,1,

Fig. 1. A fragment of the adjacency matrix

This concept is associated with the networks where all elements are connected to others through a very short path. Society often meets the term with a similar meaning - „six degrees of separation" between any two people, which means, that they know each other through six acquaintances (connected with 6 or fewer edges). The brain connection graph is also considered „small world" system.

In terms of graph analysis, one of the most important issues is the definition of "the most important" vertexes. What does "importance" mean? There may be several interpretations here. One of them explains that "importance" refers to the flow of information in the network. The concept *centrality* is introduced [1], [11]. If *centrality* is high, it means that the node has an important place in the information exchange processes in a graph.

It should be noted that this attribute has some perceptions and different detection approaches. The simplest idea is related to the hypothesis that the node with a high degree is potentially active. This measure is called *degree centrality* and is equal to the count of vertexes, connected to a given one [12]. Another view of centrality is based on the shortest paths between the vertexes passing through the given. Such a node can affect others, maintain, suspend, or stop the information transfer process. This approach was named *betweenness centrality*. There are still a number of measures related to centrality, but in this article one another is described – *closeness centrality*. It follows from the name that this centrality describes proximity, thus, the greater this centrality is, the closer this node is to all others. This value is calculated as the inverted sum of the path lengths of a node [13].

The network has another local feature - clustering. It describes the interaction between the vertex neighbours. In other words, it is mathematically possible to calculate the probability that two nodes neighbours also are closer neighbours. The resulting value is called *clustering coefficient* [14]. This value can be used to determine how fragmented the network is. Thus, at a high clustering coefficient in the brain graph, we can expect that an event (for example., stroke) will affect only one cluster, i.e. a group of regions. On low clustering, the information is

spread across the all network.

Of course, the described graph analysis parameters are only a small part of graph theory tools. But with regard to brain connections, they give a general idea of the interrelations and interactions between the regions. There is a wide range of software that can be used for mathematical and statistical processing of this graph and for calculating parameters of this graph. It can be divided into two groups: universal mathematical packages (MATHCAD, MATLAB, Wolfram Mathematica utt.), containing tool groups for working with graphs, and specialized computer programs (for example., Gephi), designed for purposeful processing and analysis of graphs.

This paper describes *NetworkX* software package. It is free software for creating and analysis of complex networks. The program for graph processing is created in *Python* programming language. Launching the created program, it represents parameters described above that gives the information about the links between brain regions.

III. Results and Discussion

NetworkX program allows you to import data from *.csv* file format. Such file is an adjacency matrix obtained from *Brainnetome Atlas*, where information about connections between 246 brain regions is stored. Program code for importing file:

```
#to import matrix from csv file
from numpy import genfromtxt
import numpy as np
mydata = genfromtxt("C:/Python/matrix.csv",
delimiter=',')
print(mydata)
```

Created program *mydata.py* is launches in the *cmd* window. Data is displayed on the screen (Fig. 3). As the matrix is large, only its part is visible on the screen. The names of the regions were imported as words *nan*.

```
[[nan nan nan ... nan nan nan]
 [nan  0.  1. ...  1.  1.  1.]
 [nan  1.  0. ...  1.  1.  1.]
 ...
 [nan  1.  1. ...  0.  1.  1.]
 [nan  1.  1. ...  1.  0.  1.]
 [nan  1.  1. ...  1.  1.  0.]]
```

Imported matrix from .csv file0

In order to "pull" the adjacency matrix from the imported data, some lines are added to the code:

```
#adjacency matrix from imported data
adjacency = mydata[1:,1:]
print(adjacency)
```

The adjacency matrix is output to the screen (Fig. 4).

```
[[0. 1. 1. ... 1. 1. 1.]
 [1. 0. 1. ... 1. 1. 1.]
 [1. 1. 0. ... 0. 1. 1.]
 ...
 [1. 1. 0. ... 0. 1. 1.]
 [1. 1. 1. ... 1. 0. 1.]
 [1. 1. 1. ... 1. 1. 0.]]
```

Fig. 2. Adjacency matrix

NetworkX programme is a console application, but there are additional extension packages that also allow visualization. Different types of graphical objects can be constructed in *NetworkX* by connecting the library *Matplotlib*. Connecting libraries, now we can create a graph from the adjacency matrix and display it:

```
#graph drowing
import networkx as nx
import matplotlib.pyplot as plt
G=nx.from_numpy_matrix(adjacency)
nx.draw(G)
plt.show()
```

When we run this code, a new pop-up window with a graph's image opens (Fig. 5). Since the number of vertexes and edges in the graph is large enough and the graph is wide, because the brain connections form a complex system, then the resulting drawing is complicated and non-informative. The vertexes and edges merge, so the graph cannot be used for further analysing.

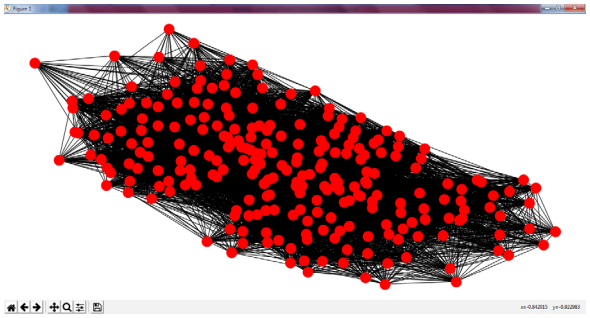


Fig. 3. Graph of brain connections

In turn, *NetworkX* has libraries with different graph parameters that can be calculated.

The first simplest parameter is node degree. It has a built-in function for calculating it - *Graph.degree*. Here are two options. If we analyse a particular vertex or some vertexes, which are similar to brain regions, then we can add a vertex number as a parameter to this function, for example:

```
G.degree[10]  # node's 10 degree
list(G.degree([0, 1, 2]))# degrees of 3 nodes
```

Alternatively, we can output a whole array of vertexes with the appropriate degrees to the screen (Fig. 6).

```
[(0, 115), (1, 107), (2, 46), (3, 38), (4, 91), (5, 83
02), (9, 96), (10, 110), (11, 113), (12, 113), (13, 10
6, 44), (17, 42), (18, 66), (19, 60), (20, 60), (21, 6
4, 49), (25, 49), (26, 104), (27, 114), (28, 42), (29,
(32, 69), (33, 70), (34, 77), (35, 76), (36, 84), (37,
(40, 88), (41, 88), (42, 79), (43, 76), (44, 53), (45,
(48, 126), (49, 107), (50, 90), (51, 93), (52, 63), (5
, (56, 59), (57, 47), (58, 70), (59, 72), (60, 55), (6
, (64, 105), (65, 98), (66, 83), (67, 80), (68, 114),
64), (72, 84), (73, 77), (74, 54), (75, 62), (76, 110)
 43), (80, 55), (81, 68), (82, 74), (83, 77), (84, 56)
 62), (88, 60), (89, 73), (90, 75), (91, 54), (92, 76)
 55), (96, 81), (97, 44), (98, 56), (99, 53), (100, 44
(103, 112), (104, 82), (105, 74), (106, 80), (107, 73)
10, 72), (111, 71), (112, 67), (113, 73), (114, 68), (
, 78), (118, 96), (119, 100), (120, 32), (121, 34), (1
76), (125, 83), (126, 87), (127, 94), (128, 21), (129,
, (132, 49), (133, 53), (134, 66), (135, 61), (136, 62
139, 40), (140, 63), (141, 56), (142, 54), (143, 38),
, 153), (147, 136), (148, 146), (149, 134), (150, 138)
(153, 138), (154, 61), (155, 64), (156, 53), (157, 57)
```

Fig. 4. Degrees of graph nodes

Other parameters were also calculated for graph analysis:

```
#Assortativity
r=nx.degree_assortativity_coefficient(G)
print("%3.1f"%r)
#Clustering coefficient
print(nx.average_clustering(G))
#Average path length in the graph
print(nx.average_shortest_path_length(G))
#Density
print(nx.density(G))
#Centrality
print(nx.degree_centrality(G))
print(nx.betweenness_centrality(G))
print(nx.closeness_centrality(G))
```

Launching such a program a number of numerical values is obtained and summarized in Table I.

TABLE I. ESTIMATED GRAPH PARAMETERS

Graph Parameters	Value
Assortativity	0
Clustering coefficients	0.69
Average path length	1.65
Density	0.37
Diameter	3

The first parameter – assortativity – equal to 0. This coefficient indicates whether high-level vertexes tend to interact with the same or similar vertexes. Since the calculated coefficient is equal to 0, this means that this graph can be called as *non-assortative graph*. In such graphs there is no correlation between vertexes that give the name of such complex systems - *uncorrelated networks*.

If you pay attention to the fragmentation of the graph, let's look at the value of the clustering coefficient obtained. Numerically it is equal to 0.69. It is difficult to assess whether this figure is considered to be low or too high. From a theoretical point of view „small – world" graphs have high clustering. The coefficients is between 0 and 1. This value reaches 0.69 for the given graph, that is noticeably larger than half, and means, that brain regions form separate groups or subgroups, in which almost all vertexes have connections to others. A more detailed view can be given by cluster analysis methods.

The characteristics of the information exchange in the graph should be analyzed separately. The obtained valu of density is 0.37. After this relationship, especially in the context of brain activity, it is impossible to say objectively and unambiguously whether the transmission of information between the regions is good or bad. Of course, this ratio is not approaching 1.0, where there are all the possible edges between nodes, but on the other hand it has gone far enough from 0. To create a full scene, let's look at how you can get from one node to the other. Average path length in the graph is 1.65. So you can transfer impulses from region *a* to region *b*, offending only a few regions on the road. Maximum path length directly from one node to other – diameter – is equal to 3. Therefore it also confirms the relatively efficient and fast flow of information in the graph.

IV. CONCLUSIONS

A method to describe brain connectivity using graph theory measurements is described in this article. There are a lot of parameters that can be calculated for the graph of connections between the brain regions. First of all, we can conclude that brain activity is a very complicated system and can be represented as a wide graph with a big number of vertexes and edges. So, it is difficult to analyse whole graph.

Brain connections graph can be called "small - world" system. The "small - world" graphs tend to contain subgraphs that have connections between almost all the nodes in them. This follows from the definition of such a property of the graph as a high coefficient of clustering. Secondly, most pairs of vertexes are connected at least in one short way. Thus, some of brain regions tend to make separate groups. It means that signals through connections will affect only one group of regions but not whole brain.

Mathematical analysis has been performed using *NetworkX* software. Calculated values have created a common presentation about the analysable graph. But for more accurate analysis, more mathematic software packages can be used to compare obtained results.

REFERENCES

[1] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems", Nat. Rev. Neurosci. 10, pp. 186–198, 2009.

[2] J. M. Perkel, "This Is Your Brain: Mapping the Connectome", LIFE SCIENCE TECHNOLOGIES.

[3] L. Fan, H. Li, J. Zhuo, Y. Zhang, J. Wang, L. Chen, Z. Yang, C. Chu, S. Xie, A. R. Laird, P. T. Fox, S. B. Eickhoff, C. Yu, and T. Jiang, "The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture", Cereb Cortex. 2016 Aug; 26(8): 3508–3526.

[4] O. Sporns, D. Chialvo, M. Kaiser, and C. C. Hilgetag, "Organization, development and function of complex brain networks", Trends Cogn. Sci. 8, 418–425 (2004).

[5] K. Börner, S. Sanyal, and A. Vespignani, "Network science". Annu. Rev. Inform. Sci. Technol. 41, 537–607 (2007).

[6] J. C. Reijneveld, S. C. Ponten, H. W. Berendse, and C. J. Stam, "The application of graph theoretical analysis to complex networks in the brain", Clin. Neurophysiol. 118, 2317–2331 (2007).

[7] C. J. Stam and J. C. Reijneveld, "Graph theoretical analysis of complex networks in the brain", Nonlin. Biomed. Phys. 1, 3 (2007).

[8] O. Sporns, G. Tononi, and R. Kotter, "The human connectome: A structural description of the human brain", PLoS Comput Biol 1(4): e42. 2005.

[9] D. J. Watts, and S. H. Strogatz, "Collective dynamics of 'small-world' networks", Nature, 393:440-442, 1998.

[10] R. Noldus and P. V. Mieghem, "Assortativity in complex networks", Journal of Complex Networks, Volume 3, Issue 4, pp. 507–542. 1 December 2015.

[11] S.P. Borgatti and M.G. Everett, "A Graph-theoretic perspective on centrality", Social Networks. 28 (4), pp. 466-484. 2006.

[12] M. E. Shaw, "Group structure and the behavior of individuals in small groups", J. Psychology,. V. 38. pp. 139-149.1954.

[13] G. Sabidussi, "The centrality index of a graph", Psychometrika. 31 (4): 581–603. 1966.

[14] T. Opsahl and P. Panzarasa (2009). "Clustering in Weighted Networks". Social Networks. 31 (2): 155–163.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 98-103

Information Technology Competency Management in the Financial Sector in Latvia

Andris Natrins
Department of Research and Development BA School of Business and Finance
Riga, Latvia
andris.natrins@ba.lv

Lasma Supe
Finance Department (Faculty) BA School of Business and Finance
Riga, Latvia
lasmasupe@gmail.com

Elina Mikelsone
Managament Department (Faculty) BA School of Business and Finance
Riga, Latvia
mikelsone.elina@gmail.com

Andris Sarnovics
Managament Department (Faculty) BA School of Business and Finance
Riga, Latvia
andris.sarnovics@ba.lv

***Abstract*—Digital transformation is penetrating financial sector. Financial sector is challenged by new start-up companies who combine finance and technology, thereby creating disruptive innovations. The authors perform a study on information technology impact on competency management in the financial sector of Latvia based on the technology management competency model.**

The authors have executed extensive academic literature review and interviews with leading fintech industry experts. The research team performed mixed research combining qualitative and qualitative methods. The process of the quantitative research was designed as a collection of publicly available financial statement data from lursoft.lv (enterprise database) for selected fintech companies and the analysis of financial data by comparing different fintech companies. The process of the qualitative research was designed as face-to-face expert interviews and the definition of central interview questions for the research based on the literature review and related sub-questions. The authors use the one-way analysis of variance technique and prove that competency management differs among fintech companies by comparing average turnover per employee of platform and credit companies. The finding persuades the authors to propose a fundamental fintech competency model to ensure competitiveness and sustainability of fintech companies.
The objective of the research is to define a fundamental set of competencies linked with information technology management for financial sector companies to maintain competitiveness.

Results prove that current set of competencies employed by fintech companies represents sufficient technical competencies. The conclusions demonstrate that fintech companies tend to have different competency management models. The research team concludes that soft skills development, data analytics using advanced data analytic tools, technology awareness, the ability and experience to use open source technology tools to develop technology solutions without deep technology competency, the ability to see the big picture, and interconnections between processes are competencies of the future.

***Keywords*—Competency Management, Competency Model, Financial Sector, Information Technology**

I. Introduction

The authors perform a study on information technology impact on competency management in the financial sector of Latvia based on the technology management competency model [1]. According to academic publications, there is a problem of competency mismatch between labor market and requirements set by companies. The employment level remains low, reaching 60.5% in the second quarter of 2018 in the European Union [2], while companies attempt to recruit competencies. Companies are mostly focused on recruiting competencies of the past and present to achieve business goals. There is a debate in the academic literature that future competency needs are often neglected due to the lack of competencies in the labor market and the lack of decent competency management by companies.

Widely is discussed interaction between financial sector and technology innovations. Financial sector companies are looking to use technology solutions to establish innovative, asset light, scalable, competitive and sustainable business models. New finance and technology-based business models are knowledge intensive.

In the last ten years, the financial services sector is experiencing radical changes on a global scale. The outcome of this process is the very rapid development of financial technologies, or fintech industries, which have effectively managed to duplicate virtually the entire spectrum of financial services with IT solutions-based alternatives.

Fintech has developed itself as a separate segment of industries with in the financial services sector of Latvia, directly affected by the "footprint" of digital transformation in this sector. By replacing the traditional for Latvia so called non-resident oriented financial services segment, fintech companies show a rapid capitalization, growth in turnover, an increase of customer numbers and a range of covered export markets.

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4168
© 2019 Andris Natrins, Lasma Supe, Elina Mikelsone, Andris Sarnovics.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

Although this segment of the financial services sector is characterized by heterogeneity in the business model (including in terms of competitiveness and sustainability), the common feature for all companies are high technology intensity and, very often, a high proportion of its own financial technology development activities.

In view of this exposure, it can be expected that it leads to structural changes in employment that may not be present. It is likely that it is not homogenous not only in the entire spectrum of financial services sector but also in the various fintech's industries.

The objective of the research is to define a fundamental set of competencies for financial sector companies to maintain competitiveness. The authors of this study considered that competencies in the fintech industries are currently fully representing changes in skills and knowledge needs and, to a enough extent, the diversity of skills and knowledge requested for the financial services sector. Therefore, the fintech industries has been selected for the analysis of the situation in this study.

Null and alternative hypotheses were formulated on the basis of the problem statement and the objective of the research (Table I).

TABLE I. HYPOTESIS STATEMENT

Null Hypothesis	**Alternative Hypothesis**
H_0: Competency management is equal among fintech companies	H_1: Competency management is different among fintech companies

The authors of the research have identified that the impact of information technology on competency management in the financial sector of Latvia is relevant for the research. Latvia is ranked 37th among 137 countries with a positive tendency towards technological readiness with highly ranked technological adaption and availability of latest technologies according to the global competitiveness index 2017 – 2018. Over the past years, Latvia has had a negative trend in the financial market development index ranking (64th spot), which demonstrates negative tendencies in the availability and affordability of financial services [3].

The global competitiveness index and Ministry of Economics financial sector development goals show importance and urgency to understand competency management synergy between technology and financial sector development to improve the competitiveness of Latvia's financial sector. The objective of the research is to define a fundamental set of competencies for financial sector companies to maintain competitiveness.

To reach the aim research team has applied 3 main data selection methods: extensive academic literature review and interviews with leading fintech industry experts, collection of publicly available financial statement data.

II. Materials and Methods

A. Background of the Study

There are various extensive academic studies about information technology management competencies but a limited number of academic studies about finance management competencies change due to digital transformation [4], [5], but little evidence on the relationship between information technology management and finance management competencies.

In the context of this study, the authors recognize that recent academic studies refer to the increased importance of competency management and measurement models as key sustainability and competency factors for contemporary companies [6]. Financial sector changes with the growing number of new fintech companies. Business models are created on the basis of synergy between information technology management competencies and finance management competencies, bring companies to the point where it is no longer possible to achieve business goals only through existing competencies leverage, but completely new sets of competencies will be brought to companies [7].

Research team investigates academic publications on digital transformation, financial sector structural change and information technology impact on financial sector development. There are very few scientific researches on competency management with respect to fintech companies. The authors research competency management based on the technology management competency model proposed by Doggett, McGee and Scott [1]. Finally, research splits fintech companies into two types based on their business profiles and validates literature research results through leading fintech expert interviews.

The dataset of this study consists of expert interviews and financial statement data of fintech companies for 2016 and 2017. The empirical part of the work is done using last two-year data as most of the fintech companies were established in 2015. The authors introduce a variable to calculate the effectiveness of competency management for different fintech companies and tests differences between fintech companies using a one-way analysis of variance.

Finally, the authors propose a competency management model for fintech companies to address their future needs. The research team has noted that there exist different competency definitions offered in various academic studies. Competency is extensively defined as a performance aspect for the combination of skills, knowledge, expertise, values, social and methodical abilities, ambitions and attitudes that are used by individuals for personal growth to perform specific tasks in an effective manner and in line with values and goals of the organization [8]- [10].

Based on the definitions collected from literature sources, research team proposes to define individual competency and organizational level competence. Individual competency is a set of ambition, skills, knowledge and characteristics a self-driven individual

uses and is able to train to achieve personal effectiveness. Organizational level competence is a set of ambitions, skills, knowledge and characteristics owned by a company through employees, which measures and predicts employee effectiveness to achieve organizational goals.

Technology management competencies have been extensively researched by Doggett, McGee and Scott [1]. The authors propose a core technology management competency model linking process, project, systems and operations through the management context that refers to self-management, people management, quality management and risk management.

Financial sector digital transformation is one of the most expeditiously researched topics across academic publications during the last decade. The authors consider "financial sector" to be a category of overall economy consisting of a variety of organizations that provide financial services to individuals and businesses. Paper treats "fintech industry" as a specific business domain, which is part of financial sector.

B. Methods Applied

Research performed mixed research combining qualitative and qualitative methods for exploring and understanding the defined problem statement.

Creswell [11], research team followed qualitative

Following guideline steps recommended by research methodology:
developed a questionnaire based on literature review;

- collected data in the field on-site through face-to-face expert interviews in the premises where experts can be observed in their natural setting;
- gathered multiple forms of data and organized additional collected sources of data into categories to get a broader understanding;
- built patterns based on categories to establish a comprehensive set of themes and then look back at themes to understand if more evidence in the form of additional data needs to be collected to support identified themes;
- explained to the participants the research team background and experiences to design potential shaping of research team interpretation and direction of the study;
- reported to the participants multiple perspectives, factors and situations sketching a larger picture that emerges by representing the holistic approach of the qualitative research.

Quantitative research was based on financial statement data analysis. The process of the quantitative research was designed as the collection of publicly available financial statement data from lursoft.lv (enterprise database) for selected fintech companies and the analysis of financial data by comparing different fintech companies.

Sample selection. In the beginning of the research identified ten fintech companies operating in Latvia for at least a 12-month period, which provided structurally different services like capital raising, payments, online credits and online auto leasing platform. Research team approached the CEO of each company to schedule interviews with CEOs, Heads of IT and Heads of HR. Fintech companies of different business types were selected to test awareness among different finance technology companies. Employees of different levels within the same company were selected for the expert interviews to test their awareness of the impact produced by information technology competency on the company's competitive position. Out of ten approached companies, six companies agreed to an expert interview. Out of six companies, two companies agreed to proceed with scheduled expert interviews for the CEO, Head of HR and Head of IT and four companies agreed to C-level expert interviews; as a result, in the research conducted nine expert interviews out of potential thirty expert interviews within the research: six C-level expert interviews, two expert interviews with Heads of IT and two expert interviews with Heads of HR.

Main research questions: (1) How do you evaluate current personnel competencies to succeed in new technology implementation in your company? (2) What competencies are lacking and what competencies are sufficient for candidates applying for positions in your company? (3) What competencies are needed to support future needs of your company?

Qualitative data analyses using content analysis but quantitative data analysis with statistical methods.

III. RESULTS AND DISCUSSION

C. Interview results

Current personnel competencies. All level experts agree that current personnel are missing big picture understanding. The experts emphasize the importance of fintech industry knowledge and self-learning ability.

The paper emphasizes that experts were extensively discussing that current personnel were missing big picture understanding. Research results proves that it is necessary to establish a communication and cooperation environment where technical teams could obtain domain knowledge from business teams and where business teams would be knowledgeable about latest technologies integrations into business processes. The research points out that big picture understanding can be obtained by establishing effective communication channels among business domains.

Competencies lacking for candidates applying for positions. All the experts share the same opinion that there are enough technically skilled and traditional specialists like financial analysts in Latvia. The credit fintech C-level experts point out that industry knowledge and experience are missing, but the platform fintech experts do not consider the lack of experience to be an

issue. The technical experts point out that technical people lack an understanding of product lifecycle and business. The credit fintech HR expert mentions that there is no knowledge of how to work with advanced data analytics using technology tools like SQL or R. The platform HR notes that critical thinking and analytical mind are missing competencies.

Research highlights that credit companies see competency gaps due to a complex business environment in which credit companies are operating. The paper points out that the opinion that there are enough competencies in the market to support platform fintech companies is due to the fact that the platform business environment has just started to develop. The paper discusses that there are fewer capital funds available for platform companies in comparison with credit companies, where the average net annual return for investors is 11.68% [12]. The research stresses that limited capital availability is decelerating the growth of platform companies, competitiveness is still low enough and, therefore, platform companies can operate with the current set of competencies. The paper emphasizes that increasing competitiveness in combination with the ambition to scale the business will motivate platform companies to employ new competencies. The research concludes that this would lead to platform companies requiring a set of competencies that would be similar to that of credit companies.

Competencies to support future needs of the companies. The credit fintech C-level experts admit that overall knowledge of all business domains, which is not limited to just one single niche, will be a key competency in the future. The credit and platform fintech experts believe that data analytics will be a future competency. The technical experts note that entrepreneurs who are able to use open source technical solutions to build new businesses are a future competency. HR experts see that the ability to work with speed and in a dynamic work environment will be important. The experts consider that mathematics, statistics and data analysis together with emotional intelligence will be a competency of the future.

The research results argue that DevOps approach has changed the technology department approach by merging development and infrastructure teams into one team. The authors notes that technology teams will merge with business teams. Results points out that there will be a joined competency set where technology people will understand business processes and business people will be able to use open source technology solutions to address business needs.

D. Financial data analysis

The authors perform financial data analysis to compare effectiveness of competency management by platform and credit fintech companies. The authors state that effective competency management leads companies to more effective operations and gives a competitive advantage. Most of the fintech companies have been operating in Latvia only starting from 2015; therefore, the only available periods for the analysis are 2016 and 2017. To compare different business types and evaluate their competitiveness, the authors choose to measure effectiveness of the companies by calculating turnover per employee (Fig. 1).

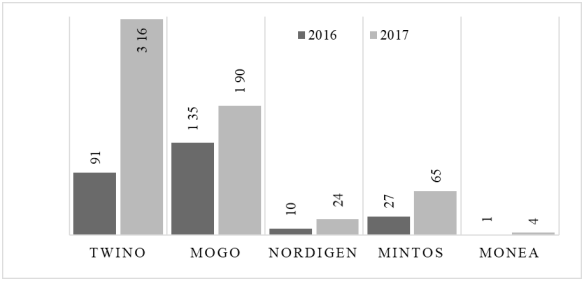


Fig. 1. Turnover per employee of platform-based companies, thousands EUR.

The research analyses the calculated results and evaluates which of the platform companies tend to improve effectiveness, and this can be related to different competency management models (Fig.2).

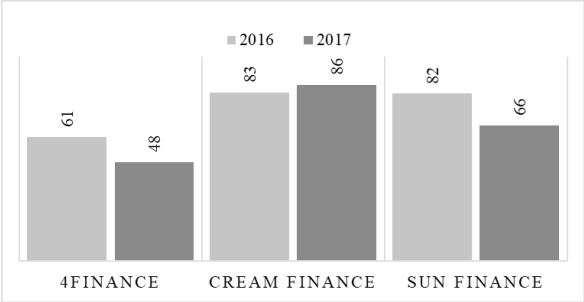


Fig. 2. Turnover per employee of credit-based companies, thousands EUR

Results points out that usually platform companies have an ecosystem that is much simpler than that of credit companies. During the expert interviews, researchers have noticed that the HR function of the platform companies is typically managed by CEO and there is no legal department, while credit companies have set up all the competencies. Research team have discussed insource or outsource competencies with the experts. All the experts agree that all core competencies will be insourced as intellectual capital will belong to the company. In the case of fintech companies, missing competencies outsourcing can be regarded as a short-term solution and not a strategic decision to maintain exceptional performance. Although knowledge-intensive fintech companies focusing on insource core competencies, fintech companies do not maintain databases of represented competencies and utilized skills, do not analyse competency gaps and do not monitor future needs. Moreover, there is little awareness about returns on investments in human capital, and the optimal structure of employed competencies that would identify overemployment points and productivity effectiveness is not analysed.

E. Competence model

As a result of the expert interviews, the paper summarizes the identified competencies (Table II) based on the technology management competencies model proposed by [1]. Moreover, the experts point out that

technical people lack business knowledge and business people need to have awareness about technical processes.

TABLE II. Competency management model

	Main Elements			
	Process	***Project***	***Systems***	***Operation***
sub elements	Self-Management	People Management	Quality Management	Risk Management
	Ambition Communication Self- training and development	Emotional intelligence Minimum viable product understanding Technology awareness	Sustainable system development UX analysis	Monitoring and measurement Data analysis Credit risk management Compliance

F. Hypotesis testing

To quantify the formulated hypothesis, the researchers selected the same sample of fintech companies as used for qualitative research for hypothesis testing and divided the selected fintech companies into platform and credit companies and calculated changes in turnover per one employee based on audited financial statement data for the financial years 2016 and 2017. Selected the one-way analysis of variance (abbreviated ANOVA) technique to test the formulated hypothesis. The ANOVA is a statistics technique based on an independent random sample to compare means between one or more populations [13].

Through hypothesis testing, the research seeks to compare average turnover per one employee yields between platform and credit companies, for which purpose data from the samples are summarized by their means and the research assumes that μ1 and μ2 are means corresponding to population and considers testing H0 (1) against H1 (2).

It is assumed under the null hypothesis that both group means are equal and that there is no difference between platform and credit companies (1).

$$H0: \mu 1 = \mu 2 \quad (1)$$

where μ is the mean value of the selected groups.

Under the alternative hypothesis, it is assumed that both group means are not equal and there is no equality between platform and credit companies (2).

$$H1: \mu 1 \neq \mu 2 \quad (2)$$

Research considers relying on three assumptions to test results if there is a difference between two industries:

1) Both population variances are equal;
2) Responses for both groups are sampled independently;
3) Response variable residuals are normally distributed.

To execute the defined assumptions and related calculations for hypothesis testing, the research has used R Studio and performed the following steps:

1) Installed and loaded R packages 'plyr', 'magrittr',' dplyr' and 'readxl'into R Studio;
2) Imported data into R Studio from Excel table;
3) Checked imported data quality;
4) Performed hypothesis testing calculations.

To test the assumption that both population variances are equal, research team makes p-value or probability calculation. The calculated p-value or probability that both industry variances are equal is 0.0044669. Statistics researches refer to $P<0.05$ as statistically significant and consider 95% as family-wise confidence level, therefore researchers reject H0 hypothesis and concludes that platform and credit companies are different.

To test homogeneity of variance between both groups, research team calculates F-value using Levene's Test. The calculated F-value is 0.6193. Statistics researches refer to a F-value >10% to consider a result as statistically significant. Based on the result of the calculation, the authors of the research cannot conclude that variances between both groups are not homogenous.

The final step is to test normality of platform and credit companies. Research team performs Shapiro-Wilk test of normality using the R function "shapiro.test()". From the output of calculations, the p-value is 0.5192 and if p-value >0.05, and normality can be assumed, implying that data of both groups are not significantly different from normal distribution.

The hypothesis test demonstrates that platform and credit companies tend to have different competency management models. Research rejects the null hypothesis and accepts the alternative hypothesis.

III. Conclusions

Companies with intensively applied technology capital experience faster productivity growth. It is crucial for companies to identify and then regularly assess which individual level competencies are required to achieve its strategic goals. Technology and knowledge intensive companies often require competencies that are not available in the market; therefore, contemporary companies are forced to ignore qualifications and recruit employees with personality traits to fit the position. A great variety of modern technical tools, frameworks and solutions are available owing to continuous technology development and technical and business managers are striving to get familiar with as much tools and technologies as possible, but in many cases, technologies are quickly replaced with new solutions due to the rapidly changing environment and continuous learning needs to be applied. Competency management must be tied to company's strategies. Properly assigned competences to new product or service creation positively contribute to overall business development. Knowledge-intensive fintech companies focusing on insource core competencies. Fintech companies do not maintain databases of represented competencies and utilized skills, do not analyse competency gaps and do not monitor future needs. Currently fintech companies are filling competency gaps by overrecruiting or assigning multiple roles to the same employee what leads to productivity decrease, affects competitiveness. Paper recommends further research in the area: (1) to

set types of the fintech companies to define fundamental competencies per fintech type; (2) to research fintech competency management from the sustainability aspect; (3) to continue to assess digital transformation impact on competency management in retail banking, identify retail banking fundamental competencies and define financial sector competency model.

Acknowledgments

Research developed with the support by BA School Business and Finance Internal Grant Program 2018.

References

[1] M. Doggett, P. McGee, P., and Scott, S. "Toward a Technology Management Core: Defining What the Technology Manager Needs to Know, " Technology Interface International Journal, vol.14, pp.70-79, Fall/Winter 2013.

[2] CSB, " In the 3rd quarter of 2018, Latvian employment rate constituted 65.3 %," October 2018. [Online]. Available: https://www.csb.gov.lv/lv/statistika/statistikas-temas/socialie-procesi/nodarbinatiba/meklet-tema/2423-nodarbinatiba-2018-gada-3-ceturksni [Accessed: Feb. 12, 2019].

[3] World Economic Forum, "The Global Competitiveness Report 2017–2018," December 2018. [Online]. Available: http://www3.weforum.org/docs/WEF_The_future__of_financial_services.pdf [Accessed: Feb. 12, 2019].

[4] M. Bontis, "Intellectual capital and performance within the banking sector of Luxembourg and Belgium," Journal of Intellectual Capital, vol. 14 pp. 286-309, 2012, https://doi.org/10.1108/14691931311323896

[5] C. Hawkes and B. Weathington, "Competency-Based Versus Task-Based Job Descriptions: Effects on Applicant Attraction," Journal of Behavioral and Applied Management, vol.15, pp.190-211, May 2014.

[6] W. Koziol, "A human capital measurement scale," Journal of Intellectual Capital, vol. 18, pp.316-329, 2017, https://doi.org/10.1108/JIC-08-2016-0085

[7] P. Gomber, R.J. Kauffman, C. Parker, and B.W. Weber, "On the Fintech Revolution: Interpreting the Forces of Innovation, Disruption, and Transformation in Financial Services," Journal of Management Information Systems, vol. 35, pp.220-265, January 2018.

[8] J. Kansal and S. Singhal, "Development of a competency model for enhancing the organisational effectiveness in a knowledge-based organisation," International Journal Of Indian Culture And Business Management, vol.16, pp. 287-301, 2018, https://doi.org/10.1504/IJICBM.2018.090909

[9] R. Colomo - Palacios, C. Casado - Lumbreras, P. Soto - Acosta, F. Garcia - Penalvo, and E. Tovar - Caro, "Competence gaps in software personnel: A multi-organizational study," Computers in Human Behavior, vol.29, pp.456-461, March 2013, https://doi.org/10.1016/j.chb.2012.04.021

[10] M. Bohlouli, N. Mittas, G. Kakarontzas, T. Theodosiou, Theodosios, L. Angelis, and M. Fathi, "Competence assessment as an expert system for human resource management: A mathematical approach," Expert Systems with Applications, vol.70, pp.83-102,March 2017, https://doi.org/10.1016/j.eswa.2016.10.046

[11] J.W. Creswell, Research design. Qualitative, Quantitative and Mixed Methods Approaches. London: SAGE Publications, Inc., 2014.

[12] Mintos, "Statistics," December 2019. [Online]. Available: https://www.mintos.com/en/statistics/ [Accessed: Feb. 12, 2019].

[13] W. Yu, H. El Barmi, and Z. Ying, "Restricted one-way analysis of variance using the empirical likelihood ratio test," Journal of Multivariate Analysis, vol. 102, pp.629-640, 2011, https://doi.org/10.1016/j.jmva.2010.11.006

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 104-108

Forecasting Missing Data Using Different Methods for Road Maintainers

Jānis Pekša
Institute of Information Technology
Faculty of Computer Science and
Information Technology
Riga Technical University
Kalku street 1, Riga, Latvia
Janis.Peksa@rtu.lv

***Abstract*—Observations collected from meteorological stations that are available to road maintainers and used for experimental purposes in this paper. Unfortunately, these observations are insufficient to make good forecasting that is needed for road maintainers. Those meteorological stations are located next to the road surface in the territory of the Republic of Latvia. The road maintainers can make forecasting using this data what is needed for the winter months. It is up to the road maintainers in winter months to process decision-making on road surface smudging with anti-slip chemical materials. The missing data in each meteorological station exists from time to time. The paper represents the possibility of using several approaches to fill out these missing data. This process is needed to be more accurate in predicting specific parameters aggregated from meteorological stations. These approaches are compared between the three closest meteorological stations available in the Republic of Latvia. The relevant data are for the winter months of 2017-2018. To conclude which is more accurate with VAS "Latvijas valsts celi" data set.**

***Keywords*— missing data, time-series, forecasting.**

I. Introduction

Meteorological stations for road maintainers today's data is needed without missing observations collected from several stations located in one region, such as the Republic of Latvia, where meteorological measurements are recorded. Using these complete data, conclusions and decisions can be drawn by observations with the same spacing, where system statistics information is stored sequentially [1]. However, the missing information in the time-series of meteorological stations is unavoidable, owing to the full observation of all the continuous processes is almost impossible [2]. If observation stops for any reason, then there is a problem with observations, in other words, if time-series data are missing they need to be filled for incomplete data. Information flows often result in missing data for many reasons, including sensor failures in meteorological stations, recording of observations errors and network lags [3]. Whereas the weather series contains total time and space characteristics, reconstruction of the missing period needs to be done carefully without characteristics of time-series statistics for interference. Widely for this purpose, the method used in the literature is the average value calculating the value in the time-series [4]. There are other methods which also considers the fleeting behavior of the time-series.

Currently, Latvia has 30 actives of 52 total meteorological stations near the road surface. However, a significant number of these stations suffering from incomplete data. Also, some time-series are suffering from a problem of inhomogeneity. These problems affect not only in the Republic of Latvia but also for the other countries' collecting meteorological time-series data [1]. Therefore, working with meteorological data must cope with these problems before any analysis is carried out. Road maintenance is very urgent in Latvia directly during the winter months, when the country roads are treated with anti-slip material. For road maintainers, forecasting from meteorological station data would allow for better decision-making, whether there is a need for anti-slip materials at a section of the road surface in future. The maintainers to make a clear-cut on the decision; therefore, requires a series of actions. Time-series should consider a specific arch. The road maintainers need predictions in the near future, which is aimed 30 minutes ahead [5]. First, there needs to be handled missing data problem. A few methods can be applied like:

- Normal ratio method – estimated by weighing the data points at various meteorological stations by the ratios of normal annual observations [6];
- k-nearest neighbors algorithm – the closest neighbor algorithm (k-NN) is a non-parameter method used for classification and regression. The input is made up of examples of the nearest training in the function space [7];
- Multilayer perceptron neural network – a class of feedforward artificial neural network. A multilayer perceptron neural network consists of at least three layers of node: the input layer, hidden layer, and the output layer. Except for the input nodes, each node is a neuron that uses a non-linear activation function [8].

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4120
© 2019 Jānis Pekša.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

Filling in missing data with existing relevant methods can continue with the next step. Second, choose the particular method due to it is most effective in predicting missing data observations. Accuracy is different, and it is fluctuating mainly of time-series windowing. Finally, the method applied can fill in missing data for meteorological stations that have missing data, forming indirect observations from VAS "Latvijas valsts celi" data set. Road maintainers can carry out future decisions by merging the observation and forecast data in their forecasting models.

The objective of this paper is to fit in the missing periods of measurement data from road surface maintenance meteorological stations dataset then requirements for the forecasting model set-up.

The paper is structured as follows. Section 2 background, Section 3 forecasting approach. Section 4 concludes.

II. Background

In order to meet the objective is the need for mathematical formulas and methods that are available in the literature. The meteorological stations in the Republic of Latvia are shown in Fig. 1.

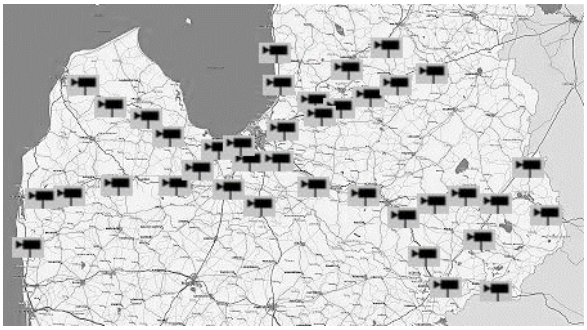


Fig. 1. Stations next to the road surface.

This paper used dataset consists of the destination station, indicated as *M*, and reference station group of all data set denoted as *Y*. The aim of the station is the one whose *x* observation(s), Y_{Mx+1}, Y_{Mx+2}, Y_{Mx+3}, …, Y_{Mx+n}. It is assumed that there is a lack of Y_{Mnone}, and they are needed to be estimated. However, there is a lack of value which is missing and needs to be filled. Thus, the missing value data capture period is artificially created. It is also assumed that the missing observations sequentially succeed one another if there are more than one of the following observations [9]. The reference stations are those that surround the target, with similar traits. Usually, for this purpose, it is used a nearby station, which has a high correlation with the objective[10]. Presuming that *M* reference stations and purpose, *t*, and if the same number of observations and including missing, *m* is recorded at all stations, the study used datasets terminology and the layout shown in Table 1.

TABLE I. The terminology of imputation methods

Denotation	Description
M	destination station
Y	all data set
x+1, x+2, …, x+n	observations
Y_{Mx+1}, …, Y_{Mx+n}	all data set with the destination station in each observation
m_1, m_2, …, m_n	missing observations
t	the target station
$m_k t$	subscript to represent missing observations recorded at the target (k=1, …, m)
$Y_{Mnone} = Y_{Mx+t}$, …, Y_{Mx+n}	missing observations at the target station, t

The following methods described below:

- Normal ratio method;
- k-nearest neighbors algorithm;
- Multilayer perceptron neural network.

A. Normal ratio method

Normal ratio method is used when the normal annual precipitation at any of the index station differs from that of the interpolation station by more than 10%. In this method, the precipitation amounts at the index stations are weighted by the ratios of their normal annual precipitation data in a relationship of the form[6], precipitation with dew point is replaced as follows:

$$P_m = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{N_m}{N_i}P_i\right) \quad (1)$$

- Pm – value at the missing location;
- Pi – value at meteorological index station;
- Nm – average annual observations at "missing data" gauge;
- Ni – average annual observations at gauge;
- n – some gauges.

Using procedure from reciprocal inverse weighting factor approach [11]:

- Divide area around gauge of interest into three parts;
- Using entries at the nearest station on each quadrant;
- Compute the missing value amount, where:

$$P_m = \frac{1}{\sum_{i=1}^{3} 1/X_i}\left(\sum_{i=1}^{3}\frac{P_i}{X_i}\right) \quad (2)$$

P_i – observations recorded by gauge *i*;

X_i – distance from gauge *i* to missing data point.

Normal ratio method is used to be able to compare with other meteorological stations in the reliability of data and their weight; the spreadsheet is used for this method calculations [12].

B. k-nearest neighbors algorithm

The k-nearest neighbors' algorithm (k-NN) is a non-parameter method used for classification and regression. The input is made up of examples of the nearest training in the function space. In the classification phase, *k* is a user-defined constant, and an unlabeled vector (query or test point) is classified by assigning a label that is most often between *k* training samples that are closest to this query point.

Usually, the distance metric used for continuous variables is the Euclidean distance [13]. The algorithm is as follows: Firstly, load the training and test data. Next, choose the value of *K*. Finally, for each point in test data a few steps need to be done: a) find the Euclidean distance to all training data points, then b) store the Euclidean distances in a list and sort it, then c) choose the first *k* points, in the end, d) assign a class to the test point based on the majority of classes [14]. For regression, k-NN predictions are the average of the k-nearest neighbors' outcome by:

$$y = \frac{1}{K}\sum_{i=1}^{k} y_i \tag{3}$$

where is the *i*th case of the examples sample and *y* is the prediction (outcome) of the query point. In contrast to regression, in classification problems, k-NN predictions are based on a voting scheme in which the winner is used to label the query. The distance weighting is formulated:

$$W(x, p_i) = \frac{\exp(-D(x,p_i))}{\sum_{i=1}^{k}(-D(x,p_i))} \tag{4}$$

Set of weights *W*, one for each nearest neighbor, determined by the relative closeness of each neighbor to the query point. Where is the distance between the query point *x* and the *i*th case of the example sample [15].

The application of the corresponding algorithm can replace the missing data from the meteorological station dataset. Since the region in question is a vast territory of the Republic of Latvia with several meteorological stations located at 64,589 km^2 of the total territory of the country [16]. The k-nearest neighbors' algorithm helps to get higher certainty comparison with other methods. The next method, which is viewed is the multilayer perceptron neural network.

C. Multilayer perceptron neural network

Multilayer perceptron (MLP) neural network uses a supervised learning method called backpropagation training. Its multilayer and nonlinear activation distinguish MLP from linear perception. It can differentiate between non-linear data.

Learning occurs in the perception by changing the weights after each of the expected results. Generalization of the least mean squares algorithm in the linear perception represent the error in output node *j* in the *n*th data point by [18]:

$$e_j(n) = d_j(n) \quad y_j(n) \tag{5}$$

where *d* is the target value and *y* are the value produced by the perception. The weight of the node is adjusted based on the corrections that reduce the error across the output that is indicated:

$$(n) = \frac{1}{2}\sum_j e_j^2(n) \tag{6}$$

With gradient descent, each weight change is:

$$\Delta w_{ji}(n) = \eta \frac{\partial \varepsilon(n)}{\partial v_j(n)} y_i(n) \tag{7}$$

where *y* is the output of the previous neuron and *n* is the training speed chosen to ensure that the scales quickly converge with the response, without fluctuations. To change the weight of a hidden layer, the weight of the output layer changes according to the activation function derivative and this algorithm denotes the activation function backpropagation [17]. Using the training can control the reliability of the data and provide for partial predictions data credibility.

The next chapter outlines the approach to the forecasting for the VAS "Latvijas valsts celi" dataset.

III. Forecasting approach

Initially, to understand the distance between meteorological stations, they can be combined in a single location, represented in Table 2.

TABLE II. The distance of the nearest station

Station name	Location (°)	Distance (km, name)
Adazi	57.0903, 24.3133	10.98 (Garkalne)
Admini	56.5931, 25.5464	36.77 (Livani)
Annenieki	56.6671, 23.0608	36.71 (Kalnciems)
Apvedcels	56.8436, 24.0192	12.84 (Dalbe)
Dalbe	56.7511, 23.8937	12.84 (Apvedcels)
Daugavpils	56.7511, 26.6438	68.90 (Kraslava)
Durbe	56.5752, 21.2939	37.10 (Rudbarzi)
Garkalne	57.0688, 24.4900	10.98 (Adazi)
Inciems	57.2741, 24.8949	12.53 (Sigulda)
Kaibala	56.6608, 24.8981	36.31 (Saulkalne)
Kalnciems	56.8360, 23.5761	10.53 (Laci)
Kraslava	55.8976, 27.2933	68.90 (Daugavpils)
Laci	56.8945, 23.7118	7.07 (Sloka)
Livani	56.4365, 26.0723	35.54 (Nicgale)
Ludza	56.3850, 28.0688	50.49 (Rezekne)
Melturi	57.2212, 25.2308	16.90 (Sigulda)
Nica	56.2145, 21.1329	41.38 (Durbe)
Nicgale	56.1551, 26.3433	35.54 (Livani)
Rezekne	56.5401, 27.2991	50.49 (Ludza)
Rudbarzi	56.6452, 21.8846	25.02 (Saldus)
Saldus	56.6710, 22.2898	25.02 (Rudbarzi)

Station name	Location (°)	Distance (km, name)
Saulkalne	56.8465, 24.4101	23.86 (Apvedcels)
Sigulda	57.1689, 24.9683	12.53 (Inciems)
Sloka	56.9295, 23.6150	7.07 (Laci)
Smiltene	57.3881, 25.9739	40.91 (Valmiera)
Stalbe	57.3781, 25.0512	14.93 (Inciems)
Strenci	57.6512, 25.8638	36.08 (Valmiera)
Talsi	57.1522, 22.6902	58.55 (Annenieki)
Valmiera	57.5035, 25.3271	21.67 (Stalbe)
Vircava	56.6401, 23.7763	14.30 (Dalbe)

The average value between the distances nearest to all distances between meteorological stations is 29.09 km. As well as the fact that standard deviation is 17.95 km not particularly surprising because a large proportion of meteorological stations do not work. Marking the smallest and largest distance of 68.90 km largest and the smallest is 7.07 km. As already acknowledged to predict from the meteorological station, the higher the distance, the more inaccurate the forecast, while splitting distances between meteorological stations, may reduce the RMSE when forecasting. In other words, windowing allows to take meteorological station time-series data and transform it into a cross-sectional format. When essentially convert time-series values into cross-sectional attributes, therefore, apply a predictive modeling algorithm to predict future values.

For experiment purposes, observations are taken between three meteorological stations that are close to each other, compared to all meteorological stations. One of the three meteorological stations will be adopted as a source of missing values, which will allow the use of the methods mentioned above to obtain missing values, or their forecasts, between two existing meteorological stations. In other words, methods will be used and then compared results with actual observations. In this way, they are getting the RMSE for each method. The interval used is 5 minutes, which consists predominantly of data entering between stations at 5 minutes interval [19]. The selected meteorological stations' names and coordinates are "Inciems" (57.2741, 24.8949), "Stalbe" (57.3781, 25.0512) and "Sigulda" (57.1689, 24.9683) located average 13.33 km from each other.

Of these meteorological stations, 392 observations are taken average 130 observation per meteorological station for a given parameter that is a dew point (C°) on winter month 15th of January 2018. The missing data meteorological station "Stalba" with 131 observations for this period is selected; part of the data is represented in Table 3.

TABLE III. "Inciems", "Stalbe" and "Sigulda" partly observation of dew point (C°) on 15th of January 2018

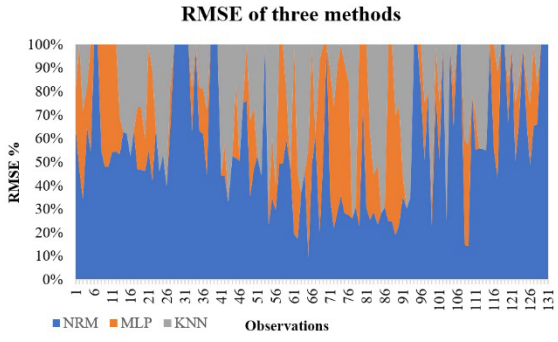
Station name	Dew point (C°)	Date time
…	…	…
Sigulda	-12.1	15.01.2018 11:41
Inciems	-10.6	15.01.2018 11:41
Stalbe	-9.9	15.01.2018 11:51
Inciems	-10.7	15.01.2018 11:52
Sigulda	-12.3	15.01.2018 11:53
Stalbe	-10.0	15.01.2018 12:02
…	…	…

The missing data is calculated using the following methods: Normal ratio method using spreadsheet for calculation, on the other hand, k-nearest neighbors algorithm is made up of examples of the nearest training in the function space using Orange Visual Programming tool (version 3.18.0) for prediction and RMSE estimation and next method MLP neural network is used with parameters neurons in hidden layers 100 then activation function for the hidden layer called "ReLu" the rectified linear unit function and finally solver called "SGD" stochastic gradient descent[20]. RMSE for each method is calculated the results are represented in Table 4.

TABLE IV. Accuracy for each method

Method	RMSE
Normal ratio method	0.35
k-nearest neighbors algorithm	0.23
MLP neural network	0.21

For Normal ratio method, RMSE is 35%, on the other hand, the k-nearest neighbors' algorithm is 23%, and finally, the MLP neural network is 21% accurate. Between k-NN and MLP is only 2% difference in accuracy. The results are also displayed graphically in order to be more readily perceived. Fig. 2 represents all three methods for the period 15th of January 2018.



15th of January 2018 RMSE for three methods of missing data prediction (dew point).

As shown in the results, the MLP neural network method has proved to be best compared to both methods.

IV. Conclusions

Meteorological stations will be required for methods that will be able to predict missing data. One of the most important reasons is sensor failures and potential interference. As well as a long distance between meteorological stations. Which makes long distances and makes significant errors in calculations, and forecasts with a high probability of error. There are a few methods to tackle this problem. Each time-series has its best or most efficient method, which can be used in practice. Before making any future forecasts, it is necessary to verify the use of the best method.

One of the well-known methods is multilayer perceptron neural network which was able to prove between the three methods as the best — presenting good results with 21% accuracy, which is relatively good but not enough. In the future, there is a need for better methods to show accuracy at the 10% mark. The road surface maintenance work in the winter months is compelled to minimize the response time on arrival at the specified road stage. The method used in the literature is the average value calculation.

Thus, the paper identifies further research direction on the generalized hybrid method on real-time forecasting for time-series. Compared to existing results and using a repeating method that can process real-time data and be able to adapt to the current situation, using modern solutions with programming language capabilities.

References

[1] Yozgatligil, C., Aslan, S., Iyigun, C. and Batmaz, I., 2013. Comparison of missing value imputation methods in time series: the case of Turkish meteorological data. *Theoretical and applied climatology*, 112(1-2), 143-167, https://doi.org/10.1007/s00704-012-0723-x

[2] Jeffrey, S.J., Carter, J.O., Moodie, K.B. and Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environmental Modelling & Software*, 16(4), 309-330.

[3] Spadavecchia, L. and Williams, M., 2009. Can spatio-temporal geostatistical methods improve high resolution regionalization of meteorological variables?. *Agricultural and Forest Meteorology*, 149(6-7), 1105-1117.

[4] Dibike, Y.B. and Coulibaly, P., 2005. Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. *Journal of hydrology*, 307(1-4), 145-163, https://doi.org/10.1016/j.jhydrol.2004.10.012

[5] Taylor, J.W., De Menezes, L.M. and McSharry, P.E., 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, 22(1), 1-16, https://doi.org/10.1016/j.ijforecast.2005.06.006

[6] Subramanya, K., 2013. Engineering Hydrology, 4e. Tata McGraw-Hill Education.

[7] Dudani, S.A., 1976. The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics*, (4), 325-327.

[8] Haykin, S.S., 2009. Neural networks and learning machines/Simon Haykin. New York: Prentice Hall.

[9] Giannone, D., Reichlin, L. and Small, D., 2008. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4), 665-676.

[10] Bagheri, M., Nurmanova, V., Abedinia, O., Naderi, M.S., Naderi, M.S. and Ghadimi, N., 2018, June. A Novel Wind Power Forecasting Based Feature Selection and Hybrid Forecast Engine Bundled with Honey Bee Mating Optimization. In *2018 IEEE International Conference on Environment and Electrical Engineering (EEEIC/I&CPS Europe)*, 1-6.

[11] Sharp, J.J. and Sawden, P.G., 2013. BASIC hydrology. Elsevier.

[12] Cuypers, W., Van Gestel, N., Voet, A., Kruth, J.P., Mingneau, J. and Bleys, P., 2009. Optical measurement techniques for mobile and large-scale dimensional metrology. *Optics and Lasers in Engineering*, 47(3-4), 292-300, https://doi.org/10.1016/j.optlaseng.2008.03.013

[13] Fabbri, R., Costa, L.D.F., Torelli, J.C. and Bruno, O.M., 2008. 2D Euclidean distance transform algorithms: A comparative survey. *ACM Computing Surveys* (CSUR), 40(1), 2.

[14] Acuna, E. and Rodriguez, C., 2004. The treatment of missing values and its effect on classifier accuracy. In *Classification, clustering, and data mining applications*. Springer, Berlin, Heidelberg, 639-647, https://doi.org/10.1007/978-3-642-17103-1_60

[15] Barrientos, R.J., Gómez, J.I., Tenllado, C., Matias, M.P. and Marin, M., 2011, August. kNN query processing in metric spaces using GPUs. In *European Conference on Parallel Processing*. Springer, Berlin, Heidelberg, 380-392, https://doi.org/10.1007/978-3-642-23400-2_35

[16] Gylfason, T. and Hochreiter, E., 2011. Growing Together: Croatia and Latvia. *Comparative Economic Studies*, 53(2), 165-197.

[17] Zhang, Z., 2018. Artificial neural network. In *Multivariate Time Series Analysis in Climate and Environmental Research*, 1-35, https://doi.org/10.1007/978-3-319-67340-0_1

[18] Murata, N., Yoshizawa, S. and Amari, S.I., 1994. Network information criterion-determining the number of hidden units for an artificial neural network model. *IEEE Transactions on Neural Networks*, 5(6), 865-872, https://doi.org/10.1109/72.329683

[19] Park, T. and Lee, S., 2004. A Bayesian approach for estimating link travel time on urban arterial road network. In *International Conference on Computational Science and Its Applications*, 1017-1025, https://doi.org/10.1007/978-3-540-24707-4_114

[20] Chaudhari, P. and Soatto, S., 2018. Stochastic gradient descent performs variational inference, converges to limit cycles for deep networks. In *2018 Information Theory and Applications Workshop (ITA)*, 1-10

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 109-113

Adaptive Kalman Filter Forecasting for Road Maintainers

Jānis Pekša
Institute of Information Technology
Faculty of Computer Science and
Information Technology
Riga Technical University
Kalku street 1, Riga, Latvia
Janis.Peksa@rtu.lv

***Abstract*—The article considers the road monitoring weather-stations which collects raw observations that are processed to be able to make the necessary forecasting for future decisions. For the road maintainers those predictions are crucial to make decisions daily. When it comes to the winter season when road safety is very important; however, the road condition is also affected by the snow and icing. In order to improve safety on the road network the road maintainers are trying to use every possible way to be able to provide it. A number of methods have been studied and compared to clarify the parameter required by Kalman filter, which can be improved by making forecasting more accurate. Several road monitoring weather-stations are merged into one region because they are relatively close to each other and it is assumed that there are common conditions in one region that may indicate changes in road conditions. The corresponding algorithms are applied for each region and then compared to each other. Adaptive Kalman filter is generalized in the relevant article in order to have a general understanding of how to correctly apply the approach. The main result of this article is a comparison with the different methods, which are finally compiled in a single table.**

Keywords*— *Kalman filter, noise covariances, time-series, forecasting.

I. Introduction

The road maintenance is a complex decision-making process that needs to be done within a specific time frame. Pastime may affect the essential operations of the road network in total. One of them is driving conditions when traffic speed is rapid. The higher the speed rate the greater the probability that accidents are possible [1]. As know, the road conditions are subject to rapid changes in surface temperature and precipitation sum. The road conditions most fluctuations are observed during the winter period; however, the road condition is also affected by the snow and icing. The road maintenance is performed for specific Road Sections belonging to a Region. The region is the Republic of Latvia road network and the dataset is taken from VAS "Latvijas valsts celi". The Region consists of 52 road monitoring weather-stations that are relatively distant from each other. To be able to respond to changes in the environment on road sections near the road surface are located road monitoring weather-stations and cameras operated by different entities. The road monitoring weather-stations collects raw observations which are processed to be able to make the necessary forecasting for future decisions. For road maintainers, those predictions are crucial to making decisions daily. The road monitoring weather-stations and cameras operated by different entities can help for the decision-making. The road maintainers controlling smart road signs are available to give warning messages to drivers on a specific stage of the road network [2]. For instance, driving conditions are evaluated according to surface temperature and precipitation sum provided by roadside weather-stations as well as weather service. However, the weather service provides data only at the regional level. In the road maintenance case, one of the adjustments evaluates a need to de-ice the roads [3]. The road conditions context element is computed using surface temperature (°C) and precipitation sum (mm). The warning adjustment is triggered only if the locking conditions are met preventing unnecessary nervousness. The road maintainers receive signals via interfaces from Enterprise Resource Planning Systems (ERPs) that have received a warning that one of the parameters has changed dramatically and reached a critical rule [4]. The critical rule depends on the definition of the road maintainers. To be able to predict future parameters forecasting comes in handy for this task.

As already known forecasting consists of pre-processing data and missing data should be completed to perform predictions successfully [5]. In the previous work, missing data are filled for the data set time-series parameters. To be effective prediction will be applied Kalman filter. The Kalman filtering, also known as linear quadratic estimation, is an algorithm that uses a series of measurements observed over time, containing statistical noise and other inaccuracies [6]. To be able to create Adaptive Kalman filter one of the parameters should be modified dynamically changing the noise parameter also known as the noise covariances. The noise covariances are found by those algorithms:

- Simple moving average – arithmetic moving

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4134
© 2019 Jānis Pekša.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

average calculated by adding recent closing prices and then dividing that by the number of periods in the calculation average [7];

- Arithmetic mean – the most common type of standard is the arithmetic mean, in other words, an average value [8];
- Bayesian average – a method of estimating the mean of a population using outside information [9];
- k-Means Clustering algorithm – a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining [10];
- Expectation-Maximization algorithm – an iterative method to find the maximum likelihood of parameters in statistical models representing the optimal weights [11].

The optimal weight algorithms are compared to each other finding the best algorithm for the Kalman filter.

The objective of this paper is to find the most suitable algorithm for the noise covariances to be applied for Kalman filter to make forecasting for few parameters of the weather-stations to trigger warnings signal for the ERP systems.

The paper is structured as follows. Section 2 background, Section 3 adaptive Kalman filter and Section 4 concludes.

II. Background

To meet the objective, the need for mathematical formulas and methods that are available in the literature are represented.

A. Simple moving average

A simple moving average (SMA) is an arithmetic mean of the variable calculated by adding the last closing prices and then dividing them by the number of time periods is calculated on average [7]. Most of the advantages:

- The SMA is a technical indicator for determining if an asset price will continue or reverse a bull or bear trend [7];

The SMA is calculated as the arithmetic average of an asset's price over some period [7];

- The SMA can be enhanced as an exponential moving average (EMA) that more heavily weights recent price action [7].

The SMA can be formulated as follows [7]:

$$SMA = \frac{A_1+A_2+\ldots+A_n}{n} \quad (1)$$

where A refers to the price of an asset at period n and n is the number of total periods. The price parameter is assumed that the price parameter is equal to the observation parameter in this case. A simple variable average is adjustable because it can be calculated for another time period by merely adding the end price of the security for several time periods and then dividing that total by the number of time periods indicating the average price of the security over the time period. A simple variable average smoothes out volatility and facilitates access to securities price trends. If the simple moving average points up, it means that the price of the security increases. If it points down, it means that the price of security is decreasing. The longer the average period, the simpler the ordinary variable. The shortest moving average is more volatile, but its reading is closer to the source data [7].

B. Arithmetic mean

The arithmetic mean is a sum of the numbers divided by how many numbers are being averaged, in other words, an average value. Formulation of the arithmetic mean can be represented as follows [8]:

$$AM = \frac{1}{n}\sum_{i=1}^{n} a_i = \frac{a_1+a_2+\ldots+a_n}{n} \quad (2)$$

where n numbers are given, each number denoted by (where $i = 1, 2, \ldots, n$), the arithmetic mean is the sum of the as divided by n.

C. Bayesian average

A Bayesian average is a method of estimating the average number of people who use external information, especially the existing beliefs that have been taken into account in the calculation. A Bayesian average method is described as follows [9]:

$$\overline{x} = \frac{C_m + \sum_{i=1}^{n} x_i}{C+n} \quad (3)$$

where if the previous mean m and constant C are used in the Bayesian average value calculation, a value is assigned that is proportional to the size of the typical data set. The value is higher if the expected difference between the datasets is small. This is less if the data sets are expected to differ significantly. This is equivalent to adding C data points of value m to the data set. It is the weighted average of the previous average m and the sample [9]. The disadvantage is that the Bayesian average can be used if the size of the dataset is small for the data selected by the experiment; they are small enough. This method can be disabled for further investigation.

D. k-Means Clustering algorithm

k-means clustering is a method for quantitative determination of vectors, initially from signal processing, which is accessible for cluster analysis in data mining. The purpose of k-means clustering is to divide n observations into k clusters where each observation belongs to a cluster with the closest mean value serving as a cluster prototype. k-means clustering can be described as a set of observations $(x_1, x_2, \ldots, x_n)$, where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into $k (\leq n)$ sets $S=\{S_1, S_2, \ldots, S_k\}$ so as to minimize the within-cluster sum of squares [10].

$$\arg\min_{S} \sum_{i=1}^{k} \sum_{x \in S_i} ||x \quad \mu_i||^2 = \arg\min_{S} \sum_{i=1}^{k} |S_i| Var S_i \quad (4)$$

where is the mean of points in S_i .

E. Expectation–Maximization algorithm

In statistics, an expectation–maximization (EM) algorithm is an iterative method to determine the maximum opportunity or maximum posterior (MAP) parameter estimation in statistical models, where the model depends on the latent variables. The EM iteration alternates with step *E* that creates a function to predict the probability of the log using the current estimation of the parameters, and the maximization *M* step, which calculates the parameters, maximizing the expected log probability found in step *E*. These parameter calculations are used to determine the distribution of latent variables in the next *E* phase [11].

Given the statistical model which generates a set X of observed data, a set of unobserved latent data or missing values Z, and a vector of unknown parameters θ, along with a likelihood function $L(\theta;X,Z)=p(X,Z|\theta)$, the maximum likelihood estimate (MLE) of the unknown parameters is determined by maximizing the marginal likelihood of the observed data [11]:

$$L(\theta;X) = p(X|\theta) = \int p(X,Z|\theta)dZ \quad (5)$$

F. The noise covariances algorithm comparison

To use the Kalman filter to determine the internal state of the process that has been assigned only a series of noisy observations, the process must be modeled according to the Kalman filter system. Kalman filters noise links consisting of covariance of Q process noise covariance and R surveillance noise. These parameters will be determined by algorithms and then compared. The Kalman filter has many materials and a lot of research on noise linking methods. In practice, Q and R are often considered as design variables. A common approach to choosing cravings, where Q is chosen as a diagonal weight matrix [6]. The above algorithms are used for comparison in this paper.

For experiment purposes, observations are taken between the closest three weather-stations that are close to each other and all the road monitoring weather-stations are separated by region $(R_1,R_2,...,R_n)$. In other words, algorithms will be used and then compared results with each other. In this way, they are getting the the Root Mean Square Error (RMSE) is used for this precise compensation. The RMSE is the standard deviation of the residuals prediction errors. The interval used is 5 minutes, which consists predominantly of data entering between stations at 5 minutes interval. The region has 30 actives of 52 total road monitoring weather-stations near the road surface, only active ones are considered. These weather-stations by region is combined as follows shown in Table 1 as the previous results indicate the nearest weather-stations.

TABLE I. The combined region from the road monitoring weather-stations

Region	Station name	Location (°)
R1	Adazi	57.0903, 24.3133
	Garkalne	57.0688, 24.4900
R2	Admini	56.5931, 25.5464
	Livani	56.4365, 26.0723
	Nicgale	56.1551, 26.3433
R3	Annenieki	56.6671, 23.0608
	Kalnciems	56.8360, 23.5761
	Laci	56.8945, 23.7118
	Sloka	56.9295, 23.6150
	Talsi	57.1522, 22.6902
R4	Apvedcels	56.8436, 24.0192
	Dalbe	56.7511, 23.8937
	Saulkalne	56.8465, 24.4101
	Kaibala	56.6608, 24.8981
	Vircava	56.6401, 23.7763
R5	Daugavpils	56.7511, 26.6438
	Kraslava	55.8976, 27.2933
	Zilupe	56.3850, 28.0688
	Rezekne	56.5401, 27.2991
R6	Durbe	56.5752, 21.2939
	Rudbarzi	56.6452, 21.8846
	Saldus	56.6710, 22.2898
	Nica	56.2145, 21.1329
R7	Inciems	57.2741, 24.8949
	Sigulda	57.1689, 24.9683
	Melturi	57.2212, 25.2308
R8	Smiltene	57.3881, 25.9739
	Valmiera	57.5035, 25.3271
	Stalbe	57.3781, 25.0512
	Strenci	57.6512, 25.8638

The regions consist from two till five road monitoring weather-stations located relatively near to each other. As a result, eight regions are created for the noise covariances algorithms calculation. Calculating results from each region use the average arithmetic approach. The road monitoring weather-stations observations per weather-station are taken into account surface temperature (°C) and precipitation sum (mm). Observations are taken from winter month 15th of January 2018, therefore, for each region representing surface temperature (°C) and precipitation sum (mm) average in each region are shown in Figure 1. Since the observation is much then cut out from 1:00 AM to 2:00 AM is reflected. As can be seen in two of the regions no precipitation sum has values only 0 mm. In the seventh region, the temperature has apparently fluctuated within one hour.

The corresponding algorithms are used for each region and their results are reflected in Table 2.

TABLE II. ACCURACY FOR EACH ALGORITHM

Region	Surface temperature (°C), RMSE				
	Simple moving average	Arithmetic mean	Bayesian average	k-Means Clustering	Expectation–Maximization
R1	0.0288	0.0502	0.1010	0.0382	0.0201
R2	0.3626	0.1052	0.2717	0.1612	0.0972
R3	0.3807	0.3525	0.3738	0.3727	0.3101
R4	0.3633	0.4372	0.3927	0.3693	0.3092
R5	0.5422	0.5577	0.5321	0.5216	0.4948
R6	0.8393	0.7994	0.7369	0.7745	0.6925
R7	0.6273	0.7065	0.6514	0.6426	0.5829
R8	0.4819	0.5016	0.4796	0.5127	0.4420

The Expectation–maximization algorithm has proven in this experiment that there is one of the most accurate compared to other algorithms. The Expectation–Maximization algorithm is used further with Kalman filter and is called "Adaptive Kalman filter", which will be able to take account of the existing algorithm factor and use it to make more accurate forecasts for the raw data observations. Each time a new portion of observations comes in from the road monitoring weather-stations, the corresponding algorithm will be called to calculate an already new parameter required by Adaptive Kalman filter to be more likely to make forecasts every time. The next step is to define the Adaptive Kalman filter, which needs to be tested in a long-term way with data set. The data set already in place to be able to determine the effectiveness of this approach, which is necessary for forecasting because, as the best way to predict, a number of data sources such as Waze, Twitter, and other useful resources are needed to face clarity and well results. Obtaining an algorithm that can be applied Kalman filter case will continue in the next section.

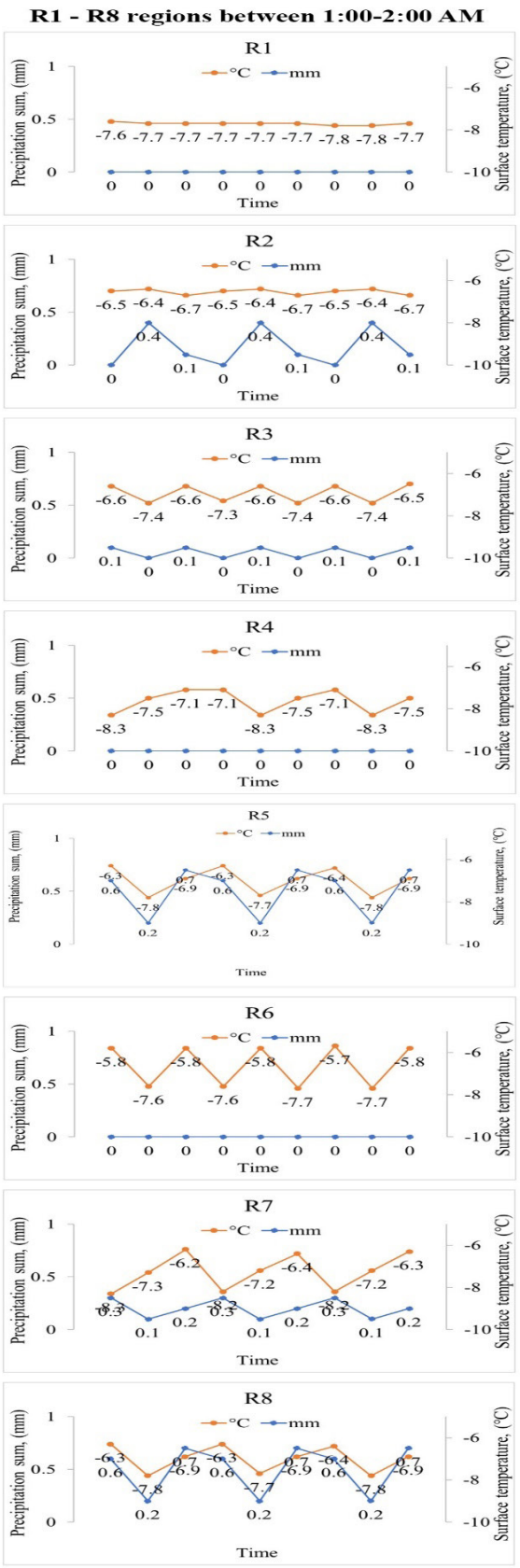


Fig. 1. R1 – R8 regions between 1:00 AM and 2:00 AM

III. Adaptive Kalman filter

The Expectation–maximization algorithm has proven in the experiment that it is best able to be more accurate for noise covariances algorithm comparison. The algorithm will be used for the Adaptive Kalman filter. It is initially necessary to define the Kalman filter to be able to work with it. Mathematical formulation of Kalman Filter as follows [6]:

$$x_k = Ax_{k-1} + Bu_k + w_{k-1} \tag{6}$$

the Kalman filter addresses the general problem of trying to estimate the state $x \in R^n$ of a discrete-time controlled process that is governed by the linear stochastic difference equation. With a measurement $y \in R^m$ that is [6]:

$$y_k = Hx_k + v_k \tag{7}$$

The random variables w_k and v_k represent the process and measurement noise respectively. They are assumed to be independent of each other, white, and with normal probability distributions: $p(w) \approx N(0,Q)$ and $p(v) \approx N(0,R)$. In practice process covariance Q and covariance of measurement noise R matrices can change with every step or measurement, but here they are constant n * n matrix. A refers to the state of the previous time move to the current stage if there is no driving function or process noise n * 1. The matrix B refers to optional control input state m * n of matrix H in the measurement equation refers to the state measurement y_k [6].

The Kalman filtering process has two stages: the forecasting step, where the next state. The system is predicted by previous measurements and the update phase where the current state of the system is assessed taking into account the measurement over a given time period. Actions are translated into equations as follows [6]:

- Prediction

$$X_k^- = A_{k-1} + X_{k-1} + B_k U_k \tag{8}$$

$$P_k^- = A_{k-1} P_{k-1} A_{k-1}^T + Q_{k-1} \tag{9}$$

- Update step

$$V_k = Y_k \quad H_k X_k^- \tag{10}$$

$$S_k = H_k P_k^- H_k^T + R_k \tag{11}$$

$$K_k = P_k^- H_k^T S_k^{-1} \tag{12}$$

$$P_k^- = P_k^- \quad K_k \; S_k \; K_k^T \tag{13}$$

where X_k^- and P_k^- are the predicted mean and covariance of the state, respectively, on the time step k before seeing the measurement. X_k and P_k are the estimated mean and covariance of the state, respectively, on time step k after seeing the measurement. Y_k is mean of the measurement on time step k. V_k is the innovation or the measurement residual on time step k. S_k is the measurement prediction covariance on the time step k. K_k is the filter gain, which tells how much the predictions should be corrected on time step *k* [6]. Adaptive Kalman filter has not been tested with experiments in this article.

IV. Conclusions

The road maintainers can analyze many sources to make decisions during the winter season to be able to serve and maintain the current road surface state more effectively. The road monitoring weather-stations allows collecting raw observations close to the road surface; allowing raw observations to be processed. To be able to predict these observations is needed. After processing the raw observations concerned, they may be used for forecasting purposes. Forecasting techniques are different types of algorithms and approaches. Studies and experiments that demonstrate the accuracy of the algorithm or approach are necessary to assess the best.

The following methods and algorithms are summarized in this article: Simple moving average, Arithmetic mean, Bayesian average, k-Means Clustering algorithm, and Expectation-Maximization algorithm. As well as an experiment with the VAS "Latvijas valsts celi" data set. The best of these algorithms is identified accordingly in this case the Expectation-Maximization algorithm shows the best RMSE. As well as the approach of Adaptive Kalman filter to make forecasting and apply the corresponding algorithm to improve forecasting results.

Further work requires an experiment with the corresponding Adaptive Kalman filter using the Expectation-Maximization algorithm with real-time data. Connecting the corresponding algorithm to raw observations and comparing it to the holding critical rules that have been defined by the road maintainers.

References

[1] [1] J. B. Edwards, "Speed adjustment of motorway commuter traffic to inclement weather," Transp. Res. Part F Traffic Psychol. Behav., vol. 2, no. 1, pp. 1–14, Mar. 1999, https://doi.org/10.1016/S1369-8478(99)00003-0

[2] [2] Y.-Y. Nguwi and A. Z. Kouzani, "Detection and classification of road signs in natural environments," Neural Comput. Appl., vol. 17, no. 3, pp. 265–289, Jun. 2008, https://doi.org/10.1007/s00521-007-0120-z

[3] [3] J. P. Gibbs and W. G. Shriver, "Can road mortality limit populations of pool-breeding amphibians?," Wetl. Ecol. Manag., vol. 13, no. 3, pp. 281–289, Jun. 2005.

[4] [4] C. Taylor, S. Pollard, S. Rocks, and A. Angus, "Selecting Policy Instruments for Better Environmental Regulation: a Critique and Future Research Agenda," Environ. Policy Gov., vol. 22, no. 4, pp. 268–292, Jul. 2012.

[5] [5] J. Honaker and G. King, "What to Do about Missing Values in Time-Series Cross-Section Data," Am. J. Pol. Sci., vol. 54, no. 2, pp. 561–581, Apr. 2010.

[6] [6] G. Bishop, G. W.-P. of SIGGRAPH, undefined Course, and undefined 2001, "An introduction to the kalman filter," byron.soe.ucsc.edu.

[7] [7] F. R. Johnston, J. E. Boyland, M. Meadows, and E. Shale, "Some properties of a simple moving average when applied to forecasting a time series," J. Oper. Res. Soc., vol. 50, no. 12, pp. 1267–1271, Dec. 1999.

[8] [8] E. Jacquier, A. Kane, and A. J. Marcus, "Geometric or Arithmetic Mean: A Reconsideration," Financ. Anal. J., vol. 59, no. 6, pp. 46–53, Nov. 2003.

[9] [9] J. Huber, K. T.-M. Letters, and undefined 2001, "On the similarity of classical and Bayesian estimates of individual mean partworths," Springer.

[10] [10] A. Likas, N. Vlassis, J. V.-P. recognition, and undefined 2003, "The global k-means clustering algorithm," Elsevier.

[11] [11] T. M.-I. S. processing magazine and undefined 1996, "The expectation-maximization algorithm," ieeexplore.ieee.org.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 114-118

Gamification Framework for Software Development Project Processes

Valerija Platonova
Information Technology Institute
Riga Technical University
Riga, Latvia
valerija.platonova@edu.rtu.lv

Solvita Berzisa
Information Technology Institute
Riga Technical University
Riga, Latvia
solvita.berzisa@rtu.lv

***Abstract*—Gamification methodology has a positive impact on software development (SD) processes, contributes to better product quality production and team involvement. To show how gamification can be used to motivate the SD project team to carry out daily routine activities and document it in PM tool the gamification framework is proposed in this paper. With this gamification framework is also tried to solve project manager challenge to get actual information in PM tool entered by the project team for correct reporting of project status and process overview. A prototype of a gaming tool has been developed, which is based on Jira's app functionality and is implemented as a plugin.**

***Keywords*—Software development project, gamification, project management, project team motivation.**

I. Introduction

Information technology (IT) software development (SD) project is a complicated process involving many participants, tools and processes [1]. Participants motivation and opportunity to work for the team is often one of the crucial moments for the successful completion of the project [2]. In SD project also is a lot of daily routine tasks that are not so interesting, but it needs to be done. One of the routine tasks is updating information in PM tool (progress, time spent) that is critical for a project manager to understand tasks and project progress. Timely available information can help the project manager correctly report status, identify risks, issues and changes necessary to achieve successful project results [3].

Many changes are made in corporate culture to build the best atmosphere for resources and increase motivation at work, e.g. working from home, flexible working hours. One change that is made is adding gamification element at work [4], [5] that allow changing process and behaviours [6]. Best practices of gamification say that the best way of using gamification in the corporate field is gamification tool usage [7], [8]. Also, in SD processes are some attempts to gamify some processes, e.g., development and testing [9].

The objective of the research is to develop a gaming framework for SD project with a target to motivate the SD project team to carry out daily routine activities and document it in PM tool. To identify requirements of the gamification framework has been reviewed the everyday life of the developer team and PM tool usage. A prototype of the gamification tool has been created based on the proposed framework and developed as a Jira plugin. The tool evaluation was carried out with focus group assessment.

The rest of the paper is structured as follows: Section 2 presents the research background of gamification and SD project daily tasks. Overview of the SD gamification framework is described in Section 3. The prototype of the gamification tool and its evaluation is given in Section 4. Conclusion and future work is presented at the end of the paper.

II. research background

As the background of this research is SD project team daily tasks (Section II.A.) that can be gamified to improve team motivation and data quality in PM tool and theory about gamification framework (Section II.B.).

A. SD project daily tasks

The tasks of software developers are to develop or maintain a software unit according to specific customer requirements, within the limits of time and budget [1]. SD process consists of certain stages, such as design, development, documentation, testing and maintenance, with the goal of developing the final product [1]. This process is defined for the chosen SD project lifecycle model that creates a difference in team composition, processes, work units, and routine tasks [1]. In general points, SD project team can be divided into three roles – programmers, analysts and testers and its main daily tasks have been summarised ([1], [10]policies, and operations of an organization, and to recommend solutions that enable the organization to achieve its goals. Business analysis involves understanding how organizations function to accomplish their purposes and defining the capabilities an organization requires to provide products and services to external stakeholders. It includes the definition of organizational goals, understanding how

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4142
© 2019 Valerija Platonova, Solvita Berzisa.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

those goals connect to specific objectives, determining the courses of action that an organization has to undertake to achieve those goals and objectives, and defining how the various organizational units and stakeholders within and outside of that organization interact. A Guide to the Business Analysis Body of Knowledge (BABOK Guide, [11]) in Table I.

Additional to daily tasks all project team, in the same way, is involved in PM activities [3]. The first, project planning [1] – estimating, tasks planning, risk identification etc. The second, project status reporting [1] – tasks progress, how many time is spent, problems etc. In most cases, some PM tool is used for status reporting. One of the project manager challenges is getting correct information from PM tools that is crucially dependent on the project team entered information. In the same time status updating in PM tool has been classified as one of the monotonous routine tasks for the project team.

TABLE I. SD PROJECT ROLE DAILY TASKS

Role	Activities
Programmer	Develop new functionality Testing Fix defects Code review
Analyst	Requirement definition and analyses Change request analysis and commenting Defect registering and analysis Testing Documentation writing Support for programmers/testers
Tester	Manual and automated testing Defect register Retesting of defects Test case writing

Gamification

Gamification is the use of game mechanics and elements in a non-game context [6]. The existing process is selected, and the game mechanics are used to motivate participation, increase engagement and loyalty [5]. To play the game, need to use certain elements of the game. Playing elements are divided into three components as indicated in Fig. 1 [12]:

- Dynamics - it can be described as the most important element of the game or "grammar" like emotional effects, a logic of events, a chronology of activities, an interaction between participants.
- Mechanics - these are the processes that prevent the game from moving forward - challenges, opportunities, co-operation and interactions, feedbacks, bonuses, game resources, etc.
- Components - these are the individual elements of mechanics and dynamics that interact, including points, roles, levels, avatars, winning tables, ratings, etc.

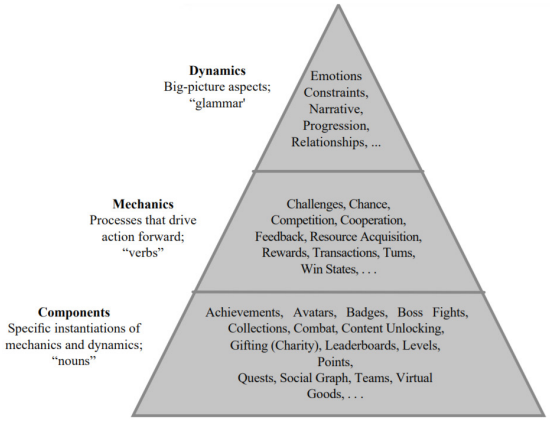


Fig. 1. SD gamification concept [12]

The main implementation of the gamification framework is tools that support it. On the product market, there are already some solutions that allow using gamification in the SD process: DevRPG [13], TEAMFEED [14], Catcher [5] and Gamiware [15]. Summary of the tools have been given in Table II.

TABLE II. SD GAMIFICATION FRAMEWORKS

Tool	Description
DevRPG [13]	***Goal:*** Gamification tool is to make the development process challenging and fun. The tool was embedded in the task management tool. ***Elements:*** Rewards, challenges, instant feedbacks, points, roles. ***Advantage:*** Good metrics to measure performance by promoting developer competition. The feedback from the team shows that everyone was impressed with their results and made the enthusiastic process more active. ***Disadvantage:*** Not mentioned.
TEAMFEED [14]	***Goal:*** Tool was developed by a team of researchers who researched a university course, testing the use of gamification in the development process, based on student loyalty. ***Elements:*** Points, leader-boards, comments, feedbacks. ***Advantage:*** Increases the amount of work done by developers, encourages more active work. ***Disadvantage:*** Not mentioned.
Catcher [5]	***Goal:*** The core functionality of the tool is to reward programmers for eliminating the code, which protects the functionality errors and improves the quality of the code. ***Elements:*** Points, reminders. ***Advantage:*** Developers are more motivated to take action for better code quality. ***Disadvantage:*** Big amount of reminders detract from work, tension during task execution
Gamiware [15]	***Goal:*** Designed to influence the ability of gamification to explore the motivation of a developer team. The main elements of the gameplay dynamics are the progress tracking and interpersonal relationships. ***Elements:*** Roles, tasks, feedbacks, leaderboards, points and levels. ***Advantage:*** Increasing motivation among participants, faster execution of tasks and better performance of tasks. ***Disadvantage:*** Hard process of implementation.

Components, tools, players, motivators together make gamification framework. Gamification framework

answers to questions: what is being gamified; why is it gamified; who are the users; how is it being gamified [16]. Multiple frameworks describe the design process [17]. One of them is 6D framework [18] with six elements or steps: define business objectives; define target behaviours; describe you players; devise activity loops; do not forget the fun; deploy appropriate tools.

III. Software Development Project Gamification framework

The SD project gamification framework is created with the target to motivate the project team to perform daily project tasks and correctly document tasks status in PM tool.

The framework has been created based on the design steps of gamification design frameworks and SD project team members daily routine tasks classified as boring. Gamification framework has been documented using the gamification canvas model [19]. Fig. 2 describes the main concepts of the SD project gamification framework.

The gamification framework main dynamic is to award with points user after he has performed activities with his assigned task (with other words, status change) according to his project role (programmer, analyst and tester). After a team member individually or team together collects the required amount of points, it can be rewarded with some prize (training or team building event).

Playing elements used in SD gamification framework:

1. Dynamics:
   a. progress – point for performed activities;
   b. relationships – team unity for goal achievement;
   c. limitations – points are given only performing activities related to role.
2. Mechanics:
   a. competition – leader board table with each team member points;
   b. feedbacks – user see his contribution to the project;
   c. prizes – achieving individual and team point goals can be exchanged with a prize;
   d. cooperation– only team working together team goal can be achieved;
   e. win status – user can see his contribution to project goal achievement.
3. Components:
   a. points – items user get for his activities. For each activity can be different points;
   b. leader boards – user ranking according to earned points;
   c. achievement – user goal achievement.

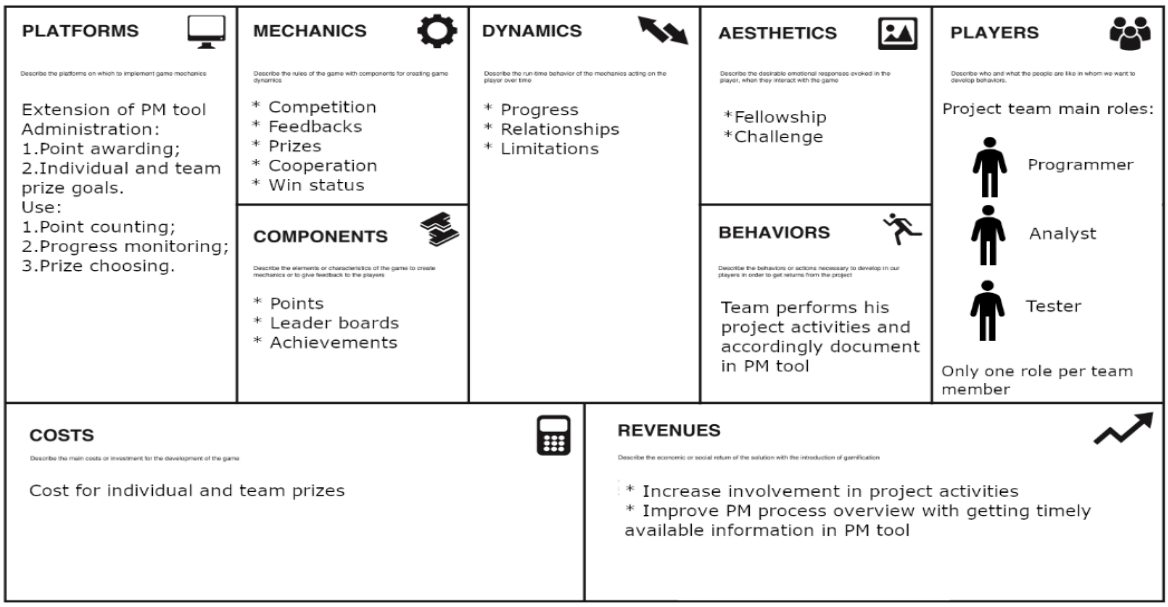


Fig. 2. SD gamification framework canvas model

- The tools that will support the SD gamification framework will be an extension of PM tool. Tool administration section needs to allow set up:
- Work items and workflow status changes performed by a defined role that are awarded to a defined number of points. Can be separately defined for individual goals and project team goals;

Individual and team prizes and their points goals.

IV. The tool from user and project perspective need to support points counting, progress monitoring and prize choosing after point goal is reached.

Application and evaluation

To demonstrate how gamification framework works it has been implemented as the Jira plugin. Gamification process has been connected to Jira project work item types, its workflow statuses and project roles.

- Settings for the SD project gamification tool:
- JIRA project work item types, e.g., user stories, bug,

task etc., according to the SD process.

- JIRA project workflow that defines work item statuses (e.g., to do, in progress, ready for QA, done, reopened, blocked) and status transactions (Fig.3 A). Need to be defined according to the SD process with correct statuses that the team need to do.
- Point configuration (Fig.3 B) where is defined points (individual and project) what each role earns by moving the task to a certain status.

Prize definition (Fig.3 C) where is configured individual and project prizes with point goals.

The user or team member earning points by doing works and updating Jira tasks status. The user can choose target prize from store (Fig. 3 D), monitor the progress of individual and project goal achievement (Fig. 3 E), and team member results in leader board (Fig. 3 F)

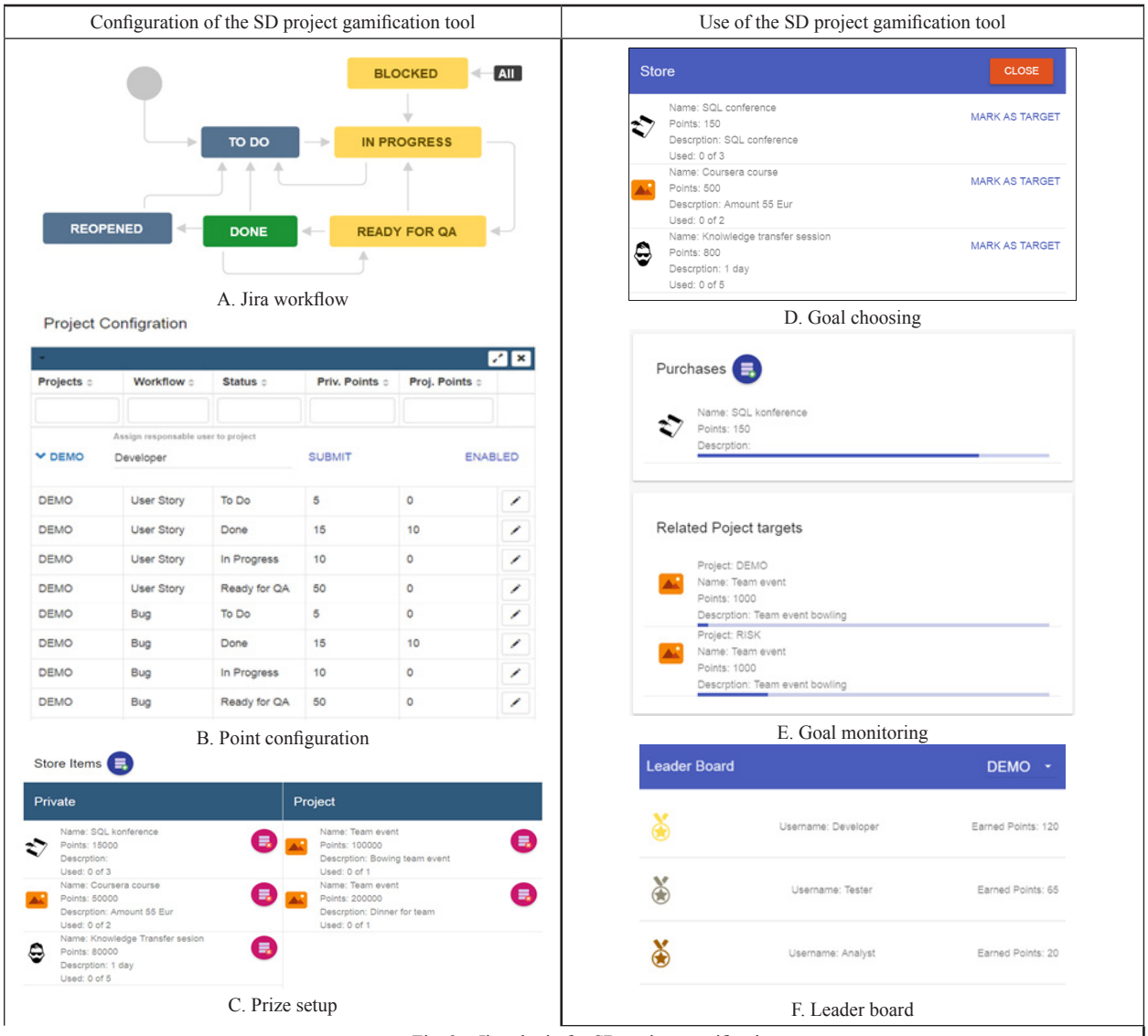


Fig. 3. Jira plugin for SD project gamification

Small focus group testing and survey have been organised to understand how good the SD project gamification framework and its usability is. Focus group consist of three project managers, one tester and one programmer. Focus group members were familiar with Jira and were introduced with gamification principles and the developed SD project gamification framework. All focus group members have a chance to use the Jira plugin. Alter they answered six questions: Benefits of tool usage? Does the tool motivate to gain new knowledge? Does the tool improve team integrity? Does the tool motivate to do daily tasks? Improvements needed? Do you recommend the tool for your team?Focus group evaluated the plugin as easy to use and can reach its objectives. However, the same time highlighted question about correct chose of prizes that is one of the standard questions during gamification framework development [6]. Ideas for the future that were recommended are: give points based on task complexity (story points or estimate) and give points for timely spent time logging.

Conclusions

Gamification of the process is one of approach on how to change people behaviours and give more fun in boring activities. With this purpose has been developed SD project gamification framework and its prototype as the Jira plugin. Because of the SD project gamification framework using, the project team member will do they

work and accordingly update information in PM tool. Motivator for team members to do this in time will be an option to get some individual or team prizes. As a benefit for project manager is not only completed project tasks but the information in PM tool that will show actual progress, bottleneck and problem areas in project processes. Of course, it will require some additional cost from the project budget, but the same time can be reviewed as an investment in team building and project risk mitigation.

As potential future works for this gamification framework extension are: 1) make more complex point awarding dependent of task complexity; 2) award points not only for status changes but other activities – e.g. time logging, feedback giving etc.; 3) introduce more components (e.g. badges) and dynamics.

References

[1] P. Bourque and R. E. Fairley, eds., The Guide to the Software Engineering Body of Knowledge (SWEBOK Guide), 3.0. IEEE Computer Society, 2014.

[2] A. Caccamese and D. Bragantini, "The hidden pyramid," in PMI® Global Congress 2015, Orlando, Florida, USA, PA: Project Management Institute, 2015.

[3] Project Management Institute, A Guide to the Project Management Body of Knowledge (PMBOK® Guide) 6th ed., PMI, 2017.

[4] J. Swacha, "Gamification in Enterprise Information Systems: What, Why and How," Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, vol. 8, pp. 1229–1233, Oct. 2016. https://doi.org/10.15439/2016f460

[5] S. Arai, K. Sakamoto, H. Washizaki and Y. Fukazawa, "A Gamified Tool for Motivating Developers to Remove Warnings of Bug Pattern Tools,"2014 6th International Workshop on Empirical Software Engineering in Practice, Nov. 2014. https://doi.org/10.1109/iwesep.2014.17

[6] J. Landsell and E. Hägglund, "Towards a Gamification Framework: Limitations and opportunities when gamifying business processes," Tech. Rep 2016. [Online]. Available: https://umu.diva-portal.org/smash/get/diva2:929548/FULLTEXT01.pdf [Accessed: 04-Mar-2019].

[7] R. Smith, "The future of work is play", в CHI 2013 Workshop "Designing Gamifica-tion", Paris, 2013

[8] "Project Management: How to Motivate Team Members" [Online]. Available: https://www.dummies.com/careers/project-management/project-management-how-to-motivate-team-members/ [Accessed: 04-Mar-2019].

[9] V. Platonova and S. Bērziša, "Gamification in Software Development Projects," Inf. Technol. Manag. Sci., vol. 20, pp. 58–63, 2017. https://doi.org/10.1515/itms-2017-0010

[10] International Institute of Business Analysis, A Guide to the Business Analysis Body of Knowledge (BABOK), 3rd ed., IIBA 2015.

[11] ISQTB "Certified Tester Foundation Level Syllabus, 2018" [Online]. Available: https://www.istqb.org/downloads/send/51-ct-fl2018/208-ctfl-2018-syllabus.html [Accessed: 25-Mar-2019].

[12] I R. Management Association, Ed., "Gamification," 2015. https://doi.org/10.4018/978-1-4666-8200-9

[13] E. B. Passos, D. B. Medeiros, P. A. S. Neto and E. W. G. Clua, "Turning Real-World Software Development into a Game," 2011 Brazilian Symposium on Games and Digital Entertainment, Nov. 2011. https://doi.org/10.1109/sbgames.2011.32

[14] L. Singer and K. Schneider, "It was a Bit of a Race: Gamification of Version Control," 2012 Second International Workshop on Games andSoftware Engineering: Realizing User Engagement with Game Engineering Techniques (GAS), Jun. 2012. https://doi.org/10.1109/gas.2012.6225927

[15] T. Barik, E. Murphy-Hill and T. Zimmermann, "A Perspective on Blending Programming Environments and Games: Beyond Points, Badges, and Leaderboards," 2016 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), Sep. 2016. https://doi.org/10.1109/vlhcc.2016.7739676

[16] A. Marczewski, "A Simple Gamification Framework / Cheat Sheet." 2012. [Online]. Available: https://www.gamified.uk/gamification-framework/ [Accessed: 25-Mar-2019].

[17] D. Riera, J. Arnedo-Moreno, A. Mora, and C. González, "A literature review of gamification design frameworks,", 7th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games), no. April 2015, 2015. https://doi.org/10.1109/VS-GAMES.2015.7295760

[18] K. Werbach and D. Hunter, For the win: how game thinking can revolutionize your business. Wharton, 2012.

[19] "Gamification model canvas" [Online]. Available: http://www.theloyaltygames.com/wp-content/uploads/2014/07/gamification_model_canvas.pdf [Accessed: 25-Mar-2019].

A Methodology to Diagnose ICT Governance Process Based on ISO/IEC 38500 Standard. Case Study: Ecuadorian Retail Organization

Andrés Robalino-López
Grupo de Investigación de Sistemas de Información, Gestión de la Tecnología e Innovación. Departamento de Estudios Organizacionales y Desarrollo Humano,Escuela Politécnica Nacional
Quito-Ecuador
andres.robalino@epn.edu.ec

Daniel Real
Escuela Politécnica Nacional
Quito-Ecuador
daniel.real@epn.edu.ec

Zanna Aniscenko
University of Seville)
Seville-Spain
zanna.aniscenko@gmail.com

***Abstract*— The purpose of this paper is to show the application of a new methodology approach to diagnose the impact of management of Information and Communication Technologies (ICT) with the reference frame ISO/IEC 38500 inside the organizations. It is based on the government's analysis of organization's technological resource management, being the strategic axis part of business strategic planning and its ICT Governance, effectiveness and efficiency, and the value of ICT for its guidelines and the risks involved in its development to optimize the use of organization's resources. In the first stage, a documentary research was carried out as an essential part of this study and an interview with an expert; Then a tool (questionnaire) was designed to collect, measure and evaluate data on aspects and components of the ISO/IEC 38500 standard. A case study of Ecuadorian retail organization was carried out. Once the information was processed, the respective indicators were obtained that helped to identify current status of ICT governance processes of the organization. With the obtained indicators we could see which principles are the ones that characterize this sector, in which they put the greatest amount of resources; be they human or financial. It was also evident that the two dimensions on which the study was based on have very different results, since emphasis was placed on what corresponds.**

***Keywords*—ISO/IEC 38500, ICT Government, Methodology Approach, Retail Sector.**

I. INTRODUCTION

The use of Information and Communication Technologies (ICT's) has become as one more phenomenon within globalization; stimulating the starting point for the optimization of managerial processes and causing changes in the organizational structure. The importance and influence of ICT's in an organization are directly linked to the characteristics of it, so ICT's will differ from a size of an organization whether it is a small and medium enterprise (SME) or a large organization.

Currently, more companies make important investments in information and communication technologies in order to meet specific needs, such as growth of the organization, seeking for greater control in their transactions, information storage and processing, etc. However, many of them do not see immediate benefits from such high investments, mainly due to the inefficient use of ICT's due to inadequate implementation or use of it.

Thus, the introduction of ICT's in companies has advantages, however it also imposes risks; this last scenario occurs because of the poor capacity and skills in ICT management by people in charge of maintaining them within the organization. Therefore, these companies take too long to recover their investment and as they are not fully exploited, they tend to stop using them and consequently it becomes a cost for the organization that cannot be recovered.

With a correct implementation and referring to governance of the ICT's, these technologies can be used to the maximum and obtain a competitive advantage that would improve the organization's economic performance and recover the investment made more quickly.

Currently technology has become as a very important element within organizations, whether it is a public or private, so they invest a considerable amount in the acquisition, maintenance and development of the same. However, there is a difference with respect to the objectives that the organization presents, since the appropriate way of aligning the strategic objectives is sought and the so-called ICT government comes into play.

The technological infrastructure and internet access platforms are the foundations on which ICT's are based, since they allow access to services such as e-business and e-commerce, among others, they are also the ones that drive changes, reducing the gap that exists around the "Information Society" [1]. Therefore, ICT requires a significant investment, which affects 10.5% of revenues for some companies [2].

A series of drawbacks have been recognized, such

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4176
© 2019 Andrés Robalino-López, Daniel Real, Zanna Aniscenko.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

as lack of vision regarding their current performance, infrastructure and personnel failures, data management, among others, as well as problems when internalizing the processes proposed in the model, such as achieving the alignment of strategies, risk management and development of value for the organization through its technology tools [3].

A. ICT government

Currently, the ICT Government, especially in companies, is spreading rapidly; half of the surveyed organizations acknowledge having implemented or are in the process of implementing elements of the ICT Government. In addition, the ICT Government is demonstrating its effectiveness and performance when it comes to obtaining the maximum value of ICT's for organizations, 65% of those responsible for ICT who have already implemented these systems recognize that they are effective [4].

To define more clearly, what corporate governance refers to, we will start with the explanation of the meaning of governance It is related to the way of governing and whose main objective is to obtain achievements that can be in different areas, such as economic, social and organizational.

It can be said that corporate governance is mainly related to management and control; however, the difference can be seen in the proposed structure, in which the distribution of the rights and responsibilities of each of the participants in the organization (general directory, managers, shareholders and others with an economic interest) stands out in the organization.

On the other hand, corporate governance can be defined as the set of principles and rules that regulate the design, integration and operation of the governing bodies of the organization, such as the three powers within an organization: Shareholders, Board of Directors and Senior Management [5].

Ross and Weill [6] mention that it is "an actively designed set of ICT governance mechanisms (for example, committees, budget processes, approvals, ICT organizational structure, charge back, etc.) that foster a behavior consistent with the mission, strategy, values, norms and culture of the organization". In addition, all this in order to get revenues for the organization.

The fact of having the implementation of ICT's; it does not represent itself a competitive advantage for organizations. It is the management of these technologies that can give an advantage to their success. According to what has been said, appropriating an ICT governance model for this management is a key element for the fulfillment of the organization's objectives [7].

In this part, it cannot be forgotten that that the development of clear policies and good practices for the security and control of ICT's is always important, in order to obtain the approval and support of the different levels of government of the organization [8]. In addition, this applies to the information systems of the entire organization, including personal computers and networks. It is based on the philosophy that ICT resources need to be managed by a set of naturally grouped processes to provide the pertinent and reliable information required by an organization to achieve its objectives. [9]

B. ICT Governance

Henriquez [2] mentions that the process of governance in this field consists of a structure of relationships and processes aimed to direct and control the organization, in order to achieve its objectives and add value while balancing the risks and returning investments on ICT's and its processes, (see Fig. 1).

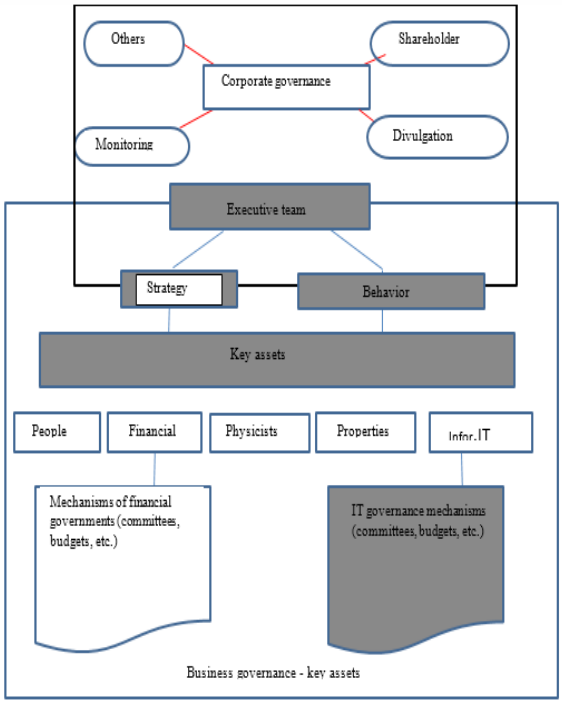


Fig. 1. ICT Governance. Adaptation of the model by Peter Weill and Jeanne W. Ross

Therefore, the governance of ICT's is a part of corporate governance, because of its structure in which decisions are made throughout its processes. It seeks to direct and control the operative part that oversees reaching the established goals, meeting the relevant objectives, adding value to the processes, and especially monitoring the returns that are obtained by ICT's.

ICT governance is the responsibility of executives and the board of directors, and consists of leadership, organizational structures and processes that ensure that the organization maintains and extends the organization's strategies and objectives [10]; The government of ICT is also defined as the structure by which the project's objectives are established, as well as it means to achieve them and to monitor the performance for them.

On the other hand, we must not forget that the ICT government puts into practice ensuring that the processes comply with the agreed standards, which will allow the organization to achieve the objectives that were proposed, obtain the information quickly and concisely, and to be able to take advantage of opportunities to gain an advantage over other competitors in the market.

It should also be emphasized that the ICT governance manages four fundamental principles that must be applied for a better performance of the government; these are: i) Direct and control, ii) Responsibility, iii) Accountability

and iv) Activities [11].

From this premise we can complement another concept of what ICT governance is: who will be responsible for decision making and who will be responsible for the accounts of how ICT has been handled, who will be the one to say how ICT management should be leaded, how it has been benefiting and developing the organization, and how this has improved the individual and collective performance of internal collaborators. On the other hand, ICT government must be carried out the following areas: i) Strategic alignment, ii) Delivery of values iii) Risk management and iv) Performance measurement.

For a good performance of any tool, in this case the ICT which has an organization, can be analyzed from the point of view of an optimal investment, the allocation of adequate resources to meet the objectives set by the area of ICT, and the purpose of reducing costs that are related to this area.

C. ICT Governance framework ISO/IEC 38500

This framework can be found in the extensive literature of ICT's, each one has a particularity and its approaches are different, but at the same time, they are complementary due to the approaches. The Government Framework that uses ICT and which we are going to focus on is the ISO/IEC 38500 Corporate Governance of Information Technologies. This norm was based on its Australian pair AS8015: 2005, which is the first norm or standard on the governments of the ICT's (ISO/IEC, 2008); Its objective is to provide a framework of principles on which organizations can base themselves to use when assessing, directing and monitoring the use of information technologies. This standard indicates six principles that form a base to create good corporate ICT governance: i) Responsibility (r): of individuals and groups; ii) Strategy (s): ICT satisfies the strategy of the organization; iii) Acquisition (a): of ICT for valid reasons; iv) Performance (p): based on supporting the business of the organization v) Fulfilment (f): with mandatory legislation and regulations and vi) Human behavior (h): response to the needs of all people in the process.

In this model, 3 areas were considered (evaluate, direct and control) as the main ones so that the roles that allow managing, planning, implementing and using ICT's can be addressed and controlled (see Fig. 2).

The model can be easily understood since the ICT government focus directly leads to the most basic business model: Plan - Build - Operate. It should be recognized that while this model is sometimes used by ICT specialists to explain aspects of the ICT cycle, it is also widely understood by business leaders and educators as the basic cycle of business management. That is the context in which it is used here [12].

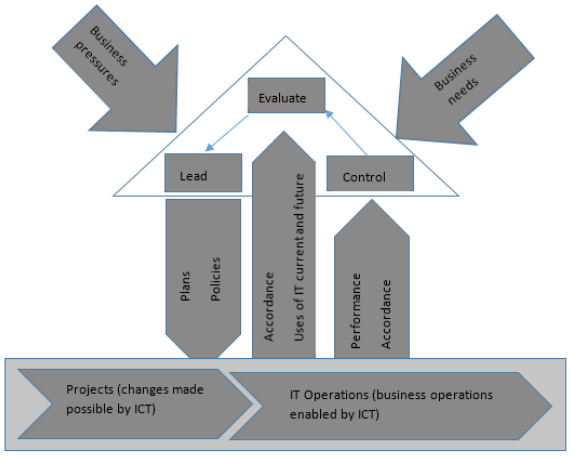


Fig. 2. Governance model IT ISO / IEC 38500

In addition, we should have clear overview at the time of choosing these technologies that will be those that provide us with some confidence when looking for effectiveness, efficiency, reliability, compliance, availability, integrity and confidentiality. Without forgetting that it is required to identify the types of resources or assets that the organization possesses, and the processes that are the necessary controls to be implemented for complying business requirements. The processes that are going to be implemented must be under supervision of the directives and executive management.

II. MATERIALS AND METHODS

A. Nature of the investigation

This research has a mixed approach (qualitative and quantitative). First, a state of art review in governance and technological management together with interviews with local experts in the sector were performed in order to get a contextualization for Ecuadorian organizations. Second, following Bosch [13] and Robalino-López and Aniscenko [14], 18 items for the six dimensions that ISO/IEC 38500 handles were used for the construction and validation of the tool. The scope of this research has a descriptive and exploratory character. In Ecuador, as far as we know, there are no studies about the ICT governance processes for commercial sector (Retail), being a completely new area for the region.

To fulfill the objectives of this study, we used a mixed design, and Non-Experimental – case of study, since no variable is going to be manipulated and only the facts and situations that interest this study will be observed. Non-Experimental research was carried out in a cross-sectional manner since the facts that are currently happening in the organization, in the area of ICT's were observed. In addition, the current situation of the Ecuadorian retail sector was described and a tool ICT-staff, distributed as follows: i) 3 Operators or users, ii) for measuring governance processes was developed.

The Organization of the "case of study" has 9 people as 5 Middle controllers or key users and iv) 1 Director or leader.

ICT-staff of organization is described as: a professional with academic level of technical education, of third

and fourth educational levels (bachelors and masters), belonging to the "Organization" in operative positions, middle managers and managers, with seniority in the position of at least one year in the organization.

B. Research tools

The research tools used in the present study were: i) Documentary research, ii) Interview and iii) Survey. Additionally, these tools were based on the activities suggested by the standard that is being handled, which in this case is ISO/IEC 38500, that is, both the survey and the interview that are based on the premises of this international standard.

The documentary research is defined as "the essential part of a scientific research process, contributing to a strategy where we observe and reflect systematically on realities using different types of documents for them [15]. The documentary research was made based on papers that manage the concepts of governance and how they focus on ICT, especially in the differentiation of corporate governance and business governance. It is also based on the principles of the ISO/IEC 38500 standard, which is what, is supported to give the guidelines for the governance of ICTs, and how these should be managed so that all the potential is used. The documentary research was carried out in order to lean about the state of art of the ICT Governance processes, focusing the study on the ISO 38500 standard, which was previously explained. The base texts for the development of the present study, in addition to all the complementary sources were [2], [16] – [19].

The interview is a fundamental tool that helps us to obtain information and allows to improve the knowledge of the processes by which the ICT governance can be established in the Ecuadorian retail sector, which works with marketing of mattresses. It also allowed us to have a better vision of the validation and contextualization of the survey to be developed [20] – [21].

The survey is an extremely important tool for research and in the case of this work it was essential because there was developed the questionnaire through which we focused on the retail sector of Ecuador. The survey was focused on the commercial sector of mattresses. The characteristics of the methodology to carry out the survey were: i) information sources collection: The primary source of information could be obtained in the area of Information Technology of the "Organization of the case of study", in which questions were organized in a survey, and it was done to the users, the supervisors and the leaders of ICT area of the organization. Totaling nine people corresponding to the organization-staff in the area of ICT's; ii) Composition of the survey: The instrument to validate was a questionnaire composed of contextual questions formulated in two stages that were designed to measure the previous state in which the implementation of the ICT's was carried out, taking into account the perception of the collaborators of the organization, and the importance that it had when the implementation or a subsequent analysis of it was carried out. For both questionnaires that were composed by 18 questions, the Likert scale from 0 to 3 points was taken as a measurement reference. The design and development of the tool forms an important part of the process in which the questionnaire was developed as an intermediate step; However, this only gave rise to the second stage from which the evaluation and adaptation of contents that should be in tune with the research of the work emerged; once this stage was finished, the validation of the tool was performed.

C. Assessment method

Following to Bin-Abbas and Bakry [22] and Robalino-Lopez and Aniscenko [14], the proposed assessment method has two questions on each control element: the level of importance of the element (w[i]); and the level of its implementation (g[i]). Four levels were considered for both questions as illustrated in Table 1. Note that the two mentioned questions concerned with each issue can be assessed by ICT-staff of the organization (Operators or users, Middle controls or key users and Director or leader). Then it is important to find averages for not only the importance but also the implementation level, as shown in Table 1. In addition, a relative measure that combines the averages of both: importance and implementation can be found. This provides this relative combined measure as a percentage value (see Table 1).

TABLE 1. ASSESSMENT METHOD: IMPORTANCE (W[I]) AND IMPLEMENTATION LEVEL (G[I]).

Levels of Scale (L=4) for both importance (w[i]) and implementation level (g[i]).			
Poor/low	Below avg.	Above avg.	Good/high
0	1	2	3

Average importance ($w_{[i]}$) and implementation level ($g_{[i]}$) of a control element *i* for *k* assessments.

$$w_{[i]} = \sum_{j=1}^{k} \frac{w_{[i,j]}}{k}; \quad g_{[i]} = \sum_{j=1}^{k} \frac{g_{[i,j]}}{k} \qquad (1) \text{ and } (2)$$

General performance indicator *A* for each principle (*r, s, a, p, f, h*) for *N* elements of each principle.

$$A_{r,s,a,p,c,h\,[i]} = \frac{\sum_{i=1}^{N} w_{[i]} * g_{[i]}}{\sum_{i=1}^{N} w_{[i]} * L}\,\% \qquad (3)$$

General performance indicator A_t: for N_t elements of the model.

$$At_{[i]} = \frac{\sum_{i=1}^{N_t} w_{[i]} * g_{[i]}}{\sum_{i=1}^{N_t} w_{[i]} * L}\,\% \qquad (4)$$

The mentioned above can be applied both to the basic ICT governance control elements mentioned in the standard and to other possible elements that may be needed for specific cases (organizational field characteristics). As will be shown below, the ICT governance's requirement controls are open to further additional considerations that may be considered. This enhances knowledge sharing and support improvement.

Table 2. Evaluation Model (Bosch)

	EVALUATE	LEAD	CONTROL
RESPONSIBILITY	**(RE1)** The models and options for assigning responsibilities. **(RE2)** The competences of those who receive responsibility.	**(RL1)** That the designed plans are carry out. **(RL2)** Those managers receive the information they need to make decisions.	**(RC1)** The assignment of responsibilities. **(RC2)** The adequate performance of assigned responsibilities (indicators).
STRATEGY	**(SE3)** Development of ICT to verify that they will support the business in the future. **(SE4)** Alignment of ICT activities with business objectives. **(SE5)** Management of risks related to the use of ICT.	**(SL3)** Design policies and plans that take advantage of the value of ICT. **(SL4)** ICT Innovation.	**(SC3)** The objectives are meet in the term with the planned resources. **(SC4)** The results to verify that the expected benefits have been achieve.
ACQUISITION	**(AE6)** Different options with ICT offers in relation to cost and risk. Risk / value analysis.	**(AL5)** The asset purchase procedure is carrying out appropriately. **(AL6)** Meeting the needs of the organization.	**(AC5)** Investments provide the expected capabilities. **(AC6)** The internal / external understanding needs the organization.
PERFORMANCE	**(PE7)** Operational proposals from ICT managers to maintain business capacity. **(PE8)** The risk of ICT in relation to the continuity of business operations. **(PE9)** The risk of the integrity of information and the protection of assets. **(NE10)** The effectiveness of ICT decisions in support of the organization	**(PL7)** Those enough ICT resources are available**.** **(PL8)** That the correct and updated information is providing to the management as support for the decisions. **(PL9)** Assign priorities and restrictions.	**(PC7)** To what extent ICT support the business. **(PC8)** The prioritization of resource allocation in relation to business objectives. **(PC9)** Compliance with established policies and rules **(PC10)** Data accuracy policies and efficient use of ICT.
ACCOPLISHMENT	**(AE11)** To what extent are the guidelines, legislation and internal rules establish. **(AE12)** Compliance with internal procedures established in the organization.	**(AL10)** Those mechanisms be establishing to verify compliance with laws, standards and standards. **(AL11)** Those policies be establishing that support the use and integration of ICT. **(AL12)** That the ICT personnel have a professional behavior and respect the procedures. **(AL13)** That an ethical use of ICT be make.	**(AC11)** Compliance and compliance (audits / reports). Timely, complete and adequate. **(AC12)** ICTs preserve privacy and strategic knowledge. **(AC13)** ICT processes.
HUMAN FACTOR	**(FE13)** That the human component is identify and take into account in all ICT activities.	**(FL14)** Those ICT activities are consistent with the human component. **(FL15)** The risks and opportunities can be identified and reported (policies and procedures) to the directors for their evaluation	**(FC14)** The perception of the importance of the human component (training). **(FC15)** The application of appropriate practices to be consistent with the use of ICT

III. RESULTS AND DISCUSSION

The growing diffusion of Information and Communication Technologies (ICT's) has contributed to the strengthening of governance. The studies presented analyze partial aspects of this integral issue, including the effects of ICT's on stability and development, advances in the field of transparency and open government, the way in which companies use ICT's to communicate with their clients. customers and the use and perception of ICT's by civil society and competition.

A. Evaluation model (BOSCH)

The management of ICT's is a part of their governance and constitutes the component for the achievement of excellence and competitiveness of organizations.

The assessment of compliance with the six principles of the ISO/IEC 38500 standard will be based on the recommendations of good ICT governance made by Bosch [13]: i) It is addressed to top management, as it indicates the way in which they should evaluate, direct and monitor the use of ICT's throughout the organization. ii) But it is also an aim at ICT managers, as it informs and guides them on how to design and implement management policies, processes and structures that support ICT governance. Managers should govern ICTs through the actions of Evaluating, Directing and Monitoring, following the model proposed by ISO 38500 [13]. Monitor, through an adequate measurement system, the appropriate policies, extremely clear procedures and established plans "Table 2".

B. Results of the evaluation

The evaluation carried out, in which the surveys were used, was developed according to the frame of reference of the ISO/IEC 38500 standard. In addition, the evaluation of the information obtained from secondary and primary sources was made, from which we took the relevant information and transformed it into annotations or citations that helped us to mark the development of this study.

Through survey, conducted results were obtained with which we can assess ICT infrastructure, whether the staff agrees or disagrees with the management of ICT's in the organization and on premise with the principles based on ISO / IEC 38500.

Table 3 shows the answers of all the principles related to the questions posed in the survey. In this part, we will analyze the percentage of compliance that the organization has in relation to the two dimensions (Implementation and Importance) that were raised in the query.

In Fig. 3, we can see that the compliance percentages are high, but if we refer to those that pass 80% we can see that the questions of the principle of responsibility, the three surpass this percentage (see Table 3).

In the principle of strategy, two questions surpass; in the principle of acquisition, performance and compliance only one question exceeds or equals this percentage; while the principle of human factor none of the three questions reaches the minimum established.

Similarly, when analyzing Table 4 (compliance percentage, importance dimension), we can observe that compliance percentages are high, but if we refer to those that pass or are to equal 80%, we can see that the questions of the strategy principle and human factor, all three indicators that surpass this percentage.

In the principle of responsibility, performance and compliance, two questions exceed the percentage, while in the principle of acquisition none of the three questions reaches the minimum established.

Table 3. Perception of compliance Implementation

Perception of compliance Implementation			
Question	Principle	Score	% compliance
P1	Responsibility	32	80,00
P2	Responsibility	34	85,00
P3	Responsibility	33	82,50
		Average	**82,50**

Question	Principle	Score	% compliance
P4	Strategy	34	85,00
P5	Strategy	30	75,00
P6	Strategy	35	87,50
		Average	**82,50**

Question	Principle	Score	% compliance
P7	Acquisition	33	82,50
P8	Acquisition	31	77,50
P9	Acquisition	30	75,00
		Average	**78,33**

Question	Principle	Score	% compliance
P10	Performance	29	72,50
P11	Performance	30	75,00
P12	Performance	32	80,00
		Average	**75,83**

Question	Principle	Score	% compliance
P13	Accomplishment	29	72,50
P14	Accomplishment	30	75,00
P15	Accomplishment	35	87,50
		Average	**78,33**

Question	Principle	Score	% compliance
P16	Human factor	30	75,00
P17	Human factor	30	75,00
P18	Human factor	31	77,50
		Average	**75,83**

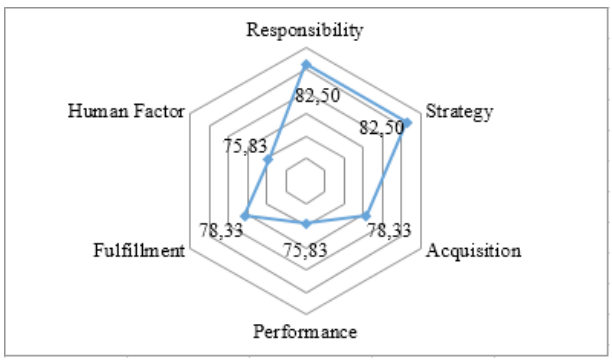


Fig. 3 Perception of compliance Implementation

Results obtained have a direct relationship with respect to the definitions made by the international standard ISO / IEC 38500. With the survey, conducted results were obtained with which we can assess whether the staff agrees or disagrees with the management of ICTs in the organization and on premise with the principles based on ISO/IEC 38500.

Table 4. Perception of compliance Importance

Perception of compliance Importance			
Question	Principle	Score	% compliance
P1	Responsibility	33	82,50
P2	Responsibility	30	75,00
P3	Responsibility	34	85,00
		Average	80,83

Question	Principle	Score	% compliance
P4	Strategy	34	85,00
P5	Strategy	32	80,00
P6	Strategy	33	82,50
		Average	**82,50**

Question	Principle	Score	% compliance
P7	Acquisition	31	77,50
P8	Acquisition	31	77,50
P9	Acquisition	30	75,00
		Average	**76,67**

Question	Principle	Score	% compliance
P10	Performance	34	85,00
P11	Performance	31	77,50
P12	Performance	33	82,50
		Average	**81,67**

Question	Principle	Score	% compliance
P13	Accomplishment	33	82,50
P14	Accomplishment	30	75,00
P15	Accomplishment	34	85,00
		Average	**80,83**

Question	Principle	Score	% compliance
P16	Human factor	32	80,00
P17	Human factor	32	80,00
P18	Human factor	33	82,50
		Average	**80,83**

Similarly, when analyzing Fig. 3 (compliance percentage, implementation dimension), we can see that the compliance percentages are high, but if we refer to those that pass 80% we can see that the questions of the principle of responsibility, the three surpass this percentage; In the principle of strategy, two questions surpass this percentage; in the principle of acquisition, performance and compliance only one question exceeds or equals this percentage; while the principle of human factor none of the three questions reaches the minimum established.

In Fig. 4 (compliance percentage, importance

dimension), we can observe that compliance percentages are high, but if we refer to those that surpass or are equal to 80%, we can see that the questions of the strategy principle and human factor, all three surpass this percentage.

In the principle of responsibility, performance and compliance, two questions exceed the percentage, while in the principle of acquisition none of the three questions reaches the minimum established.

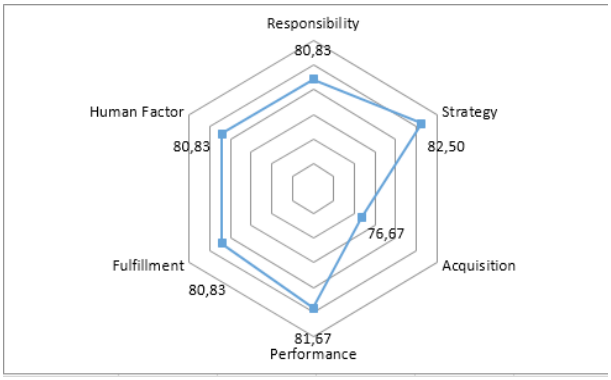


Fig. 4 Perception of compliance Importance

IV. CONCLUSIONS

The study proposes the use of new methodology approach to analyze the governance processes of ICTs within organizations. The approach is showing in a case study of the organization in the Ecuadorian retail sector.

The aim is to use the ISO/IEC 38500 model as a basis in conjunction with an adhoc model (Bosch model) to have a broader vision of the processes involving ICTs in Ecuadorian retail companies. Indicators that were generated allowed to obtain a first diagnosis of the use of ICTs within the organization and show the implementation of the proposed approach.

The design of the tool for the collection of information is based mainly on the proposed model. This model was developed to obtain a broader perspective of how the ISO/IEC 38500 standard should be evaluated and implemented within organizations. The tasks in the management process: assess, direct and control are fundamental to understand the good use of the ICT resource within the organization.

The diagnosis made in the organization of the case study shows the use of an important amount of resources, (human and economic) within the area of technologies; However, there are considerable gaps in the implementation and import perspectives in the model: This reveals that there is a gap between the real state of the processes and their ideal state. The results provide a wider and more useful view for the top management of the organization.

ACKNOWLEDGEMENTS

This work has been funded by Escuela Politécnica Nacional of Ecuador, under the Research Projects PIJ-16-10 and PIMI-16-02.

REFERENCES

[1] S. Cedeño and A. Robalino, "Rediseño de la infraestructura del proveedor de servicios de internet ONNET SA para la optimización del servicio en el Distrito Metropolitano de Quito," Tesis, Escuela Politecnica Nacional del Ecuador, Quito, Ecuador, 2008.

[2] S. Henriquez, "Implantación de Gobierno TI. NETWORK-SEC," 2011.

[3] A. I. Almanza Gomez, "La aplicación de Cobit en las organizaciones ¿vale la pena el esfuerzo?," Tesis, Universidad Militar Nueva Granada, 2012.

[4] A. Fernández Martinez and F. Llorens Largo, "Gobierno de las TI en las Universidades Españolas," Univ. Almer., 2011.

[5] R. Salvochea, "Mercados y buen gobierno corporativo. Buenos Aires: La Ley.," 2012. [Online]. Available: https://www.virtuniversidad.com/greenstone/collect/administracion/archives/HASH2cad.dir/doc.pdf. [Accessed: 10-Feb-2016]

[6] J. W. Ross and P. Weill, "IT Goverance: How Top Performers Manage IT Decisions Rights for Superior Results," IT Goverance How Top Performers Manag. IT Decis. Rights Super. Results, no. Issue: Harvard Business School Press Boston, Massachusetts, Pages: 1-10, pp. 1–13, 2004.

[7] C. E. Marulanda Ecehevrry, M. López Trujillo, and C. A. Cuesta Iglesias, "Modelos de Desarrollo para Gobierno TI. Models of development for IT government," Sci. Tech., vol. 1, no. 41, pp. 185–190, 2009.

[8] N. E. O. E. P. Fitzgerald, "Cobit," 2009. [Online]. Available: https://www.slideshare.net/nikifitz/cobit-1154456. [Accessed: 10-Feb-2016].

[9] J. Ocrospoma, "Investigación Cobit," 2014. [Online]. Available: http://cobitcbt.blogspot.com/. [Accessed: 16-Feb-2016].

[10] R. Moeller, Sarbanes-Oxley internal controls effective auditing with AS5, CobiT and ITIL, John Wiley. Hoboken: John Wiley & Sons, Inc., 2008.

[11] ITGI, "Global Status Report on the Governance of Enterprise It (GEIt)— 2011," Gov. An Int. J. Policy Adm., p. 70, 2011.

[12] M. Toomey, "Waltzing with the Elephant: (Bailando el Vals con el Elefante)," pp. 1–59, 2016.

[13] A. Bosch Pujol, "Herramientas para la implantación del gobierno de las TI: ISO 38500," p. 12, 2008.

[14] A. Robalino-López amd Z. Aniscenko, Gap analysis of ICT governance in organizations: A simple implementation approach since ISO/IEC 38500, Cladea Annual Conference Proceedings, 2015.

[15] P. Bravo, Luis; Ramirez, Tulio; Mendez, La Investigación documental y bibliográfica, Panapo. Caracas: Panapo, 1987.

[16] I. L. Muñoz Periñán and G. Ulloa Villegas, "Gobierno de TI – Estado del arte," Sist. Telemática, vol. 9, pp. 23–53, 2011.

[17] A. ; F. L. L. Fernández Martinez, "Gobierno de las TI para universidades," 2007, vol. 136, no. 1, pp. 23–42.

[18] Z. Aniscenko, A. Robalino-López, T. E. Rodríguez, & B. E. Pérez. ¨Regional cooperation in dealing with environmental protection. E-government and sustainable development in andean countries¨. The Vide. Tehnologija. Resursi - Environment, Technology, Resources, 1 13-19. 2017. https://doi.org/.17770/etr2017vol1.2578

[19] A. Robalino-López, V. Ramos, X. Unda, J.L. Román, "Gestión empresarial y análisis de ventajas competitivas. Portafolio de negocio de las telefónicas en Ecuador," Revista CienciAmérica 6 (3), 17-22. 2017.

[20] A Robalino-López, V Ramos, X Unda and A Franco, University's contribution to industries in the creation of a tool to diagnose innovation management processes. INTED 2017 Proceedings, 2351-2360, 2017.

[21] A Robalino-López, V Morales, X Unda, Z Aniscenco. "Procesos de desarrollo e innovación: aplicación de una metodológica para medir el nivel de innovación en el contexto organizacional ecuatoriano" Brazilian Journal of Development, vol. 5, pp. 4550-4567, Apr. 2019.

[22] H. Bin-Abbas and S. H. Bakry. "Assessment of IT governance in organizations: A simple integrated Approach". Computers in Human Behavior 32, p. 261–267, 2014.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 126-131

Development and Analysis of Authentication Method for Iot Devices Software on the Network Using Blockchain Technologies

Shtennikov Dmitry Gennadievich
Faculty of Software Engineering and Computer Systems
ITMO Universaty
Saint-Petersburg, Russia
dshtennikov@gmail.com

Kuzmin Kirill Sergeevich
Faculty of Software Engineering and Computer Systems
ITMO Universaty
Saint-Petersburg, Russia
wee3dood@gmail.com

Letov Nikolay Evgenievich
Faculty of Software Engineering and Computer Systems
ITMO Universaty
Saint-Petersburg, Russia
letov.nikolay@gmail.com

***Abstract* — Currently, there is an unsolved problem of IoT device authentication in networks to ensure their security. Due to low performance, work with traditional methods of protection is complicated and therefore a different approach is required. The article proposes a method for authenticating IoT devices (devices and software) by verifying their data and then entering it into the data storage. The implementation of a data storage using blockchain technology and the comparison of its efficiency with a classical database are considered. The aim of research is evelopment of the IoT device authentication method, analysis of the effectiveness and applicability of the developed method and comparison of various technological approaches to solving the problem. Obtained results is evaluation parameters of the resulting system and methods for constructing systems based on this method.**

Keywords— IoT, Software integrity, Blockchain.

I. Introduction

For different software, there is a problem of software integrity. Each manufacturer and developer solves the issue of protecting his software in his own way - encryption, hashing or simply refusing to operate on potentially dangerous devices - this task currently remains unsolved and the goal of the work done is to try to propose a solution that allows, in theory, to provide the necessary security level for mobile devices. [3]

Building a general architecture for the IoT a very complex task, mainly because of the extremely large variety of devices, link layer technologies, and services that may be involved in such a system. [9] An important factor in such systems is the timely identification of problems with devices. For this reason, there is a time response requirement - the shorter the response time, the sooner you can determine the cause of the fault.

II. MATERIALS AND METHODS

A prototype of a distributed system has been designed for software verification when updating and monitoring integrity for mobile devices to solve this problem.

The main purpose of such a system is to provide storage of authentication data and issue them upon request of the client, as well as to ensure that it is not possible to replace recorded records. The technology of private blockchain is quite suitable for this task - it allows you to store information about system states and software, and provides the ability to log changes to the system, so you can accurately determine where the failure occurred. [4]

With different goals of private blockchain as low cost of transactions and the high TPS (TPS- transactions per second) rate there is a problem – low security private blockchain. Comparing with an open blockchain, where more than half of all nodes have to be compromised, more than half of the validator nodes are enough to corrupt a private blockchain, and this is a smaller number. For now, the solution to this issue is the anchoring mechanism - the hash of the entire system is periodically recorded in a public blockchain. If the data of the block chains are changed, the system hash does not coincide with the one stored in the public network, which will reveal a substitution. [5]

The anchoring mechanism allows you to enhance the security of a private blockchain, but the implementation requires a public blockchain, which means that the solution will not be autonomous and energy consuming, since you still need to bring proof of work to the public blockchain.

As one of the options, it is proposed to consider a system of private blockchains with cross-anchorage. This architecture implies the presence of several private blockchains, which will periodically save their state in the chains of neighbors.

For this reason, the more blockchains are in the system, the more protected it becomes.

It is worth noting that in the case of anchoring the state of each private blockchain to all other blockchains of the

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4066
© 2019 Shtennikov Dmitry Gennadievich, Kuzmin Kirill Sergeevich, Letov Nikolay Evgenievich.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

system, the number of connections between nodes would be equal to where K is the number of connections; N is the number of nodes. Thus, the increase in the number of links in the system will grow almost quadratically with respect to the number of nodes. Looking at this phenomenon from the point of view of practical implementation, this means that network traffic will also grow steadily until the moment comes when the network's physical bandwidth is not enough for quickly use anchoring, and therefore the system will be under threat.

To resolve this issue, it is reasonable to divide the total space of blockchains into several anchorage zones. Blockchains that are in the same zone will be trustees for each other - each of them will send a hash of its state to blockchain trustees.

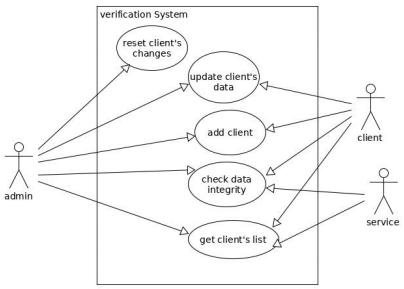


Fig 1 Use case diagram.

The overall system architecture shown on the picture:

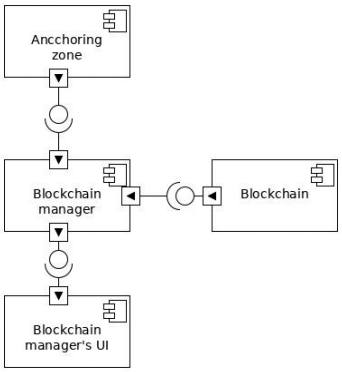


Fig. 2. System architecture.

The blockchain module implements all the functionality related to the blockchain structure itself, including: creating a blockchain, adding blocks, checking integrity, etc.

Module of zones of anchoring - this module provides opportunities for creating zones of anchoring and interaction of blockchains inside them.

The blockchain manager module is a control module for creating and managing blockchains, anchorage zones, and system configuration settings.

Scenarios for adding data about a client device and getting a list of clients shown in a Fig. 3, 4 and are not of particular interest - there are ordinary operations with a database.

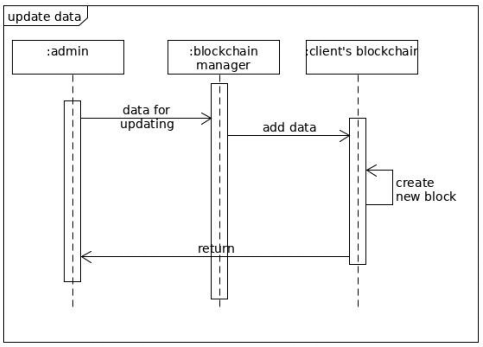


Fig. 3. Client add data scenario

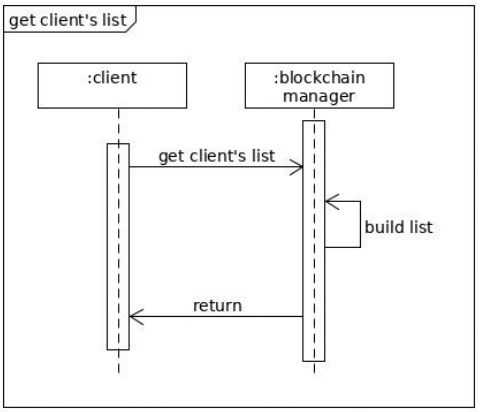


Fig. 4. Client list receiving script

The scenario of correcting client data looks the same as adding data, since the blockchain structure does not allow a simple change of data, since a change of one block will entail a violation of the connectivity of the entire system.

Of particular interest are the operations of adding a new client, creating bindings blocks and checking the integrity of customer data. In the first case, assumed change the system configuration — add or change anchorage zones, in the second and third, the anchorage mechanism is implemented and the integrity of the client's blockchain is checked, and anchor entries in the blockchain validators of the specified chain.

The sequence diagram of adding a client shown in Fig. 5.

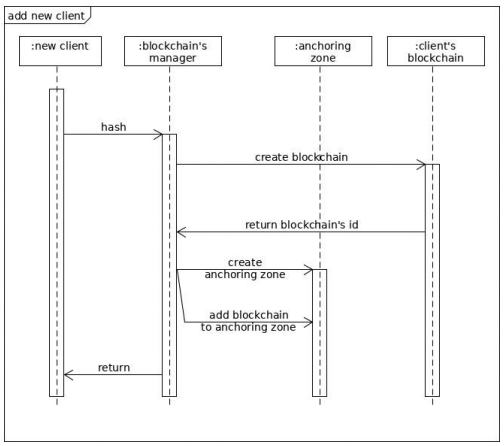


Fig. 5. Sequence diagram of adding a client

When we add a new blockchain, the blockchain manager creates a new anchoring zone and adds the created blockchain as the chain check, and all the blockchains known up to this point as the validators. This algorithm implements the principle "every with each", which is not applicable for a large number of chains, but, the simplicity of implementation and the ability to simply change the algorithm without changing the zone mechanism is quite suitable for implementation in the prototype.

The mechanism for creating anchor blocks is simple - when performing a binding in each zone, a block with a hash of the bounded blockchain added to each blockchain validator.

Scheme of process shown on a fig. 6.

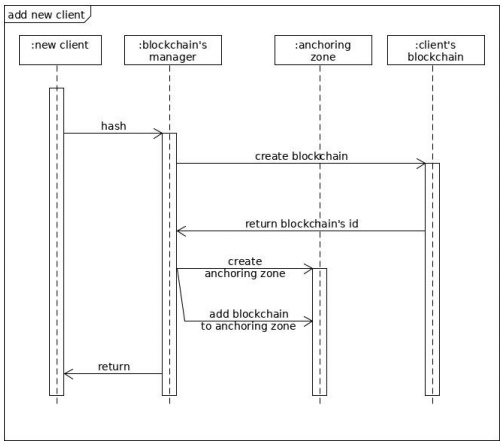


Fig. 6. Creation anchoring blocks

Integrity checking is a key mechanism for the entire system. Take a look in more detail:

1. There is a request to check the integrity of the blockchain N
2. The blockchain manager finds a zone in which blockchain N is a verifiable.
3. The last anchor block of the blockchain N is requested from each validator in this zone.
4. The current hash is requested from the blockchain being checked.
5. The resulting hash of the blockchain N comparing with the hashes of validators.
6. If more than 2/3 of the validator hashes coincided with the current hash, then the blockchain is considered verified. Otherwise, compromised.

This algorithm is simple, but effective - in order to quietly replace the data, you need to replace the data in blockchains, where N is the number of validators in the anchorage zone.

A schematic description presented in Fig. 7.

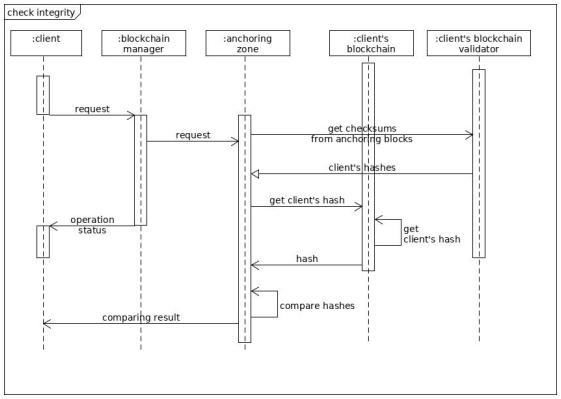


Fig. 7. Sequence diagram of checking the integrity of the client's blockchain.

Since clients are low-performance devices, they do not participate in the blockchain architecture themselves. Clients prepare data for processing and interact with the system as follows:

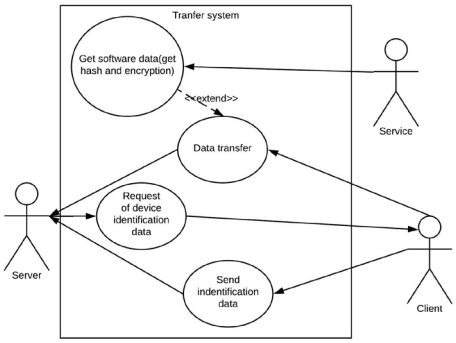


Fig. 8. Use case of client side

The client part of the system consists of 3 blocks, each of which contains its component implementation:

- Hashing block;
- Encryprion block;
- Data transfer block;

The authentication and encryption algorithms are used to protect the confidentiality and integrity of IoT data. [10]

The system works according to the following principle:

1. Made request to the client device.
2. The device returns device identification
3. The device collects a hash of predefined files / information and additional device parameters.
4. Data encryption occurs
5. Transfer data to the server.
6. Server decrypts data

The system operation sequence is presented in more detail in the Fig 9

.:

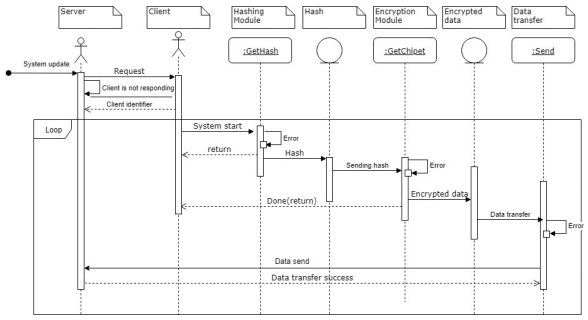


Fig 9 – sequence diagram

The server, after receiving and processing data, transfers them to the verification system.

The transactional information of IoT applications will remain secure, because all transactions are protected using cryptographic encryption. [11].

All process is shown in Fig: 10

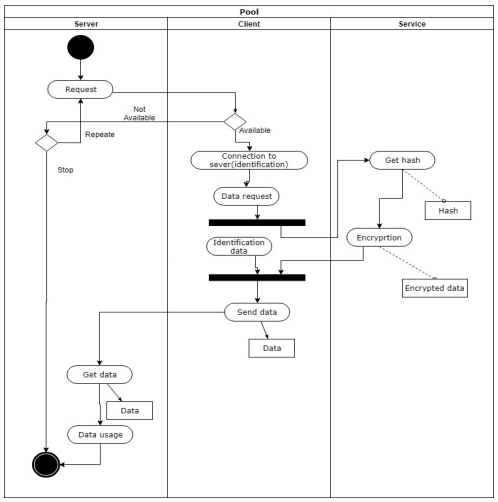


Fig –10 Sequence diagram

III. Results

As encryption algorithms, we will use the following algorithms:

- SHA256;
- SHA512;
- SHA3-256;
- SHA3-512;
- Tiger;
- Whirlpool;

MD5 is not usable due to the possibility of collisions. The following are the results of testing various algorithms for different sizes of input data (Figure 11-14):

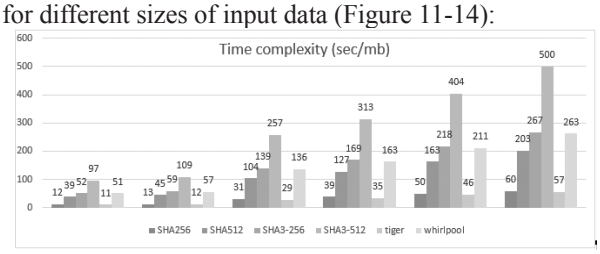


Fig 11 – Hash algorithms test

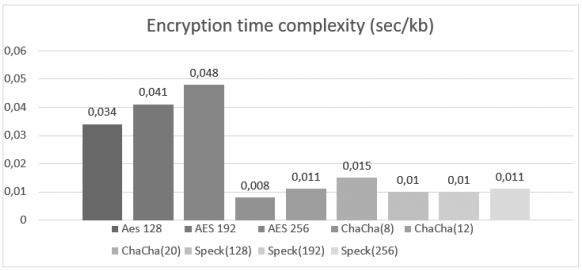


Fig 12 – Encryption test

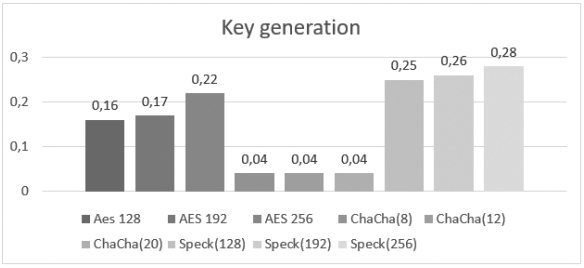


Fig 13 – Time generation test

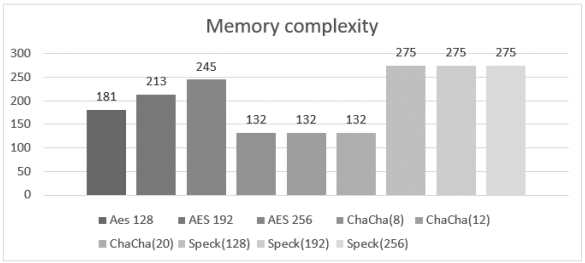


Fig 14 – Memory complexity test

Encryption algorithms are used for testing:

- AES 128/192/256;
- ChaCha 8/12/20;
- Speck 128/192/256;

The main indicators of encryption in this experiment are:

- Encryption / decryption speed
- Key generation time
- Memory usage

The client-side architecture of the resulting client interaction system is as follows:

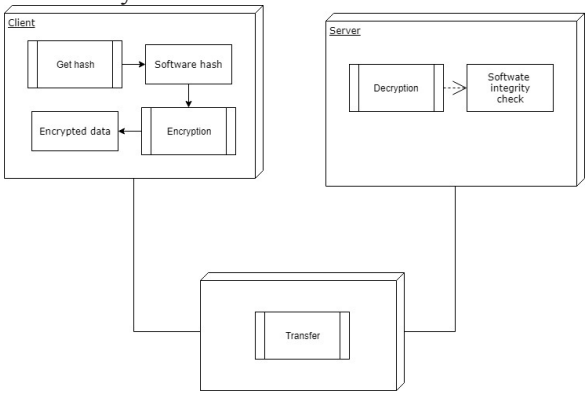


Fig 15– Client-side archetecture

Since the hardware platform itself for the implementation of the client part of the system and testing is different in capabilities, the implementation for different devices is different.

The hardware can be divided into two parts, Arduino-compatible devices [7] (such as Trekduino and ESP 8266) and with the UNIX system on board (Raspberry pi 2 and 3[8])

Accordingly, for each of these groups, the implementation requires different combinations of technologies.

In the case of Arduino compatible devices used:

- C / C ++ programming language
- Arduino cryptography Library

In the case of Raspberry 2 and 3:

- Python 3.6 programming language
- Unix-based encryption and hashing algorithms
- Socket / Django

Server implementations are used:

- Python 3.6 programming language
- Socket server

The server part of the system consists of three classes:

- Blockchain
- Blockchain Manager
- Anchorage space

Consider these classes in more detail:

The blockchain class describes, in fact, the very structure of the blockchain with support for all its functions, such as

- Adding a block;
- Integrity check;
- Getting a block;
- Getting the hash of the entire chain.

This class contains the path to the repository, and the current number of blocks, and has a unique identifier.

The class of anchorage zone (space) is an entity that describes the connections between different blockchains using the following functionality:

- Adding a blockchain to anchorage space for verification (who binds);
- Adding a blockchain to the anchorage space for validation (where bindings are made);
- Removal of the verified blockchain from the space;
- Remove a validating blockchain from the space;
- Creating bindings between blockchains;
- Check bindings between blockchains.

The anchorage zone class fields are two lists of blockchains: verifiable and validating.

The class manager provides the interaction of the external environment with the storage and manages the storage. Its features:

- Create blockchain
- Creating an anchor zone
- Storage reconfiguration
- Remove blockchain
- Remove anchorage zone

The fields are the lists of blockchains and anchorage zones.

The server class is a conductor of the system's functionality for the external environment - it implements functions for working with the main storage and interaction with the network.

Class diagram is shown (fig. 16)

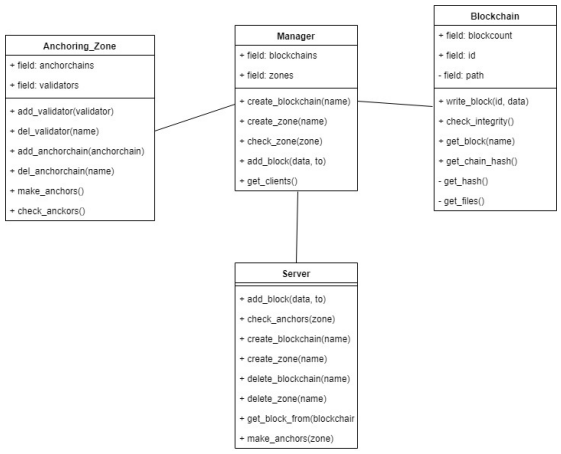


Fig 16 – class diagram.

Since it implies the ability to run the system on different devices, for the development language chosen Python3, combining power, speed of development, brevity and cross-platform.

CONCLUSIONS

The developed system allows you to ensure the security of data about the device and its software when updating and monitoring software. Each individual blocks were tested independently and the results of their testing are presented in this paper. In the future, it there is a plan to carry out testing and evaluation of the entire system as a whole with various parameters of the software and hardware.

References

[1] Donald E. Knuth. The Art of Computer Programming, vol.3. Sorting and Searching, Second Edition (Reading, Massachusetts: Addison-Wesley, 1998), xiv+780pp.+foldout. ISBN 0-201-89685-0

[2] Melanie Swan. Blockchain: Blueprint for a New Economy O'Reilly, 2015

[3] Blockchain and IoT: interaction perspectives and problems on the way of development [Online] https://forklog.com/blokchejn-i-iot-perspektivy-vzaimodejstviya-i-problemy-na-puti-razvitiya/ [Acessed: 27 Feb, 2019]

[4] Bitcoin: A Peer-to-Peer Electronic Cash System [Online] https://bitcoin.org/bitcoin.pdf [Acessed: March 2, 2019].

[5] The method of guaranteeing trust in blockchains [Online]. https://habr.com/post/338696/ [Acessed: Feb 25, 2019]

[6] Arduino vs. Raspberry Pi: Mortal enemies, or best friends? [Online] https://www.digitaltrends.com/computing/arduino-vs-raspberry-pi/ [Acessed: Feb 24, 2019]

[7] Arduino official site [Online] URL: https://www.arduino.cc/

[8] Raspberry PI official site[Online] https://www.raspberrypi.org/

[9] A. Zanella, N. Bui, A. Castellani, L. Vangelista and M. Zorzi, "Internet of Things for Smart Cities," in IEEE Internet of Things Journal, vol. 1, no. 1, pp. 22-32, Feb. 2014 http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6740844&isnumber=6810798

[10] H.C. Pöhls, JSON sensor signatures (JSS): End-to-end integrity protection from constrained device to IoT application, in: Proc. 9th Int. Conf. Innovative Mobile Internet Serv. in Ubiquitous Comput., IMIS'15, 2015, pp. 306–312

[11] G. Prisco, Slock. it to introduce smart locks linked to smart ethereum contracts, decentralize the sharing economy, Bitcoin Mag. (Nov) (2015) (online), Available: https://bitcoinmagazine.com/articles/sloc- it-to-introduce-smart-locs-lined-to-smart-ethereum-contractsdecentralize-the-sharing-economy-1446746719, (Accessed:02 February 2019)

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 132-135

Modeling of the Financial System Using the Concept of Vacuum Polarization

Dmitry Alexandrovich Shustov
Dep. of computer science, modeling and human sciences
Financial University under the Government of the Russian Federation, St. Petersburg branch
St. Petersburg
Russia
dim_shu@mail.ru

Mikhail Yuryevich Volokobinsky
Dep. of computer science, modeling and human sciences
Financial University under the Government of the Russian Federation, St. Petersburg branch
St. Petersburg
Russia
volokobin@mail.ru

Olga Anatolyevna Pekarskaya
Dep. of computer science, modeling and human sciences
Financial University under the Government of the Russian Federation, St. Petersburg branch
St. Petersburg
Russia
pekarskaya.olga@mail.ru

***Abstract*—The possibility of using the theory of deterministic chaos is considered. Vacuum polarization is a fundamental type of physical processes. The average value of any physical quantity must remain zero. Taking this assumption as a fundamental law, we propose a model that describes the processes of variation (depreciation) of the value of money. At the same time, we consider the processes of variation of the initial amount of funds introduced into the economic system by an independent "generator" (for example, the US Federal Reserve System) in a situation when the amount of returned funds must exceed the amount of given credits. Since the amount of returned funds must exceed the amount borrowed by the value of the refinancing rate, either a constantly increasing debt is formed, or an increase in the money introduced to the economy. The relationship of the "depreciation" of money is also investigated in this model**

Keywords—cash, depreciation, financial system, model, regularities, trends.

I. Introduction

The existence of various macroeconomic schools: classicism, Keynesianism, monetarism, reflect, at a minimum, the complexity of the formation of a model of an economic system.

In different models, the role and patterns of monetary circulation are determined differently. In particular, the model of circulation of money and

inflation is formed in different ways. At different historical stages, these models turn out to be untenable.

Taking into account that the economic system represented as the social practice of people under the division of labour and the existence of property rights to the results of labour, aimed to satisfy the needs and based on the equivalence (determination of the equivalent) of exchange processes, money initially acts as an equivalent, and are essentially models of these processes.

It is logical to assume that socio-economic systems, as a part of the surrounding world, are subject to general objective physical laws. In particular, the fundamental type of physical processes - vacuum polarization - can serve as the basis for modelling monetary circulation in socio-economic systems.

The starting point of the work is the assumption that socio-economic systems, as a part of the surrounding world, obey the universal objective physical laws. In particular, the fundamental type of physical processes - vacuum polarization - can serve as the basis for modeling the properties of socio-economic systems.

In accordance with this type, the average value of any physical quantity must remain zero (constant, by changing the reference point may be reduced to zero).

This approach is well known in physics in various formulations, such as, for example, the law of conservation of matter, the law of conservation of energy [2,3].

The aim of this work is to develop a model of money circulation, using the concept of vacuum polarization: money is generated artificially as the equivalent of goods and services produced by the economy, and cyclically "returns" to the generator.

II. Materials and Methods

When modelling the economic system, the subject of research in this study is the patterns of change in the monetary equivalent of labour results and needs that are satisfied in the course of practical activity of people in market economy.

Research Methods - Mathematical Modelling.

Since at present time money circulation is determined by an independent participant (regulator) of economic systems (for example, the US Federal Reserve System), which credits all economic activity, funding money according to the conditions of returning of greater amount of money than taken on credit, by the value of the lending rate .

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4045
© 2019 Dmitry Alexandrovich Shustov, Mikhail Yuryevich Volokobinsky, Olga Anatolyevna Pekarskaya.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

In such a situation, to simulate the process of cash flow in the economy, taking into account the fundamental property of vacuum polarization, the following prerequisites (axioms) must be fulfilled:

1. The number of participants satisfying their needs is constant.
2. The amount of money given on credit is allocated at the beginning of the period and does not change over the period of one cycle (usually a year)
3. Money, as an exchange equivalent, is returned to an independent regulator.
4. The average price of economic transactions is defined as the quotient of the division of cash volume in the economy, divided by the number of participants in the economic system.
5. Since the economy does not have additional physically existing funds needed to repay the cost of the loan, they are taken as a new loan in the next step.

III. Results

It is quite obvious that the model describes the properties of a dynamic system that may be described in phase space [4], [5] with the help of a system of differential equations:

$$\frac{dX}{dt} = \boldsymbol{F}(\boldsymbol{x}) \tag{1}$$

In our case it is a discrete system:

$$\frac{x}{t} = F(x) \tag{2}$$

where Δx is the change in x over a period of time Δt.

At a formalized level, the model is as follows:

Oi - the amount of money contributed to the economy in the form of a loan (for example, the Fed printed $ 1,000,000 for i = 1 first period)

Sti - lending rate (for example, 10% for i = 1 period)

WPi - received proceeds from the sale for i -th period (must be equal to the amount of money received from the Regulator)

Koli - the number of participants on the i-th period.

Cenai is the price of purchased goods and services for the i-th period reduced to one participant. Price is determined by the ratio:

$$Cena_i = \frac{O_i}{Kol_i} \tag{3}$$

Dolgi - creditors debt not secured by cash available in the economy on the i-th period

Infli - the absolute value of the "depreciation" of money (relative to the first period) on the i-th period:

$$Infl_i = \frac{\left(\frac{WP_i\ (1+St_i)}{Kol_i}\right)}{Cena_1} \quad 1 \tag{4}$$

$Year_{ihf}$ is the depreciation value relative to the previous period:

$$Year_{ihf} = Infl_i \quad Infl_{i-1} \tag{5}$$

For i> 1

Then the solution of equation (2) can be obtained in tabular form, using the EXCEL processor.

A fragment of the calculation is shown in Fig. 1.

Period	O	Em	St	WP	Cena	Kol	Dolg	Infl	Year_ihf	
0	1000000	=Dolg	0,1	=O	=WP/Kol	100000	=O*(1+St)-WP	=(WP*(1+St)/Kol)/$F$2-1	0	temp
1	=B2+C2	=Dolg	0,1	=O	=WP/Kol	=G2*(1+temp)	=O*(1+St)-WP	=(WP*(1+St)/Kol)/$F$2-1	=I3-I2	0,1
2	=B3+C3	=Dolg	0,1	=O	=WP/Kol	=G3*(1+temp)	=O*(1+St)-WP	=(WP*(1+St)/Kol)/$F$2-1	=I4-I3	
3	=B4+C4	=Dolg	0,1	=O	=WP/Kol	=G4*(1+temp)	=O*(1+St)-WP	=(WP*(1+St)/Kol)/$F$2-1	=I5-I4	
4	=B5+C5	=Dolg	0,1	=O	=WP/Kol	=G5*(1+temp)	=O*(1+St)-WP	=(WP*(1+St)/Kol)/$F$2-1	=I6-I5	

Fig. 1. Fragment of tabular model with calculation formulas

In the tabular model, the additional variable WPi is used to give a probability for further investigation of the influence of additional factors.

The variable "temp" allows to include an additional factor in the model - the share of the increment of system participants (market growth rate).

At the first stage the situation, when the market volume is assumed to be constant, lending is introduced "at one time" at the beginning of the period, loan repayment is carried out "at one time" - at the periods' end, is modeled.

To return the interest on the loan at the end of the period, additional funds are issued in the amount of the refinancing rate, which are given in the form of an increase of credit volumes.

Interest Rate does not change.

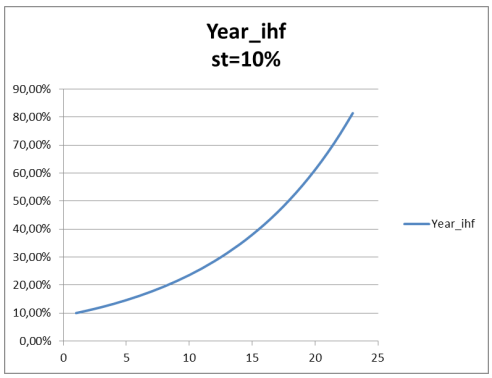
For example, with the initial amount of cash in the amount of $ 1,000,000 at 10% per annum for one year, introduced into the economic system where the number of participants is equal to 100,000, in one year it will be necessary to return to the Regulator $ 1,100,000, but only 1,000,000 is practically available. That is that during the next year it will be necessary to "reprint" (emmit) an additional $ 100,000 into the economy.

But one year later it will be necessary to return 1 121 000 $, etc.

The data fragment and the graph of "depreciation" of funds by periods is shown in Fig. 2.

Perio	O	Em	St	WP	Cena	Kol	Dolg	Infl	Year_ihf	
0	1000000	100000	10,00%	1000000	10	100000	100000	0,00%	0,00%	temp
1	1100000	110000	10,00%	1100000	11	100000	110000	10,00%	10,00%	0,00%
2	1210000	121000	10,00%	1210000	12,1	100000	121000	21,00%	11,00%	
3	1331000	133100	10,00%	1331000	13,3	100000	133100	33,10%	12,10%	
4	1464100	146410	10,00%	1464100	14,6	100000	146410	46,41%	13,31%	
5	1610510	161051	10,00%	1610510	16,1	100000	161051	61,05%	14,64%	
6	1771561	177156	10,00%	1771561	17,7	100000	177156	77,16%	16,11%	
7	1948717,1	194872	10,00%	1948717	19,5	100000	194872	94,87%	17,72%	
8	2143588,8	214359	10,00%	2143589	21,4	100000	214359	114,36%	19,49%	
9	2357947,7	235795	10,00%	2357948	23,6	100000	235795	135,79%	21,44%	
10	2593742,5	259374	10,00%	2593742	25,9	100000	259374	159,37%	23,58%	
11	2853116,7	285312	10,00%	2853117	28,5	100000	285312	185,31%	25,94%	
12	3138428,4	313843	10,00%	3138428	31,4	100000	313843	213,84%	28,53%	
13	3452271,2	345227	10,00%	3452271	34,5	100000	345227	245,23%	31,38%	
14	3797498,3	379750	10,00%	3797498	38	100000	379750	279,75%	34,52%	
15	4177248,2	417725	10,00%	4177248	41,8	100000	417725	317,72%	37,97%	
16	4594973	459497	10,00%	4594973	45,9	100000	459497	359,50%	41,77%	
17	5054470,3	505447	10,00%	5054470	50,5	100000	505447	405,45%	45,95%	
18	5559917,3	555992	10,00%	5559917	55,6	100000	555992	455,99%	50,54%	
19	6115909	611591	10,00%	6115909	61,2	100000	611591	511,59%	55,60%	
20	6727499,9	672750	10,00%	6727500	67,3	100000	672750	572,75%	61,16%	
21	7400249,9	740025	10,00%	7400250	74	100000	740025	640,02%	67,27%	
22	8140274,9	814027	10,00%	8140275	81,4	100000	814027	714,03%	74,00%	
23	8954302,4	895430	10,00%	8954302	89,5	100000	895430	795,43%	81,40%	

(a)

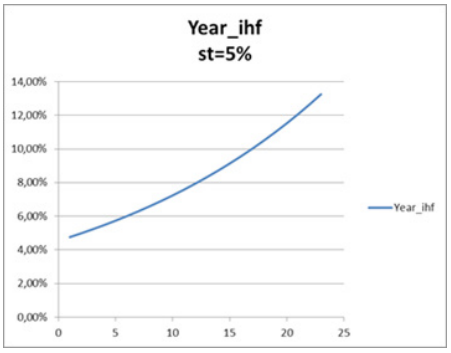


(b)

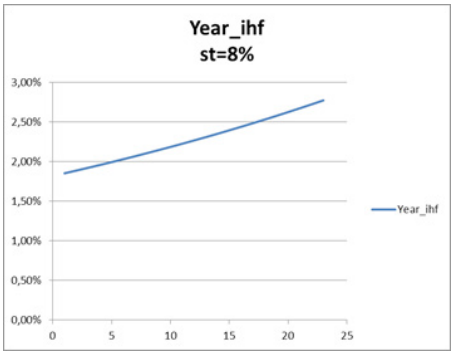
Fig. 2 Tabular model and schedule "depreciation"

Obvious is the conclusion about the increase in the rate of "depreciation" of funds, ahead of the refinancing rate.

If the number of participants in the economy increases (the model takes into account discrete - over the years - increase in the market) in an amount not exceeding the refinancing rate, the cash depreciation will decrease to zero. A number of graphs for the values of growth rates of 5% and 8% are shown in Fig.3.



(a)



(b)

Fig. 3 Schedule of "depreciation" for the market growth rate of 5% and 8%

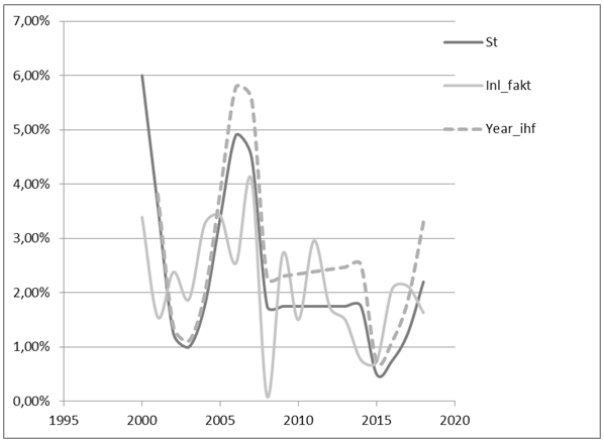
As we see, the rate of "depreciation" decreases from 80% up to 14 and 2.8, respectively.

Let's build a model of cash depreciation on the basis of statistical data on the Fed's refinancing rates and data on US inflation over the past 20 years [4], [5].

The simulation results are shown in Fig.4

Period	O	Em	St	WP	Cena	Kol	Dolg	Infl	Year_ihf	Inl_fakt
2000	1000000	60000	6,00%	1000000	10	100000	60000	0,25%	0,00%	3,39%
2001	1060000	38160	3,60%	1060000	10,6	100000	38160	6,00%	5,75%	1,55%
2002	1098160	13727	1,25%	1098160	10,9816	100000	13727	9,82%	3,82%	2,38%
2003	1111887	11118,87	1,00%	1111887	11,11887	100000	11118,87	11,19%	1,37%	1,88%
2004	1123006	19652,6	1,75%	1123006	11,23006	100000	19652,6	12,30%	1,11%	3,26%
2005	1142658	38621,86	3,38%	1142658	11,42658	100000	38621,86	14,27%	1,97%	3,42%
2006	1181280	57882,74	4,90%	1181280	11,8128	100000	57882,74	18,13%	3,86%	2,54%
2007	1239163	55762,34	4,50%	1239163	12,39163	100000	55762,34	23,92%	5,79%	4,08%
2008	1294925	22661,19	1,75%	1294925	12,94925	100000	22661,19	29,49%	5,58%	0,09%
2009	1317587	23057,77	1,75%	1317587	13,17587	100000	23057,77	31,76%	2,27%	2,72%
2010	1340644	23461,28	1,75%	1340644	13,40644	100000	23461,28	34,06%	2,31%	1,50%
2011	1364106	23871,85	1,75%	1364106	13,64106	100000	23871,85	36,41%	2,35%	2,96%
2012	1387977	24289,61	1,75%	1387977	13,87977	100000	24289,61	38,80%	2,39%	1,74%
2013	1412267	24714,67	1,75%	1412267	14,12267	100000	24714,67	41,23%	2,43%	1,50%
2014	1436982	25147,18	1,75%	1436982	14,36982	100000	25147,18	43,70%	2,47%	0,76%
2015	1462129	7310,645	0,50%	1462129	14,62129	100000	7310,645	46,21%	2,51%	0,73%
2016	1469440	11020,8	0,75%	1469440	14,6944	100000	11020,8	46,94%	0,73%	2,07%
2017	1480460	18505,75	1,25%	1480460	14,8046	100000	18505,75	48,05%	1,10%	2,12%
2018	1498966	32977,26	2,20%	1498966	14,98966	100000	32977,26	49,90%	1,85%	1,63%

(a)



(b)

Fig. 4. Simulation results with real Fed rates for the period 2000–2018.

The simulation results show the principal workability of the model, despite of the difference in the prerequisites - a real increase in total dollars (national debt) does not allow to accept as constant the volume of domestic market. However, modelling according to the US Federal Reserve for inflation and comparison with the limited market model shows comparable results in terms of trends, since the growth of the US market is due to the expense of other countries financed by the Fed, therefore the US market itself can be assumed as "limited".

IV. Confirmation

Correlation analysis of data on inflation and refinancing rates of the USA over the past 18 years allows us to estimate the existence of a linear relationship between sets of values of random variables.

Since the value of the refinancing rate and the inflation value for the simulated period may be considered as random variables, the analysis of the existence of a linear relationship between the real rate (Sti) and the value of the money depreciation Year_ihf was undertaken. The results of the analysis performed in EXCEL are shown in Fig. 5.

St	Inl_fakt						
0,06	0,0339		**Correlation**				
0,036	0,0155		r	0,53		Significance level	0,05
0,0125	0,0238		n	19		tкр	2,109816
0,01	0,0188		|tr|	2,59			
0,0175	0,0326						
0,0338	0,0342		**Hypothesis testing**				
0,049	0,0254		Variables dependent?			True	
0,045	0,0408						
0,0175	0,0009						
0,0175	0,0272						
0,0175	0,015						
0,0175	0,0296						
0,0175	0,0174						
0,0175	0,015						
0,0175	0,0076						
0,005	0,0073						
0,0075	0,0207						
0,0125	0,0212						
0,022	0,0163						

Fig. 5. Simulation results with

Comparing the values of t-statistics with a critical value allows us to state a linear relationship between the refinancing rate and the inflation rate based on official data [6], [7]

V. Conclusions

A currency circulation model based on the concept of vacuum polarization has been developed and implemented.

The study of the model confirms the existence of a link between “depreciation” and the value of the interest rate of the regulator’s lending.

The results of implementation of the model with real US FRS refinancing rates during 18 years showed a correlation between the Fed rate and the inflation rate, which allows us to make a conclusion that the model works in principle.

It is important to underline that this model can be used to study the effect on the “depreciation” of funds of other important factors, such as expansion and market contraction (change in the number of participants), withdrawal of funds from economy (decrease in the amount of funds)

It should be noted that the rate of inflation, calculated in a number of countries using different methods, may not correlate with the “depreciation” of money, caused by the principal parameter - the value of the refinancing rate.

REFERENCES

[1] E. J. Dolan et al.. Money, banking and monetary policy / Per. with English. V. Lukashevich, etc.; under the General ed.- Moscow, -448.. 1996 Available: https://economy24info.com/kredit-obraschenie-denejnoe/dengi-bankovskoe-delo-denejno-kreditnaya.html

[2] . S. P. Kuznetsov Dynamical chaos and hyperbolic attractors. From mathematics to physics — Moscow, Izhevsk: Regular and chaotic dynamics, Izhevsk Institute of computer research, 2013.- 488 c.— Available: http://www.iprbookshop.ru/28886.html

[3] M.B. Ignat’ev Kiberneticheskaja kartina mira. Slozhnye kiber-fizicheskie sistemy [Cybernetic picture of the world. Complex cyber-physical systems]. St. Petersburg, SUAI Publ; 2014. 472 p.

[4] M. B. Ignat’ev ; A. E. Karlik ; E. A. Iakovleva ; V. V. Platonov Linguo-Combinatorial Model for Diagnosing the State of Human Resources in the Digital Economy 2018 XVII Russian Scientific and Practical Conference on Planning and Teaching Engineering Staff for the Industrial and Economic Complex of the Region (PTES) https://ieeexplore.ieee.org/document/8604217

[5] D. A. Shustov. The possibilities of using the linguo-combinatorial approach in the modeling of economic processes. Information-Communication-Society (ICO–2019): Proceedings of the XVI all-Russian scientific conference / St. Petersburg, 24 – 25 January 2019.: Publishing house of Etu, 2019. - Vol. 1. P. 390-394

[6] Reference materials > inflation > US Inflation Rate Federal reserve: 2000 – 2019, http://global-finances.ru/stavka-frs-ssha, [Accessed March 27, 2019

[7] World Finance. Inflation in the US, https://fin-plus.ru/info/inflation_index/usa/, [Accessed March 27, 2019

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 136-140

Fuzzy Logic Procedure for Drawing up a Psychological Profile of Learners for Better Perception in Courses

Krasimir Slavyanov
Department Computer Systems and Technologies
"Vasil Levski" National Military University
Shumen, Bulgaria
k.o.slavyanov@gmail.com

***Abstract*—This article offers an original classification procedure based on Mamdani fuzzy inference system (FIS) dedicated to compute multiple criterions each from different type of psychological profiles. The modelling and information analysis of the FIS are developed to draw a general conclusion from several psychological criterions in order to provide better pre-course lecturer preparation and thus better students' perception. Simulation experiments are carried out in MATLAB environment.**

Keywords*—fuzzy inference system, fuzzy rule, membership function, psychological profile.

I. Introduction

It is important for short-term courses to teach the learning material in the most understandable way for learners. From this point of view, a brief survey or study could be used to quickly classify learners and, accordingly, select the most receptive style of teaching to achieve quick results with minimal effort.

People learn knowledge in different ways and the new technologies can be successfully implemented for the education improvement [1]. Howard Gardner has developed the theory of eight different types of intelligence, and it is widely used in education by then until now. Gardner believes that: All human beings possess all eight types of intelligence in varying degrees. Every person is intellectually built on different way. We can improve education by working with many types of intelligence of their students. The types of intelligence are as follows: linguistic, logical-mathematical, visual-spatial, musical, body-kinetic, interpersonal, natural and intrapersonal [2].

The basic idea in Gardner's theory about the many types of intelligence is that each person learns in their own way. Some students need to read text to understand the information, while others need to hear information in the form of a lecture. Other best learn some information through charts or pictures. This does not mean that they study only one specific way, but they are usually stronger in some respects and weaker in others. All students possess all kinds of intelligence. The predominant intelligences in the group can be quickly detected with artificial intelligence technologies like Fuzzy Inference Systems / FIS/ for implementation of a fuzzy rules ([2], [3], [4]), like for resolving similar problems ([5], [6], [7], [8], [9], [10]). In this paper, one possible way to analyse a pre-course survey results with a Fuzzy Inference System is explained. Furthermore, this approach can be successfully implemented in Human Resources Management systems ([11], [12], [13]) or easily adopted to some existing HR methods [14].

II. Materials and methods

Preconditions.

Taking this paper thematic under consideration, the well-detailed Bloom's Taxonomy of Educational Objectives can be used for the input variables in the dedicated FIS. According to that taxonomy, the educational learning objectives can be classified into levels of complexity and specificity. These levels are determined as follows: Remembering; Understanding; Applying; Analysing; Evaluating; Creating [15]. For the current purpose a cross table between Gardner's intelligences and Bloom's Taxonomy is developed and shown on Table 1.

The psychological profiles for the FIS procedure are taken to be the types of intelligence. The Classification of Educational Goals in each level are filled with the corresponding action for each type of intelligence. This action can be valued between zero and one. The resulting profile for a student is proposed to be one of a seven - for all of the Gardner's intelligences.

Building a system for making the intelligence type decision based on fuzzy logic.

The fuzzy logic tools that can be used (in Matlab environment) are of two approaches for implementing a decision-making system (Sugeno and Mamdani types of Fuzzy Inference). The membership functions in FIS tools are of many types regarding the way in with each entrance variable is forming the affiliation degree for the output.

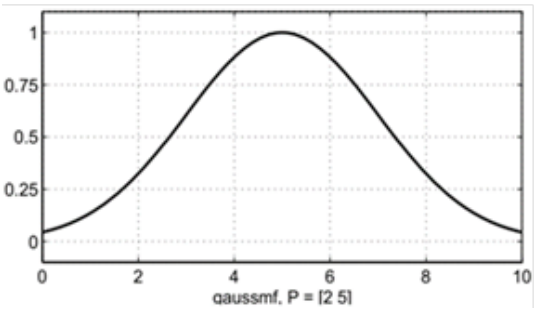
Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4073
© 2019 Krasimir Slavyanov.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

TABLE 1

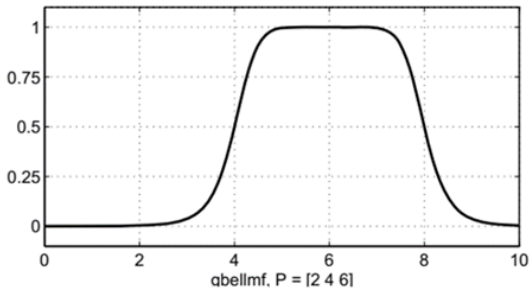
GARDNER'S INTELLIGENCES /VERTICALLY/ AND BLOOM'S TAXONOMY /HORIZONTALLY/

	REMEMBERING	UNDERSTANDING	APPLYING	ANALYSING	EVALUATING	CREATING
Verbal Linguistic	Dialogue or monolog performance	Fill in a characterization sheet	Write a letter, notes	Identify the key features of the task	Write a review	Write methodology
Logical-mathematical	Construct a rules	Draw up a grid of attributes	Interpret directions and draw up plans	Convert the plans, measure and divide up the desk board	Supervise the accuracy with mathematics	Design the set to be built precisely
Visual Spatial	Find pictures	Make a picture, sketch	Design figures, graphical models	Interpret the directions, light	Draw result charts	Create graphics with visual impact
Bodily Kinesthetic	Follow instructions	Develop appropriate movements and gestures	Gesture performance, moves,	Compare the moves	Evaluate changes in time and movements	Divide work into steps
Musical	Look for music or sound	Select appropriate sound effects	Operate the sound for the performance	Discuss in sounds, intonation	Adding sounds in results expression	Compose like for orchestra
Interpersonal	Keep a logbook of all others activities	Group, discuss, forum	Interact with other, team work	Analyze interactions and development	Assess connections in results	Design each module in group manner
Intrapersonal	Write a recount of everything	White a self-learning path	Take the lesson on an emotional journey	Analyze each object's behavior	Find Independencies	Design in one object behavior
Naturalist	Write a list of locations	Use natural world examples	Perform a scene of the task outdoors	Analyze the nature	Evaluate with nature similarity	Design recycling systems and backups

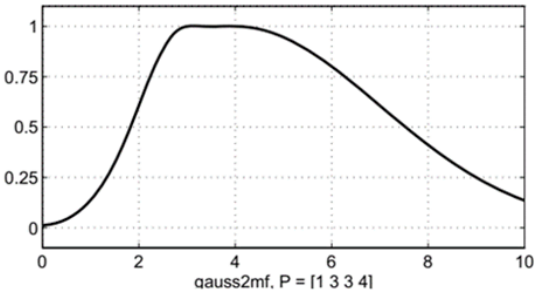
The experiments in Matlab environment are carried out with the fuzzy logic tool and the use of Gaussian type of membership functions, because of the factors: the chosen type of classification; the possible similarity in many student results; universality of application of Gaussian functions; evenness of the shape; values other than zero for all points. The subdivision of the Gaussian curves by their form includes a simple Gaussian curve and a two-way combination in type "bell" (Fig. 1).



(a)



(b)



(c)

Fig.1 Graphics of the functions belonging to the Gaussian distribution – plain (a), type "bell" and combination (c).

The degree of membership of a learner to the Gardner's intelligences is from zero to one, by the value of membership. In this way, the fuzzy inputs can be associated with a positioning rule in the membership function area. To form the resulting Gardner's intelligence for each learner, the decision making method for the FIS is chosen to be of Mamdani's type [16], because the output variable is designed to be number and the output membership functions are fuzzy sets like the input ones. The maximum value in every membership function is used for the aggregation model for the FIS input variables.

Following the Bloom's Taxonomy for the learner's complex intelligence profile, each of the levels' profile is selected to make the final defuzzification decision of the type "largest of maximum". These parameters of the FIS are designed to summarize the results of the six Bloom's Taxonomy levels for every learner examined. The system build in Matlab is shown on Fig.2.

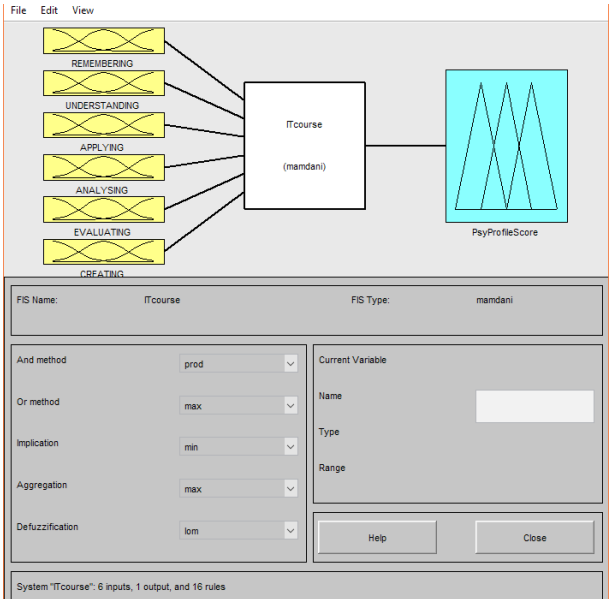


Fig. 2. Block diagram of the FIS for learner profile decision making by means of the Mamdani method, built in Matlab.

The input variables on the left side of the system are six, corresponding to each of the Bloom's Taxonomy levels used to classify the learner in the resulting Gardner's intelligence psychological profile. Each of the input variables is made of membership functions developed in order to correspond for the psychological profile. The output variable is a combination of all the rules applied to the input variables and is the person's intelligence type that best satisfied the rules.

The applied shape of the curve membership function to each input fuzzy set is of Gaussian combination type. The results for each of the Bloom's taxonomy levels (Remembering; Understanding; Applying; Analysing; Evaluating; Creating) should fall within eight fuzzy sets defined by the functions of belonging in the range of 0 to 8. Each sector is formed in manner 0-1, 1-2, and 2-3 until 7-8 in the same sequence as the output values of the Gardner's intelligence indicators (table 1.). For example the results for the Bloom's Understanding level are designed to fall in the membership functions: Fill tables; Draw grid; Make pictures; Develop movements; Sounds; Forum; Self path; Nature examples (fig. 3)

The tailor made form of these functions is to use maximum space in the range 0-8 to fall into the fuzzy sets for higher percentage results from the input value. Therefore, the increase of the impact of the indicator is designed to increase the degree of belonging to the elements of the fuzzy set. Similar are the parameters of input variables for the results for the other Bloom levels. In this paper, they are not described in detail but all are followed by the crossings in table 1. The fuzzy variable for output is developed of fuzzy sets, again in the range 0-8, given the intelligence type.

The membership functions are with the Gaussian distribution of type "bell", designed to summarize at maximum the results for each student's score (variables), classified by Gardner rules. The rules developed for the operation of FIS are 16, divided into 2 groups, the weight of every rule of the groups is 1.(fig. 5).

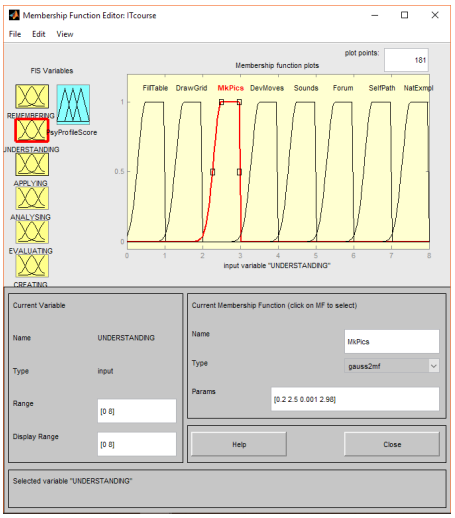


Fig. 3. Diagram describing the membership functions of fuzzy set corresponding to the variation in criteria for level Understanding

The functions for the fuzzy sets of the output variable – Profile Score, are shown on figure 4.

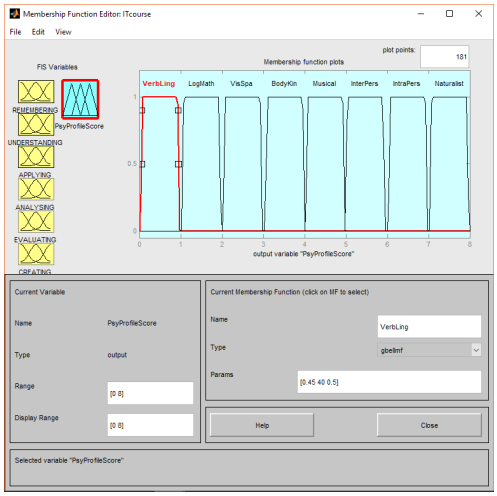


Fig. 4. Diagram of membership functions of Gaussian type "bell" for the fuzzy sets of the output variable corresponding to the student's intelligence type score.

For the rules from 1 to 8, each learner's intelligence type is described with selected fuzzy set functions of input variables that describe it's Bloom level characteristics (Remembering; Understanding; Applying; Analyzing; Evaluating; Creating) (Table 1).

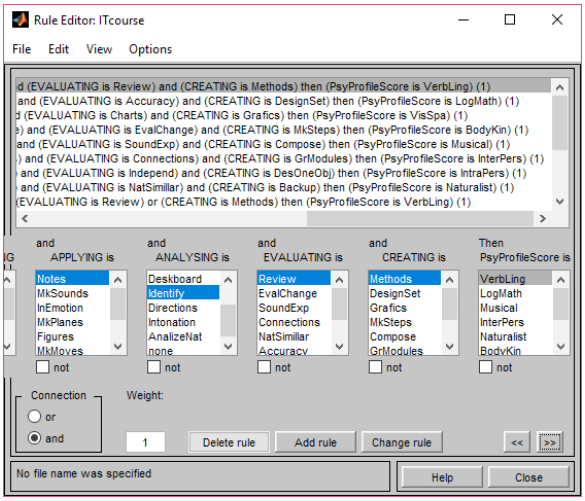


Fig. 5. Tool with Fuzzy system rules for psychological profile decision-making process of a system with fuzzy logic simulated in Matlab

The operation used for the various fuzzy sets, which are obtained for Bloom level characteristic is a logical "AND" to reflect the intersection of these fuzzy sets. One of the rules is used for example:

(Remembering==Rules) & (Understanding==DrawGrid) & (Applying==MkPlanes) & (Analyzing==DeskBoard) & (Evaluating==Accuracy) & (Creating==DesignSet) => ***(PsyProfileScore=LogMath)*** (1)

The second eight rules are designed to assess the results of the learner that specify not strong but acceptable results for each Bloom level. The logical operator "AND" is used in these rules for operation between the selected fuzzy sets, therefor to address potential some level exceptions allowed by Gardner's methodology. Following the description, the set of these 8 rules can be explained, with one of them:

(Remembering==Rules) || (Understanding==DrawGrid) || (Applying==MkPlanes) || (Analyzing==DeskBoard) || (Evaluating==Accuracy) || (Creating==DesignSet) => (PsyProfileScore=LogMath) (2)

This fuzzy inference system is developed to summarize the results of six Bloom levels, to make fast result as a value corresponding to the combination of degrees of matching the learners' attitude with the Gardner's intelligence types.

III. Results and Discussion

The experimental results are produced for an exemplary learners' psychological intelligence assessment on five Bloom levels as criterions. The psychological profile according the Gardner's intelligence types is decided upon its classification by fuzzy logic. The numerical simulation results of the experiment for a learner's profile with FIS result PsyProfileScore = 2.88 is depicted on figure 6.

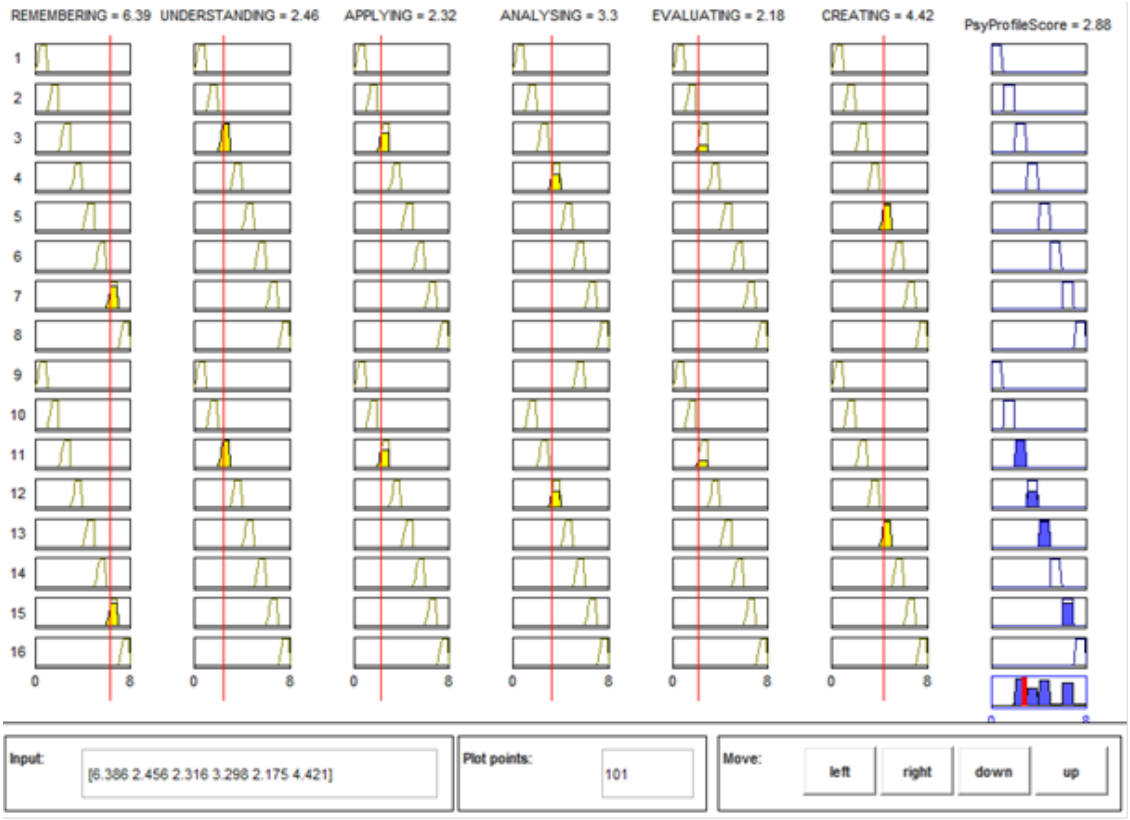


Fig.6. Graphically Matlab FIS expressed results for one exemplary learner in a short-term course. The score 2.88 (blue column) is connected to the output variable with the number 3. The learner's psychological profile in that case is assessed as Visual Spatial.

The simulation is resulted to the output variable with this number even if not relevant to all of the predefined rules. The learner's psychological profile for a course is classified Visual Spatial, one of the Gardner's intelligences.

IV. CONCLUSION

The Fuzzy Inference System is developed to summarize the results of all predefined Bloom's taxonomy levels in accordance with a system of rules is established. The input variables are designed to produce pre-course results about the most of the learners' psychological profiles. The flexible approach in formulating the fuzzy membership functions according the Bloom's taxonomy is important for the possible implementation of the developed Fuzzy logic procedure in similar education cases.

A proper analysis of the FIS results can be therefore successfully used in order to indicate which teaching technologies can be implemented for the education improvement purposes in short-term courses. A FIS procedure similar to the offered one can be designed to optimize some other real-time dependent signal processing assessment systems ([17], [18]).

References

[1] Atanasov, V.T., Smart Educational Cluster Conceptualization, Proceedings of International Scientific Conference, Shumen, October 02th -04th, 2018, pp.173-181, ISSN 2367-7902 http://www.aadcf.nvu.bg/scientific_events/dft2018/V.%20T.%20Atanasov.pdf

[2] Gardner H., Frames of Mind: The Theory of Multiple Intelligences, New York : Basic Books, ©1983. ISBN-978-0-465-02433-9, https://doi.org/10.1177/001698628502900212

[3] Zadeh, L.A., Outline of a new approach to the analysis of complex systems and decision processes, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 3, No. 1, Jan. 1973, pp. 28-44. https://doi.org/10.1109/TSMC.1973.5408575

[4] Zadeh, L.A., Fuzzy logic, Computer (Volume: 21, Issue: 4, April 1988), pp. 83 – 93. https://doi.org/10.1109/2.53

[5] Al Ganideh S. F.; El Refae G. A., Socio- psychological variables as antecedents to consumer ethnocentrism: A fuzzy logic based analysis study, 2010 Annual Meeting of the North American Fuzzy Information Processing Society, 12-14 July 2010, ISBN: 978-1-4244-7858-3, 2010,1-6. https://doi.org/10.1109/NAFIPS.2010.5548274

[6] Lefly M., Austin V., Match'em: using fuzzy logic to profile criminals, Proceedings of 6th International Fuzzy Systems Conference, 1997, vol.1, 305 – 311. https://doi.org/10.1109/FUZZY.1997.616386

[7] Di Nuovo A., Di Nuovo S., Buono S., Cutello V., Benefits of fuzzy logic in the assessment of intellectual disability, 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2014, 1843 – 1850. https://doi.org/10.1109/FUZZ-IEEE.2014.6891834

[8] Zhang J., Wang X., Mahfouf M., Linkens D.A., Fuzzy Logic based Identification of Operator Functional States Using Multiple Physiological and Performance Measures, 2008 International Conference on BioMedical Engineering and Informatics, 2008 , Vol.1, 570 – 574. https://doi.org/10.1109/BMEI.2008.181

[9] Petrova A. K., Methods and algorithms of decision making support in the process of experts' qualification evaluation for the completion of centers of professional certification, 2017 IEEE II International Conference on Control in Technical Systems (CTS), 2017, 384 – 386. https://doi.org/10.1109/CTSYS.2017.8109575

[10] Fernandez S., Ito T., Driver classification for intelligent transportation systems using fuzzy logic, 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), 2016, 1212 – 1216. https://doi.org/10.1109/ITSC.2016.7795711

[11] Hosseininezhad F., Hosseininezhad F., Balalpour M., A Fuzzy Expert System for performance evaluation of HRM with 360 degree feedback approach (Case study: An Iranian IT company), Computer Sciences and Convergence Information Technology (ICCIT), 6th International Conference, Dec. 2011. https://ieeexplore.ieee.org/document/6316664

[12] Doctor F., Hagras H., Roberts D. Callaghan V., A type-2 fuzzy based system for handling the uncertainties in group decisions for ranking job applicants within Human Resources systems, Fuzzy Systems, FUZZ-IEEE 2008, 2008, 481 – 488. https://doi.org/10.1109/FUZZY.2008.4630412

[13] Slavyanov K.O., An algorithm of fuzzy inference system for human resources selection tools, "Society. Integration. Education", Proceedings of the International Scientific Conference. Volume V, Rezekne, May 25th-26th, 2018. 445-454. http://dx.doi.org/10.17770/sie2018vol1.3311

[14] Vasileva M., Lambeva M., "Intelligent methods for attestation of scientific and teaching staff. Rasch model", Proceedings of the 11th International Scientific and Practical Conference "Environment. Technology. Resources", June 15-17, 2017, Rezekne, Latvia, Volume II, pp 171 – 175, http://dx.doi.org/10.17770/etr2017vol2.2665

[15] Bloom B.S., Engelhart M.D., Furst E.J., Hill W.H., Krathwohl, D.R., Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive domain. New York: David McKay Company, 1956

[16] Mamdani E.H., Assilian S., An experiment in linguistic synthesis with a fuzzy logic controller, International Journal of Man-Machine Studies, Vol. 7, No. 1, 1975, pp. 1-13. https://doi.org/10.1016/S0020-7373(75)80002-2

[17] Nikolov, L. "Wireless Network Vulnerabilities Estimation", International Scientific Journal "Security and Future", Vol. 2 (2018), Issue 2, pg(s) 80-82 http://confsec.eu/sbornik/cs-2-2018.pdf

[18] Bedzhev B., Tasheva Z., Stoyanov B., Bogdanov R., An algorithm for lossless signal processing (2007) Proceedings of the Third Workshop - 2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2005, art. no. 4062184, pp. 498-502. https://doi.org/10.1109/IDAACS.2005.283033

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 141-145

Neural Network Classification Method for Aircraft in ISAR Images

Krasimir Slavyanov
Department Computer Systems and Technologies
"Vasil Levski" National Military University
Shumen, Bulgaria
k.o.slavyanov@gmail.com

***Abstract* — This article offers a neural network method for automatic classification of Inverse Synthetic Aperture Radar objects represented in images with high level of post-receive optimization. A full explanation of the procedures of two-layer neural network architecture creating and training is described. The classification in the recognition stage is proposed, based on several main classes or sets of flying objects. The classification sets are designed according to distinctive specifications in the structural models of the aircrafts. The neural network is experimentally simulated in MATLAB environment. Numerical results of the experiments carried, prove the correct classification of the objects in ISAR optimized images.**

***Keywords*— artificial neural network, engine position, reference model.**

I. Introduction

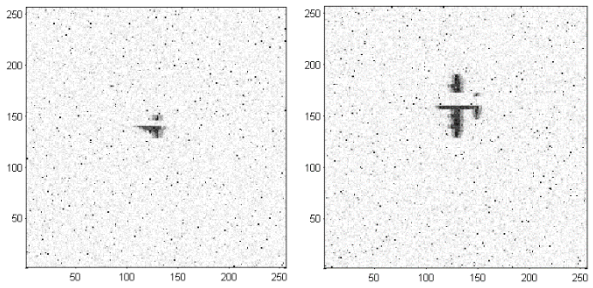
For the Inverse Synthetic Aperture Radar Systems (ISAR) the main purpose of the observed object's image is to be further classified and fully recognized as a type, model, specifications and owner even if the friend-foe system onboard or radio communications at all are not operational or used properly. In order to provide easy access to the ISAR principles of work a special dedicated airplane flight and ISAR simulation systems can be used to provide the environment for the ISAR system improvements ([1], [2], [3], [4], [5]).

Of particular importance for the proper recognition of objects in ISAR images is the classification of objects in the images according to particular details ([6], [7], [8], [9]). Several approaches can be used to analyse different data obtained for flying objects. The more successful classifications of the object, so they can be successfully merged with fuzzy logic tools to enable full recognition of objects according to specific rules [10].

II. Materials and methods

Preconditions and design.

For the purposes of the study, an ISAR simulation approach has been used ([1], [2], [6], [11]) in order to generate simulated ISAR images with additive Gaussian noise with constant zero mean and variance 0.01 and "salt and pepper" noise with density 0.015 (figure 1).



(a) (b)

Fig. 1. Reconstructed images in presence of additive noise for the aircrafts Rafale (a) and C-130 H (b) [5].

To achieve the most accurate classification results, it is necessary for the experiments several optimization procedures to be applied to the ISAR image after its generation. It is of particular importance in this research to obtain the most detailed images ([11], [12], [13], [14], [15]), shown of figure 2 and to maintain a target-specific database with reference models [6].

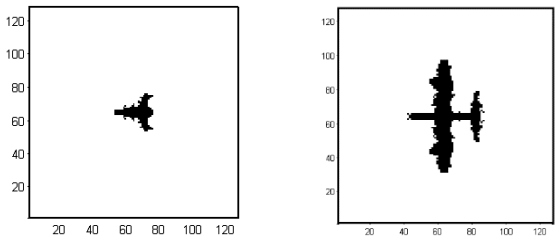


Fig. 2. Optimized ISAR images (128x128 pixels) of the aircrafts Rafale (a) and C-130 H (b).

This article proposes a method by which the observed aircraft is classified according to the location of its engines. R is the number of reference models to be compared. Patterns are binary matrices whose elements are a numerical representation of graphically described models of exemplary planes with different engine positions. With

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4074
© 2019 Krasimir Slavyanov.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

three models, the three different engine versions available, respectively, at the rear of the aircraft, on its wings and the variant where the engines are embedded in the fuselage of the airplane front or rear and do not stand out in its profile. Fig. 3 presents 9 ($R=9$) exemplary graphically-detailed solid aircraft models that form the model base in the numerical experiments. On figure 3 (a), (b), (c) are presented models with the positioning of the wing motors, in Fig. 3 (d), (e), (f) - the position of the engines at the rear of the airplane and in Fig. 1 (g), (h), (i) - engines built into the fuselage.

The object patterns are placed exactly in the middle of the frame - horizontally and vertically. Model matrices with the size 128x128 elements are formed as follows: If the pixel (i, j) of the graphical solid model is part of the structure, then the corresponding element (i, j) of the matrix is assigned a value of 1, otherwise the element is 0.

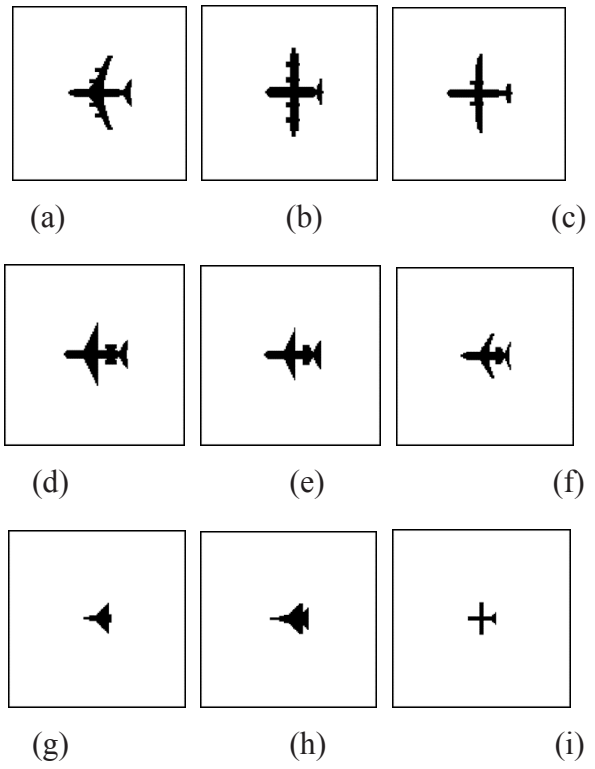


Fig.3. Graphical reference models of airplanes with engine position of types wing (a, b, c), rear (d, e, f) and built in fuselage (g, h, i).

For each reference model, a multiplication of the binary matrix of the original image S is performed with each of the R_F ($F=\overline{1,9}$) matrices.

$$S_m = SR_m \quad (1)$$

where the number of the current comparison model is $m=\overline{1,R}$.

The resulting matrices S_m will have non-zero elements only at the positions in which the S and R_m matrices have both non-zero elements. In this way, pixel address registration is performed by the matrices of a given reference model that coincide with the pixels of the image of the object being compared.The resulting R counts S_m are compared, and one of them has the maximum value. Its number corresponds to the model number (engine position) with which the object being compared has shown a maximum degree of matching between the non-zero positions of the matrices S and some of the matrices R_m.

A dataset is composed of 16 aircraft models (128x128 pixels) is used, the same like in [6] representing the planes: Eurofighter Typhoon, Pilatus 9M, Rafale, Mirage 2000, MiG-29, Gripen, Falcon 2000, F-22, F-18, F-16, C-130 H, Bombardier Q400, Boeing-747, Boeing-737, Boeing-707 and Embraer Legacy 600.

Neural network design and training

The basic idea in the proposed neural network classification approach is to compare the pixel intensity of the observed image with a set of predefined reference patterns of objects with different engine locations. In the process of realizing this approach, the task is to use the previously transformed and recognizable image in an input vector to be classified by a neural network as similar to one of the classes in the formed database.

To implement the described method, the training dataset shown on fig.1 is used. Reference matrices containing the pixel intensity values of the reference images are converted into vectors of 16384 elements by taking the respective column elements (128x128). The resulting vectors form a common "training" matrix called Training that is designed with dimension (16384x16) - 16 columns as the number of objects in the database subject to the recognition procedure and 16384 lines as the number of pixels in 1 image.

At this stage, a matrix of the "desired result" called Target, which is needed for the neural network training process, is built too. The matrix is with dimension 3x6 - 3 rows for the three types of objects classified by their size and 6 columns, because for each type of engine position (on the wings, rear and built in the fuselage) three exemplary reference models are designed. The location of the non-zero element in the main diagonal of each column corresponds to the class number that associates the corresponding recognition result.

In the third stage of the design, in accordance with the given task, a neural architecture consisting of two layers is constructed, which is built with the tools of the programming language Matlab (fig.4) and modelled in the Simulink environment (fig.5).

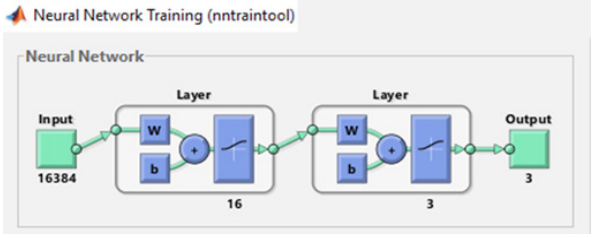


Fig.4. Neural network block diagram implemented in MATLAB.

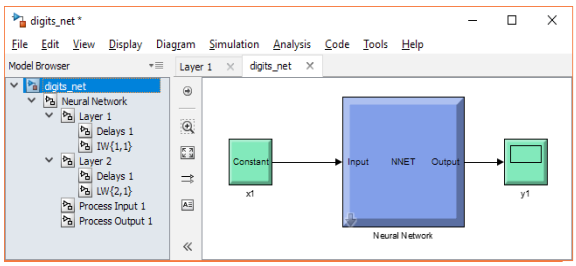


Fig.5. Schematic of Neural Network designed in Simulink.

According to the approach, the neural network is designed with backpropagation architecture of the two-layer error that performs associative memory functions in the presence of input sequence disturbances (Fig.6).

The first layer of the neural network is "hidden" and is made of 16 neurons with logarithmic-sigmoidal transmission function. These neurons form the subclasses, some of which are classified as the input vector. The inner structure of this layer is depicted in figure 7.

A delay line "Delays 1" is included in the layer structure, which converts the elements of the input sequence into an input vector.

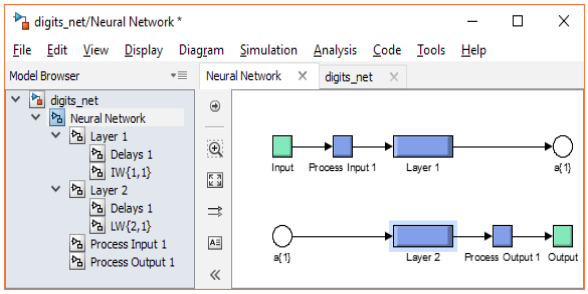


Fig. 6. Diagram of a two-layer Neural Network with right signal propagation and back propagation of the error realized in Simulink.

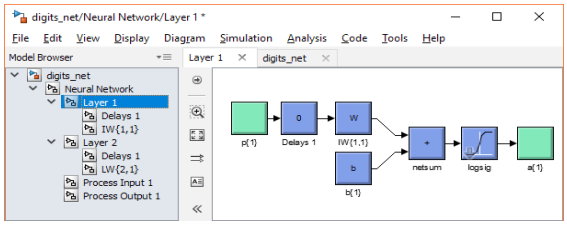


Fig.7. Structure of the first layer of Neural Network built in Simulink.

The logarithmic sigmoidal transmission function ensures high sensitivity and high resolution in the recognition process (Fig.8).

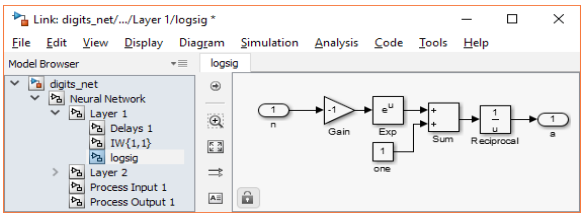


Fig.8. A structural scheme describing the mathematical model of Logarithmic-sigmoidal transfer function in Simulink.

On Figure 9 is presented the structure of the mathematical model and the physical realization of a logarithmic-sigmoidal type of function - the unfolded structure of the neuron inputs in the layer where the weight matrix IW is made by sixteen weighing vector weights whose specific values are determined at the stage of training of the neural network.

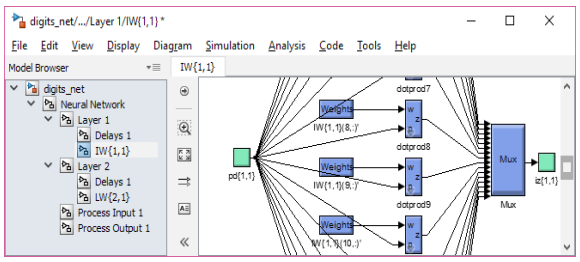


Fig.9. Structure of the input matrix of the first layer of Neural Network

The second layer contains neurons again with a logarithmic sigmoidal activation function. The number of neurons in this layer is set according to the final number of desired classes; in this case, the number is 3. The number of the neuron "winner" will correspond to the class ("on the wings", "rear" and "built in the fuselage") to which the current input vector is determined. The role of this layer is to classify the results of the first layer and to summarize and reduce them to the user-defined number classes. Its structure is analogous to the structure of the first layer with the stated differences. On Figure 10 is presented the structure of the mathematical model and the physical realization of a logarithmic-sigmoidal type function - the unfolded structure of the neuron inputs in the layer where the weight matrix LW consists of three weight weights, the specific values of which are determined at the stage of training of the neural network.

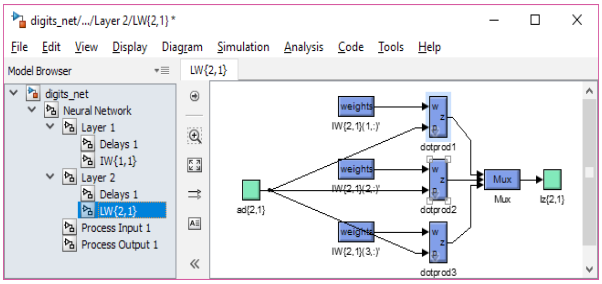
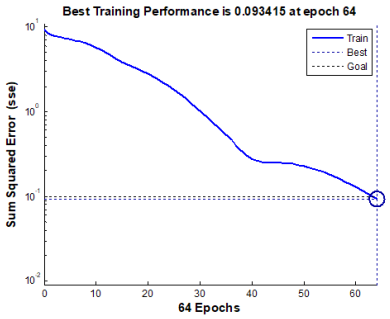
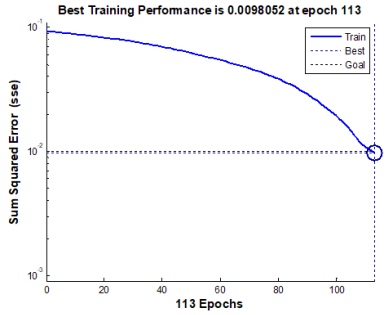


Fig.10. Structure of the input matrix of the second layer of the Neural Network

The fourth stage of the neural architecture realization involves a neural network training process, which essentially consists of adjusting the coefficients of the weight matrices of the neurons of the two layers. Inline algorithms and procedures for automated self-learning of Matlab neural networks are used to facilitate this task. The algorithm used for learning is backpropagation of the error. The admissible error level is 0.1, and the function of calculating this error is sse (sum squared error). The training is limited to 1000 epochs, and the results are presented in figure 11 (a). Network training is continued at higher requirements - the threshold for the permissible error level is reduced ten times to 0.01. The learning outcomes are presented in Figure 11 (b), the training process is considered complete.



(a)



(b)

Fig.11. Graphics of learning - the desired permissible error 0.1 (a) and 0.01 (b) for noise-free learning reached for 64 and 113 epochs respectively.

III. Results and Discussion

The evaluation for each object in the ISAR image is resulting in a number from 0 to 1, expressing the classification of the airplane /by its engine location/ according the neural network design and training. The results of 3 experiments with 3 aircraft models, carried out in Matlab environment are described on table 1. The results are improved advancing with repetitions of the experiments with the same airplane model according to the neural network's work.

Figure 12 (a) shows the graphical results of the third simulation with a C130 H reference model with initial parameters of the trajectory of the flying object as described in [10]. The result of the operation of the Neural Network - the position of the engines is on the wings (WingJet). The graphical results for the model Falcon 2000 (Rear Jet) and Rafale (Build In Jet) are depicted respectively on figure 12 (b) and (c).

Table 1 Results

Aircraft model used	Simul.№1	Simul.№2	Simul.№3
C-130 H	**0.72574**	**0.793476**	**0.823917**
	0.142658	0.06961	0.059254
	0.014988	0.019725	0.019126
Falcon 2000	0.081169	0.07581	0.021735
	0.706703	**0.796599**	**0.924063**
	0.016546	0.012988	0.010172
Rafale	0.010236	0.015908	0.017213
	0.009062	0.003987	0.003661
	0.989968	**0.993362**	**0.993443**

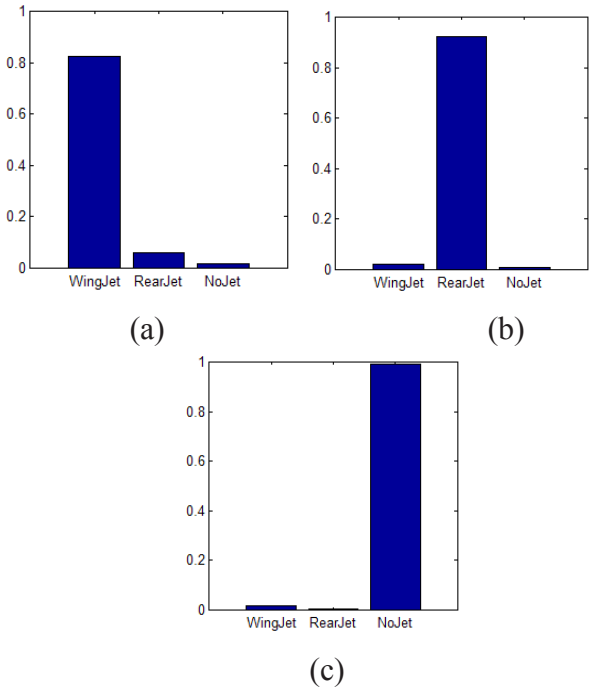


Fig.12. Neural network graphical results (for simulation № 3) for the models: C-130 H (a), Falcon 2000 (b) and the Rafale (c).

IV. Conclusion

In this article, the decision-making processes concerning the type of object being identified are modelled by processing the results obtained from the proposed method. The resulting radar images can be classified into several larger groups or pluralities of planes, depending on the location of the engines, at the stage of their recognition. In this way, partial information about the observed object can be obtained even if it is not fully recognized.

The classification method can be used for any other silhouette distinguishable details in different type of aircraft structures analyzed in ISAR images. In combination with different classification approaches like speed and power detection, airflow and temperature variations etc. the problem of complete identification of an aircraft without any communication lines can achieve more grounded results.

References

[1] Lazarov A., Minchev C., Correlation-autofocusing-spectral 2-D ISAR image reconstruction from linear frequency modulated signals, Digital Avionics Systems Conference, 2002. Proceedings, p. 27-31, Oct. 2002. http://dx.doi.org/10.1109/DASC.2002.1052984

[2] Lazarov A. , Minchev C., ISAR Signal Modeling and Image Reconstruction with Entropy Minimization Autofocussing, 25th Digital Avionics Systems Conference, IEEE/AIAA, 2006. http://dx.doi.org/10.1109/DASC.2006.313697

[3] Wei Q.; Hongzhong Z.; Jianxiong Z.; Qiang F., "High-Resolution Fully Polarimetric ISAR Imaging Based on Compressive Sensing", Geoscience and Remote Sensing, IEEE Transactions on (Volume:52 , Issue: 10), 02 January 2014. http://dx.doi.org/10.1109/TGRS.2013.2295162

[4] Xiaojian Xu; Narayanan, R. M., "Three-dimensional interferometric ISAR imaging for target scattering diagnosis and modeling", Image Processing, IEEE Transactions on (Volume:10 , Issue: 7), Jul 2001. http://dx.doi.org/10.1109/83.931103

[5] Dimitrov D., Kirov M., Using multi-carrier probing signals for detecting non-linear objects, Scientific research and educa-

tion in the air force -AFASES 2016, p. 247-250 http://dx.doi.org/10.19062/2247-3173.2016.18.1.33

[6] Slavyanov K., Minchev C., An algorithm for object classification procedure for ISAR images, Proceedings of 12th International Symposium on Applied Informatics and Related Areas, November 9, 2017, Székesfehérvár, Hungary, p.70-74. http://ais.amk.uni-obuda.hu/ais2017/proceedings/2017/AIS_2017_paper_17.pdf

[7] Benedetto F., Riganti Fulginei F., Laudani A., Albanese G., Automatic Aircraft Target Recognition by ISAR Image Processing based on Neural Classifier, International Journal of Advanced Computer Science and Applications,Vol. 3, No.8, pp 98-103, 2012.

http://dx.doi.org/10.14569/IJACSA.2012.030816

[8] Boulay T., Lagoutte J., Mohammad-Djafari A., Gac N., A Fuzzy-Logic Based Non Cooperative Target Recognition, Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference, Nov. 2012.

http://dx.doi.org/10.1109/SITIS.2012.66

[9] Tang, N.; Gao, X.-Z. ; Li, X., "Target classification of ISAR images based on feature space optimisation of local non-negative matrix factorization", Signal Processing, IET (Volume:6, Issue: 5), July 2012.p. 494 - 502

http://dx.doi.org/10.1049/iet-spr.2011.0286

[10] Slavyanov K., Minchev Ch., An Algorithm of Fuzzy Inference System for ISAR Image Classification, Environment. Technology. Resources, Proceedings of the 11th International Scientific and Practical Conference.Volume II, 154-158, http://dx.doi.org/10.17770/etr2017vol2.2599

[11] Minchev C., Slavyanov K., An opportunity for improved modeling and information analysis in ISAR systems, Machines. Technologies. Materials., Vol. 9 (2015), Issue 6, p. 24-27 https://stumejournals.com/journals/mtm/2015/6/24

[12] Parmar J. ; Patil S."Performance evaluation and comparison of modified denoising method and the local adaptive wavelet image denoising method", Intelligent Systems and Signal Processing (ISSP), 2013 IEEE http://dx.doi.org/10.1109/ISSP.2013.6526883

[13] Russo F., "A fuzzy approach to performance measurement of grayscale image denoising algorithms", Intelligent Signal Processing (WISP), 2015 IEEE http://dx.doi.org/10.1109/WISP.2015.7139159

[14] Otsu N., A Threshold Selection Method from Gray-Level Histograms, IEEE Trans., Systems, Man, and Cybernetics, vol.9, no.1, p.62-66, 1979. http://dx.doi.org/10.1109/TSMC.1979.4310076

[15] Minchev C., Slavyanov K., An algorithm for ISAR image optimization procedure, Bulgaria, National Military University, Proceedings of Internatinal scientific conference, 2015, ISSN 2367-7902, p.177-184. http://www.aadcf.nvu.bg/scientific_events/papers/NS_2015.pdf

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 146-153

The Basics of Event-Related Management of Safety and Quality in Economics

Evgeny Dmitrievich Solozhentsev
Honored Scientist of the Russian Federation,
St. Petersburg State University of Aerospace Instrumentation
Professor of the Institute of Business Technologies
URL: www.topeconomics; mail: esokar@gmail.com

Abstract. **The work forms a new breakthrough scientific trend in the management of safety and quality in the economics based on the introduction of new objects and management criteria, new knowledge and models, new tasks, special software and digital management.**

The paper offers the new management objects: State authorities, socioeconomic systems and projects, process quality management in the socio-economic life of a person, the safe living space. We use the management criteria of safety and quality.

The paper introduces the new knowledge for management in economics: the methodology of safety and quality, new types of Boolean events-propositions, scenarios of systems failure, new types of logical-probabilistic models, results of research of various systems and educational program.

New management tasks are also considered, special software are described and the essence of digital management in economics is stated.

Keywords: total safety management, managed objects, business enterprises, economics, safety and quality criteria, event-related digital management, structural complex systems, new knowledge and tasks, invalidity, special software.

I. INTRODUCTION

Many papers describe the unsatisfactory management of State and economics. Many economists are trying to find a way to overcome this critical situation.

The present theory of economic management is inadequate. A long time there are no fundamental achievements in State and economics management. Economics management is performed without mathematical methods and models on the basis of ephemeral concepts and aims, using "unwritten rules", "manual management", or "give me more money" principle, promises, ephemeral programs of economic growth and reviving industry.

After analysis of ephemeral management methods, management objects, State officials, army and policy structure, the education system, economic and academic sciences, a conclusion was made about impossibility to remedy the situation without a new outlook, new knowledge and new tasks in economics management.

The presented study outlines the scientific foundations of a new breakthrough trend in management of safety and quality in structural complex systems in economics.

Aim is the creation of a new breakthrough scientific trend in economics "Event-Related Management of Safety and Quality in Economics" on the basis of new knowledge and new tasks solution.

New objects and criteria of management are State authorities, socioeconomic systems and business activities, quality management of processes in the socio-economic life of a person and the safe living space. Safety and quality of system have been selected as management criteria.

New knowledge. For management in economics the new knowledge are introduced: the methodological and methodical foundations of safety and quality management, new types of Boolean events-propositions, scenarios of systems failure, new types of logical-probabilistic (LP) models.

New tasks. For management in economics the new tasks are proposed: modeling, analysis and management of one system and a group of logically connected systems (models); management of the State and evolution of systems; assessment of management system quality.

Digital management in economics is defined as a technology of wide introduction of innovations and new tasks solution on the basis of the unified set of new knowledge, methods, models, tasks, technologies and software. Digital management has been connected with innovations and investments.

Scientific novelty. The new breakthrough scientific trend "Event-Related Management of Safety and Quality of Structural Complex Systems in Economics" is created. In management of economics we introduce: the methodological and methodical foundation, new objects and criteria of management, new knowledge and tasks, special software and the set of LP-models for the management of safety and quality of systems.

Practical value. We offered the way to overcome critical situations for economies and successful development. Authorities have to develop and use new knowledge and solve new tasks in economies. We have created the LP-models for the State analysis and management of economic systems under safety and quality criteria. Special software was adapted for management purposes. Additional education program for

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4039
© 2019 Evgeny Dmitrievich Solozhentsev.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

economists was developed. We proposed a unified system of knowledge, models, and software for event-related digital management.

Publications. We are engaged in research of event-related economic management for almost 20 years. We have published 8 books, including 3 books in English (Springer and Cambridge Scholars Publishing) and more than 30 papers, including 10 Scopus publications. Our results were presented and discussed at International Scientific Schools "Modeling and Analysis of Safety and Risk in Complex Systems" (MASR 2001–2016, Saint-Petersburg, IPME RAS) and use in educational program of the State University of Aerospace Instrumentation.

The fundamentals of a new scientific trend were published in a large number of papers in Russian and English with the Scopus index in different years and were scattered for economists. Papers describe methodology of event-related digital management of systems, new objects and criterion, new knowledge and tasks, substantiations on digital management. In the paper, these results are brought together.

Concepts and statements of eminent scientists about management of economics. This study was performed for the first time. In development a scientific trend in economics, choice a mathematical apparatus and constructing risk models of systems, separate concepts, principles and statements of eminent scientists about management in economics were used.

Norbert Wiener and John von Neumann, the founders of cybernetics, believed that mathematical methods for management of socio-economic systems should be based on logic, set theory and combinatorics.

Albert Einstein, the famous scientist, wrote that no problem can be solved at the same level at which it arose.

Rudolf Kalman, the author of the Kalman's filter, believed that the problem "data à model, explaining data" should be considered as basic for any branch of science.

Nobel laureate James Buchanan considered the relationship of government, corruption and public opinion in the State.

Nobel laureate J. Heckman proposed an analytical apparatus of mathematical statistics for the analysis of socio-economic processes in the State.

The Nobel dynasty was guided by the principle of social justice. Nobel family spent a significant part of their profits for workers: they paid decent wages, built houses, kindergartens and schools, provided free medical services.

Li Keqiang, the Prime Minister of China, said, technological innovation and innovation in management are the same.

Kate Raworth, Professor at Oxford University, proposed a new economic model in the form of a "safe living space" ring (doughnut). Outside the ring – the ecological limits of the earth are violate: dangerous climate change, water pollution, etc. Inside the ring - the resources for a good life: food, clean water, housing, energy, education, etc. are lack.

Albrecht W., Wernz G., Williams T., American lawyers, believed that everyone was capable of fraud if circumstances put pressure and there is no sufficient control.

Isaac Newton, the great English scientist, believed that concrete examples are no less instructive than theory.

Williams Okkam, an English philosopher, believed, we should not complicate the model unnecessarily. A simple model is more likely to be correct. Sets and logic allow you to build the most simple and transparent models.

Robert Stevenson, an American scientist, believed that errors in projects (development programs) of complex systems and projects are inevitable and operational tests are necessary to identify and eliminate errors.

Stephen Robbins and Mary Coulter, authors of the textbook "Management", set out the components of system management, which include the functions of planning, organization, management and control, which can be the basis for assessment the quality of management systems in economics.

I. Ryabinin proposed the logical probabilistic calculus for reliability theory in structural complex systems in engineering [1]. We applied this theory to economic systems by introducing multi-states of the system instead of two states (refusal / failure) in technology.

N. Hovanov developed a method of randomized summary indicators for ranking systems and synthesis the probability of events by non-numerical, incomplete and inaccurate expert information, which is used to construct safety and quality LP models in economics [2].

The author defended his doctoral thesis in the USSR "Fundamentals of Building Systems for Automated Development of Complex Machines" (1983, Institute of Cybernetics, Academy of Sciences of Ukraine, Kiev) [3]. The theory and results of thesis are developed for management in economic systems.

II. METHODOLOGICAL FOUNDATIONS OF MANAGEMENT IN ECONOMICS

The fundamentals of a new scientific trend were published in a large number of papers with the Scopus index and were scattered for economists [4–8]. They include methodological foundations of event-related management of systems, new objects and criterion, new knowledge and tasks, substantiations on digital management. In the paper, these results are brought together, excluding some mathematical descriptions and limiting themselves to references.

We propose the methodological foundations of management of structural complex systems in economics, which are considered as new knowledge in management of economics. They are formulated in the form of provisions that we will consider as event-propositions and new knowledge in management. These provisions are common to all economies and states. To develop a new scientific trend "Event-Related Management of Safety and Quality in Economics", the following statements on management methodology in economics were proposed and used:

1. It is impossible to increase efficiency in economics without a new outlook, new knowledge and solving new tasks.

2. All the troubles of the economics are caused by management.
3. Safety and quality are the prerequisites of existence of all systems. We propose to perform management in economics using safety and quality criteria.
4. New scientific tend in the management of safety and quality in economics should use new knowledge: methodological and methodical bases of management of economics, new Boolean event-proposition, risk scenarios for system failure, new risk models, new tasks, special software and assessment of the management systems quality.
5. New objects for management should be: public authorities, socio-economic systems, process quality management in the socio-economic life of a person and a safe living space. The management of these systems should be part of the management of the economics and the State.
6. To manage economics, system structure should be used – to introduce connections *AND, OR, NOT* between the elements of the system and the purpose of the system.
7. The management of systems (objects) in the economy should be carried out according to the scheme: modeling → analysis → management. Management on quality and safety criteria is carried with the participation of the manager.
8. The safety of the system is defined by the concepts of "risk" and "acceptable safety". The quality of the system is determined by the invalidity of system's indicators.
7. Consider management and optimization are close in meaning: optimization, according to safety and quality criteria, is management; management by criteria is optimization.
8. The model of invalidity should be built on the invalidity indicators of one state of the system.
9. Systems and their components should be associated with events and logical variables.
10. Consider internal events as factors of success/failure of management in economics are: theory and methods of management, objects of management, managers (officials), army, education system, economic and academic sciences.
11. Consider the management of the system's evolution as the management of the movement along the program trajectory at the stages and correction when deviating from it.
12. To manage the system in time for signal events with a correction of the probabilities of initiating events in LP-models of safety and quality.
13. Digital management is considered as a technology of wide and rapid introduction of new scientific trend in economics to solve important problems.
14. The technology of the digital management in economic systems increases their efficiency, safety and quality.
15. Accept the dialectic of the subjective and objective in invalidity: the system requirements are set subjectively, and the compliance of the system with these requirements is objective factor.
16. Calculate significances of initiating events and manage changes of the probabilities of these events, investing funds, improving staff skills, changing the structure of the system and performing reforms.
17. The system's management modeling in the economics found that without scientists and public opinion the problems of economics cannot be solved.

Boolean propositions were the basis for the creation of mathematical logic. They were developed in events in engineering and served as the basis for the theory of reliability. Boolean propositions are almost never used in economics.

The main requirements for any systems are their quality, efficiency and safety. Hence, a mathematical apparatus and models are needed to quantify the criteria for efficiency, safety and quality of systems. To the greatest extent, the apparatus of logical-probabilistic calculus is suitable for this, besides providing a unified approach to the calculation of these criteria in all objects and systems. System security is defined by the concepts of "risk" and "acceptable security", the quality of a system is determined by the invalidity of its indicative events, its effectiveness – by the mathematical expectation of the risk of loss of assets or by price of the system on the market.

Methodical bases of safety and quality management of structural complex systems in economics are formulated in the form of provisions that will be considered as event-propositions and new knowledge in management.

1. Any database can be transformed into knowledge base in the form of a system of logical equations.
2. LP-analysis of safety and quality is transparent and this is the advantages of LP-models.
3. The LP-model of the invalidity (quality, safety) of any system can be always built on the basis of invalid indicators of its state, taking into account the indicators of the external and internal environment of functioning the system.
4. When constructing LP-model of system safety and quality, external and internal influencing factors should be separately identified. This will allow construct an accurate LP-model, combining several systems with a correct account of repeated external event factors.
5. Logical variables become dependent when they are placed into one logical function. Therefore, we must perform the orthogonalization of logical functions so that the terms in the logical function are independent.
6. For each system, the following safety and quality models should be consistently constructed: structural, logical and probabilistic models.
7. The problem is solved for any logical complexity of systems.
8. The dynamics of LP risk models is ensured by the correction of the probabilities of initiating events when signal events occur.
9. The connection of various systems should be ensured by correct consideration of repeated events included in the LP-models of safety and quality of systems.
10. It is necessary to control and analyze the ephemerality of internal initiating events and factors of success /

failure of the management system in the economics.

11. The overcoming critical situation in economics and every system should be performed on the basis of new knowledge and the solution of new tasks.
12. In order to manage the safety and quality of every system (object), it is necessary to select successively: objects and criteria, new knowledge, LP-models, new tasks, LP-analysis methods for calculating contributions and significance of the initiating events, the technique of LP-management of system's state and system's evolution.
13. To estimate the quality of the management system on the LP-model.
14. Be informed with examples of computational research on safety and quality management.
15. Install special software Arbiter and Expa, having certificates, for modeling safety and quality.
16. Pass a licensed additional education program for economists in the management in economics.
17. Connect to the computer network event-related digital management.
18. To analyze the relationship of event-related digital management with innovations and investments.
19. Use the unified set of models, knowledge, tasks, and special software for event-relate management.

III. THE EPHEMERAL MANAGEMENT IN ECONOMICS

In managing economics, internal influencing events-factors of success/failure were identified: theory and methods of management, objects of management, managers (government officials), education system, economic and academic sciences.

Internal influencing events-factors of economics management in Russia were analyzed. All events-factors of management system are ephemeral, i.e. they are illusory phantoms. The following aspects of the economic safety of Russia continue to deteriorate: the national debt is growing, the scientific and technological potential is decreasing, the industrial foundations of the national economics are being destroyed, difference in income of various groups of population is growing very rapidly, finances are flowing out of the country, etc. We have performed the analysis of the ephemeral economic system of management in Russia in order to cope with the critical economic situation [4].

Management methods. Management use ephemeral concepts and targets, is based on promises and slogans, appeals, ephemeral programs of economic growth, increasing labor efficiency and industrial revival; management by "unwritten rules", "manual management", "give me more money" principle. This situation inevitably leads to corruption.

Objects of management include aims, objectives, processes in economics. A lot of factors are ephemeral: GDP, the share of machine building in industrial output, investment volume and expenses on science, percentage of GDP. Being taken together, these factors make the problem of management as multi-criteria one, which is impossible to decide. Each of these factors depends on lots of various other factors and is not a criterion itself. If we build the time series for such factors, we will not get an answer what and how we are going to manage.

State officials are called ephemeral due to the following reasons: there are 1,4 times more state officials for every 10,000 people in Russia than in developed countries and 2,5 times more than in the countries with the average level of economic development; the salary of state officials is 14–15 higher than employees with university degrees; the increase in the numbers and salaries of state officials did not make the economics more effective. Due to these reasons managers are not interested in changing management in economics.

The army and the police are called so because they are ephemeral as state officials due to their large strength, high salaries and pensions and low efficiency of their activities**.** So, army and police are not interested in changing management in economics.

The educational system. We have destroyed the industry-specific principle of education, which was used in the Soviet Union when the educational system served not the interests of particular people, but trained the specialists which the country really needed. School leavers chose the university according to personal preferences and they knew in what position and where they would work later. There was no hype about certain jobs. It was believed that the market itself will determine which universities are going to develop. Actually, it was not the market, but the universities administration, which had their own ideas about market demands. Every university started training future lawyers, economists and state officials. Universities earned money on education fees.

The economic science. The "homo economicus" model dominates at present. This is an autonomous individual who seeks only maximum self-interest. In order to simplify the analysis, political and social factors are not considered. However, the production of new products and services requires non-standard solutions based on cooperation. This presupposes the expansion of social links between employees and employers. By this way non-financial social risks are increased. The importance of social environment and the standard of living is increasing.

The academic science. The situation in the academic science is really dramatic. The science community has lost the unity; it looks like islands with very few connections between them. Complicated problems are not tackled. The science community is getting older. The gap between generations of researchers is becoming wider. We must create favorable conditions for researchers. Destruction of science will lead to the decay in education and to the problems with sustaining the present complex infrastructure and developing new technologies. Man-induced disasters will become a usual thing.

A way out from critical situation. Ephemeral theory and ephemeral management objects cannot change management in economics in Russia and other countries. The ephemeral state officials, army, the educational system are not interested to change the situation and science cannot do this [4].

System safety and quality requirements are the main conditions of surviving as a nation. The fate of Russia depends on the fate of the Russian science. The main problems of the Russian economic science include the clan system, plagiarism of dissertations and red tape. Applications for the grants of the Russian Foundation of Fundamental Research supporting the development of the new research area have been rejected for over 10 years. The papers submitted to leading economic journals have also been rejected under various pretexts, though during the same period 10 Scopus papers and 2 books devoted to this topic were published in abroad.

The fate of Russia depends on the emergence of new knowledge and tasks in economics [4–8]. Ephemeral methods and objects of management in economics are typical not only for Russia, but in other countries, as well. The difference lies in the fact that such components of the management system as state officials, army and policy, the educational system, economics and academic science are in satisfactory condition, which provides the leading technological position of these countries.

IV. THE ESSENCE OF EVENT-RELATED DIGITAL MANAGEMENT

Studying a lot of literature sources, we found *digital economic management* has not been touched upon**.** This can be explained by the fact that there are no mathematical models in management of economics. Event-related digital management is based on new knowledge, LP-models and corresponding software [9]. The introduction of digital management is necessary due to a number of reasons. Problems of management in economics are the most vital and widely spread both at the top level of management, and at the level of regions, cities and enterprises. Problems of management have a complex interdisciplinary character, novel mathematical tools and are characterized by the use of intellectual knowledge. Problems of management have a high arithmetic and logical computational complexity and cannot be solved without special software. Event-related digital system management in economics employs new knowledge, unified techniques, mathematical models and special Software [4, 9].

Digital management in economics is considered as a technology of wide and fast introduction of the new scientific trend for solution important problems of economics. Digital management provides system management based on new knowledge, LP-models of safety and quality, technologies and special software. Digital management has a single unified set of methods, models, new knowledge, new tasks, technologies and software.

Links between the digital management system, innovations and investments are illustrated in Fig. 1. Here 1, 2, 3 – educational systems, economic and academic science; 4 – new knowledge and tasks; 5 – innovations for systems and management, 6 – management systems; 7 – systems in economics, including mining and selling raw materials; 8 –investments from systems.

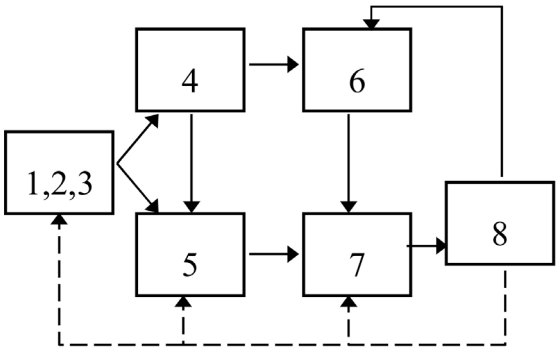


Fig. 1. The links between digital management, innovations and investments.

Specialists and investments are needed for the implementation of event-related digital management. The educational system trains specialists for systems, economic and academic sciences. The latter play a role in teaching students and develop new knowledge, innovations and technologies, improving the criteria of safety and quality of systems.

Investments derive from innovations in industry and system management. They are distributed for the education, economic and academic sciences; development of systems.

Digital economics and digital management. We had been conducting research in "Event-Related Management of Systems in Economics" more than ten years before the importance of developing "digital economics" was acknowledged nationwide. When "digital economics" had been acknowledged as the economic development technology, the place of the new scientific direction in digital economics was determined and its name was adjusted to "Event-Related Digital Management of Safety and Quality Systems in Economics". Due to the novelty of the research field, its complex character and computational complexity, there were practically no opportunities for large implementation. The situation changed when the government started support the development of digital economics.

V. NEW OBJECTS OF MANAGEMENT IN ECONOMICS

We study the following structural complex systems (objects) in economics [4, 9].

State authorities (in brackets – the number of objects)*:* ministries (21), State agencies (35), services agencies (15), State corporations (2), State extra-budgetary funds (3), State Duma, Federation Council, governments and the legislative assemblies of the regions and cities.

Socioeconomic systems and projects. The State allocates the budget to socioeconomic systems (SES) and projects. The State incurs losses in SES and projects due to corruption, drug addiction, bribery, embezzlement, decisions taken by unwritten rules and not by law, excessive expenses on social and military projects. The SES groups can be identified:

Group SES-1 contains SESs with highest importance for the State, aimed at reducing the loss of funds and increasing revenues: management of innovations in the

country, regions and companies, management of risk in banks and capital reservation by Basel III, management of production systems quality and product quality by WTO requirements, monitoring and management of crediting process in banks, opposition to bribery, corruption and drug addiction; assessment the management system quality.

Group SES-2 includes SESs which are complex for the State and the regions and depend on a number of ministries, departments and agencies (the systems of education, healthcare; agriculture, ecology, industry, communications, sport, transport, finances, economic development, energy).

Group SES-3 consists of the enterprises whose success depends mainly on their desires and capabilities. They include industrial, service, trade, transport, banking, educational, medical, etc. companies. In our papers we analyzed the management of risk and efficiency of a restaurant, management of crediting process in banks, etc.

Processes quality management in socio-economic life of a person. Such processes include: treatment of diseases, training at the institute, child care and others. A characteristic feature of these processes is the participation of several subjects and their respective infrastructures. For example, the process of cataract eye surgery is reviewed and analyzed with the aim of improving its quality. Management is performed according to the criterion of quality, taking into account the following event factors: 1) medical personnel qualifications and hospital infrastructure, 2) the patient's condition and the infrastructure outside the hospital in the postoperative period, 3) deficiencies in the organization of treatment in the State. Evaluation, analysis and quality management of a cataract surgery-treatment process is performed on event-related logical-probabilistic model. Structural, logical and probabilistic models of the quality of eye cataract surgery are built. The estimation of the probabilities of initiating event-factors is performed by the method of randomized summary indicators for non-numeric, inaccurate and incomplete expert information by the method of statistical data identification on the results of surgery-treatment success.

Safe space of humanity. Kate Raworth from Oxford University notes: in the 21th century the economics strived to be a science based on an erroneous picture of humanity [7]. The dominating model — "rational homo economicus", self-interested, isolated, calculating tells us more about economists than about common people. The loss of the goal encouraged the task of achieving permanent economic growth. Kate Raworth reconsiders the foundations of economy. She presents the economic model, which consists of two rings. If we go beyond the outer ring, we will leave the ecological limits of the Earth and face climate change, the depletion of the ozone layer and water pollution. If we go beyond the inner ring, we will not have enough resources for good life: food, clean water, accommodation, sanitary conditions, energy, education, healthcare and democracy.

VI. NEW KNOLEDGE IN MANAGEMENT IN ECONOMICS

Methodological (19 concepts and positions) and methodical (19 definitions and solutions) foundations of management in economics were described earlier as new knowledge. The following new knowledge is also introduced for managing economics: Boolean utterance events, system failure scenarios, and LP risk models of systems.

Boolean events-propositions for management. The ideas of "events-propositions", which were developed in the works of G. Boole, P. Poretsky, S. Bernstein, A. Kolmogorov, V. Glivenko and I. Ryabinin, were elaborated further and extended. To manage the economics seven new types of Boolean events-propositions were introduced**:** about the failure of subjects and objects, signal events, invalidity events, conceptual and indicative events, latent and repeated events, groups of incompatible events [4-6].

The logical addition of events forms a derivative event. In management tasks in economics the probabilities of success / failure, danger / safety, validity / invalidity of events are used.

The system failure scenario is a description of events-propositions affecting system failure, as well as their logical connection with each other and system failure.

New types of system failure risk LP-models. For the purposes of system management in economics we use structuring**,** establishing logical connections *AND, OR, NOT* of the elements with each other and with the target of the system. Seven new types of safety and quality models were proposed and tested: hybrid models of system failure; invalid models of risk and quality; conceptual models of predicting system development; indicative models of system state danger, models used for management of system's state and evolution; models for system management quality assessment.

LP-risk models. LP-models should be used for comprehensive analysis and management of one system. Different systems are connected by repeated events-propositions. Criteria of safety and quality are management criteria.

Safety and quality criteria are simply calculated and analyzed on the probabilistic risk model obtained after orthogonalization of the logical risk model of the system. The problem is solved for any complexity of the logical model.

The dynamic character of the LP-model of systems is ensured by the correction of the probabilities of events – proposition for signal events, which indicate the necessity of changing the probabilities of initiating events in LP-models. The probabilities of probabilistic model are adjusted by non-numerical, inexact and incomplete expert information.

VII. THE NEW CRITERIA OF MANAGEMENT IN ECONOMICS

Event management of the safety and quality of structural complex systems in economics is performed

according to the criteria of safety and quality. Safety of a system is defined by the concepts of "risk" and "acceptable security", the quality of a system is determined by the invalidity of its indicator events, effectiveness of a system – by the mathematical expectation of the loss of assets or by market price of the system. The criteria of safety and quality have the following advantages [4, 5, 9]:

1. Providing the main requirements for any systems – a quantitative assessment of the criteria of safety and quality.
2. Use for calculations and analysis the criteria of the known logical-probabilistic calculus.
3. The use of the methodological foundations for managing both the quality and safety of systems.
4. The use of a single unified approach for the construction and study of models of quality and safety for all structural complex systems in economics.
5. The possibility of combining any number of different systems (models) into one joint system (model) and solving new actual tasks in economics.
6. The possibility of creating the digital management of safety and quality in economics based on a unified set of knowledge, models, tasks and software.
7. The possibility of forming a new breakthrough scientific trend in economics and the economic sciences "Event-Related Management of Safety and Quality in Economics".

VIII. NEW TASKS IN MANAGEMENT IN ECONOMICS

New tasks are designed for:

• Theoretical forecasting and development of system evolution programs;

• Management of state and development of functioning systems;

• Operational management of systems in case of unexpected events.

With the designations: *M* – modeling, *A* – analysis, *C* – management, the scheme of management process of the economic system in general form can be represented:

$$(M_1 \rightarrow A_1 \rightarrow C_1) \rightarrow (M_2 \rightarrow A_2 \rightarrow C_2) \rightarrow \ldots \rightarrow (M_i \rightarrow A_i \rightarrow C_i) \rightarrow \ldots. \rightarrow (M_n \rightarrow A_n \rightarrow C_n),$$

where $i = 1,2, \ldots, n$ are stages or management time.

The created and adapted Arbiter and Expa software systems allow solve following new and previously unknown management tasks in economics:

1) Modeling, analysis and management (MAM) of the safety and quality of one system;
2) MAM of safety and quality of several systems combined into one common system (logical model);
3) MAM on common logical model of safety and quality of systems connected by logical operations *OR, AND, NOT,* with different outcomes;
4) MAM on common logical model of different systems with the correct account of repeated initiating events.
5) LP-management of system's state. We propose to perform the LP-management of the state of the systems, based on the quantitative LP-analysis of initiating events contributions in the following sequence: perform a quantitative risk analysis of initiating events contributions to safety and system risk, make decision about changing the probabilities of significant events, allocate resources to change probabilities of selected events, increasing staff experience.
6) LP-management of system's evolution [3, 8, 14, 15]. We propose to perform management as a complex object control (Fig. 2). The system's movement is managed along the prescribed trajectory and corrections are made in case of deviations: $j=1, 2,\ldots,n$ – stages of development; R_j – criteria of safety or quality of the system, U_j – managing actions (resources), W_j – correcting actions (resources). A system is transferred from its initial state *A* into the final state *B* along the chosen trajectory *A – B*.

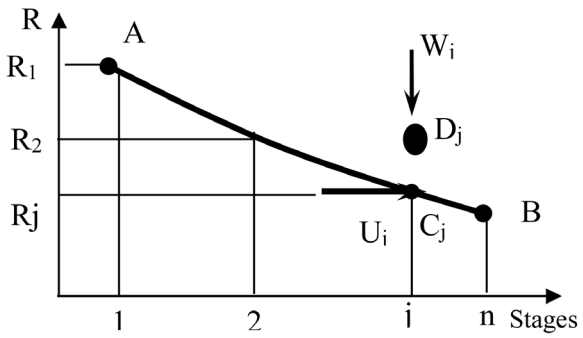


Fig. 2. Trajectory of system's evolution and control.

Structural, logical and probabilistic risk models are developed for a system. The criteria *R* is calculated at each stage; contributions of events into system risk are analyzed. While developing the program of system's evolution management the values of *R, W, U* at stages *n* are determined. Resources are required for the realization of *R, W, U, n*. The LP-model of the process failure is built by the logical addition of failure at development stages.

7) The assessment of the quality of management systems in economics is made by the LP-model, built by the structural management model, which includes events-propositions about invalidity of the functions of planning, organization, management and control. Events-propositions have a measure of invalidity in the interval [0, 1]. Each function consists of events-propositions for sub-functions [16].

IX. SPECIAL SOFTWARE FOR MANAGEMENT IN ECONOMICS

Systems in economics have a lot of events. The orthogonalization of the logical function of system risk for obtaining the probabilistic function of safety and quality for real systems is possible only with the help of special software. For the purposes of digital management of systems in economics the following special certified software can be used: *Expa* – for the synthesis of probabilities of events-propositions [17]; *Arbiter* – for structural logical modeling [18], as well as a licensed training course for further education of economists [19].

Works [8 – 15] provide about 30 examples of using software *Arbiter* and *Expa* for various objects and systems. The results of system management with real data helped us establish the following facts: socioeconomic problems cannot be solved without scientists and public opinion. To improve the efficiency of the country's innovation system, reforms in education, science and economics are needed.

XI. CONCLUSION

The results of this study are follows:

1. The new breakthrough scientific trend in economics and economic science "Event-Related Management of Safety and Quality in Economics" is formed.
2. New objects (systems) of management in economics: public authorities, socio-economic systems, processes quality management of the socio-economic life of person, safe living space are introduced.
3. New knowledge for management in economics: methodological and methodical bases of safety and quality management, Boolean events-propositions, systems failure risk scenarios, LP-models of safety and quality of systems, examples of real systems management are introduced.
4. We offered the criteria of management in economics having some advantages.
5. New tasks in economics for modeling, analysis and management of one system and a group logically related systems (models) with different goals are offered.
6. The essence of the digital management of safety and quality of systems in economics is stated. The relationship of event-related digital management with innovations and investments is considered.
7. The special software *Expa* and *Arbiter* for safety and quality management in the economy are described.

REFERENCES

[1] Ryabinin I.A. Reliability and safety of structurally complex systems (2nd ed). – SPb.: Publishing House of St. Petersburg University. 2007. 276 p.

[2] Hovanov N., Yadaeva M., Hovanov K. Multicriteria Estimation of Probabilities on the Basis of Expert Non-numerical, Inexact and Incomplete Knowledge / European Journal of Operation Research. –Vol. 195. N 3. 2007, p. p. 857 – 863.

[3] Solozhentsev E. D. The Basics of Building a System of Automated Debugging of Complex Machine-Building Objects / Author's Summary of the Doctoral Thesis. Kiev: Ukraine Institute of Cybernetics, 1982.

[4] Solozhentsev E. D. Ephemeral and digital management of safety and quality in economics / Problems of risk analysis. Vol. 15, No. 5, 2018, p. p. 58–77.

[5] Solozhentsev E. D. The Management of Socioeconomic Safety. – Cambridge Scholars Publishing, 2017. 255 p.

[6] Solozhentsev Eugene D. Top-economics. Economic safety management. – St. Petersburg: GUAP, 2015. 250 p.

[7] Raworth Kate. DOUGHNU ECONOMICS: Seven Ways to Think Like a 21st – Century Economist. – Publisher: Cornerstone. Publication Date: 06.04. 2017. 284 p.

[8] Solozhentsev E. D.. Risk Management Technologies with Logic and Probabilistic Models.-Dordrecht, Heidelberg, New York, London: Springer. *2012,* 328 p.

[9] Solozhentsev E. D. On the issue of digital management of state and economics. – Problems of risk analysis, Vol. 14, 2018, N 6, p. p. 39–43.

[10] Solozhentsev E. D. Top-economics: management of socio-economic safety / Special Issue: Management of Safety in Socio-Economic Systems, Int. J. of Risk Assessment and Management, Vol. 21, 2018, Nos. 1/2. p. p. 65–88.

[11] Karasev Vasily, Solozhentsev Eugene. Hybrid logical and probabilistic models for management of socio-economic safety / Int. J. of Risk Assessment and Management, Vol. 21, 2018, Nos. 1/2. p. p. 89–110.

[12]. Solozhentsev E. D. Logic and probabilistic risk models for management of innovations system of country / Int. J. of Risk Assessment and Management. Vol.18. 2015, Nos.¾. p. p.237–255

[13] Solozhentsev E. D., Mityagin S. A. Logic and probabilistic risk models for assessment and analysis of the drug addiction problem in a region / International J. of Risk Assessment and Management Vol. 18, 2015.No. 1, p. p. 1 – 20.

[14] Solozhentsev E. D. I3-Technologies for Economics. – SPb.: Nauka, 2011, 386 p.

[15] Solojentsev Evgueni D. *Scenario Logic and Probabilistic Management of risk in business and Engineering.* Second Edition. – Springer. 2009. 450 p.

[16] Robins S. P., Coulter M. Management, 6-th edition. – M.: Williams Publishing House, 2002. 880 p.

[17] EXPA. Certificate of state registration of computer programs, No. 2018612197. Expert System Expa. Date of issue: February 13, 2018. Authors: E. Solozentsev, V. Alekseev, E. Karaseva.

[18] ARBITR. (*PC ASM SZMA*, basic version 1.0) certified for use at the facilities of Rostekhnadsor of the Russian Federation for a period of 10 years. Certification passport number 222 of February 21, 2007.

[19] License N 2556 of 09/05/2018 for the implementation of educational activities issued by the Committee on Education by the Government of St. Petersburg to *SPIK SZMA* Joint-Stock Company (the course "Event-related digital management of safety and quality in economics" by E. Solozhentsev).

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 154-158

Artificial Intelligence in Workplaces and How It Will Affect Employment in Latvia

Ilmārs Šukjurovs
MBA
Author "Turība" University
Riga, Latvia
ilmars.sukjurovs@gmail.com

Rosita Zvirgzdiņa
Dr. oec. asoc. prof.
Supervisor "Turība" University
Riga, Latvia
rosita@turiba.lv

Silva Jeromanova-Maura
MBA
Co. author Turība" University
Riga, Latvia
silva.jeromanova-maura@silja.biz

***Abstract*—Already, artificial intelligence (AI) is all around us, from self-driving cars and drones to virtual assistants and software that translate or invest. Impressive progress has been made in AI in recent years, driven by exponential increases in computing power and by the availability of vast amounts of data, from software used to discover new drugs to algorithms used to predict our cultural interests. Digital fabrication technologies, meanwhile, are interacting with the biological world on a daily basis. Engineers, designers, and architects are combining computational design, additive manufacturing, materials engineering, and synthetic biology to pioneer a symbiosis between microorganisms, our bodies, the products we consume, and even the buildings we inhabit. Companies in the technology, media, and telecommunications industry expect AI to have a significant impact on product offerings in the next five years. This research contains the data from a pilot survey from Latvian business executives' expectations for AI and robotics for the next 5 years. The aim of this research is to find out if AI and robotics will make significant impact on workplaces in Latvia in the next five years.**

***Keywords*—artificial intelligence in business, economic theory, employment, Latvia.**

I. Introduction

Modern information technologies and the advent of machines powered by artificial intelligence (AI) have already strongly influenced the world of work in the 21st century. Computers, algorithms and software simplify everyday tasks, and it is impossible to imagine how most of our life could be managed without them. However, is it also impossible to imagine how most process steps could be managed without human force? The information economy characterised by an exponential growth replaces the mass production industry based on economy of scales [1]. The so-called fourth industrial revolution is about to kick in. And it will happen at an exponential pace. Klaus Schwab wrote an article in the World Economic Forum what the fourth industrial revolution is. The First Industrial Revolution used water and steam power to mechanize production. The Second used electric power to create mass production. The Third used electronics and information technology to automate production. Now the Fourth Industrial Revolution is building on the Third one, the digital revolution that has been occurring since the middle of the last century. It is characterized by a fusion of technologies that is blurring the lines between the physical, digital, and biological spheres. There are three reasons why today's transformations represent not merely a prolongation of the Third Industrial Revolution but rather the arrival of the Fourth and distinct one: velocity, scope, and systems impact. The speed of the current breakthroughs has no historical precedent. When compared with the previous industrial revolutions, the Fourth is evolving at an exponential rather than a linear pace. Moreover, it is disrupting almost every industry in every country. And the breadth and depth of these changes herald the transformation of entire systems of production, management, and governance [2].

Expectations for artificial intelligence (AI) are sky-high, but what are companies actually doing now? The gap between the ambition and execution is large at most companies. Three-quarters of executives believe AI will enable their companies to move into new businesses. Almost 85% believe AI will allow their companies to obtain or sustain a competitive advantage. But only about one in five companies has incorporated AI in some offerings or processes. Less than 39% of all companies have an AI strategy in place. The largest companies — those with at least 100,000 employees — are the most likely ones to have an AI strategy, but only a half has one [1].

Countries with high labour costs are more likely to implement AI into their daily tasks, but how about Latvia? There is a pilot survey being conducted right now to find out what business executives in Latvia think about AI in their companies, what their expectations are, how informed they are about the AI opportunities, and how they value the AI possibility to replace low qualified employees.

II. Materials and methods

Theoretical Framework

The name behind the idea of AI is John McCarthy, who began research on the given subject in 1955 and assumed that each aspect of learning and other domains

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4151
© 2019 Ilmārs Šukjurovs, Rosita Zvirgzdiņa, Silva Jeromanova-Maura.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

of intelligence can be described so precisely that they can be simulated by a machine [3]. Artificial Intelligence is a broad spectre computer science, and it could use some classifications and definitions to set apart its fields of science. Till this day there are no types of AI that would be multi-purpose programs. AI is a single purpose only – one that can play chess, cannot drive a car etc. However, Google DeepMind technologies are on its way to general purpose AI.

Artificial intelligence describes the work processes of machines that would require intelligence if performed by humans. The term 'artificial intelligence' thus means 'investigating intelligent problem-solving behaviour and creating intelligent computer systems' [4].

There are two kinds of artificial [5].:

1. Weak artificial intelligence: The computer is merely an instrument for investigating cognitive processes – the computer simulates intelligence.
2. Strong artificial intelligence: The processes in the computer are intellectual, self-learning processes. Computers can 'understand' by means of the right software/programming and are able to optimise their own behaviour on the basis of their former behaviour and their experience. (Shultz, 2016) This includes blockchain which is automatic networking with other servers or machines, and it leads to a huge scaling effect.

In general, the economic use of AI can be separated into six categories [5].:

1. Automation: What makes a system or process function automatically. For example, robotic process automation (RPA) can be programmed to perform high-volume, repeatable tasks that humans normally performed. RPA is different from IT automation in that it can adapt to changing circumstances.

2. Machine learning: The science of getting a computer to act without programming. Deep learning is a subset of machine learning that, in very simple terms, can be thought of as the automation of predictive analytics. There are three types of machine learning algorithms:

a. Supervised learning: Data sets are labelled so that patterns can be detected and used to label new data sets

b. Unsupervised learning: Data sets aren't labelled and are sorted according to similarities or differences

c. Reinforcement learning: Data sets aren't labelled but, after performing an action or several actions, the AI system is given feedback

3. Machine vision: The science of allowing computers to see. This technology captures and analyses visual information using a camera, analogue-to-digital conversion and digital signal processing. It is often compared to human eyesight, but machine vision isn't bound by biology and can be programmed to see through walls, for example. It is used in a range of applications from signature identification to medical image analysis. Computer vision, which is focused on machine-based image processing, is often conflated with machine vision.

4. Natural language processing (NLP): The processing of human -- and not computer -- language by a computer program. One of the older and best-known examples of NLP is spam detection, which looks at the subject line and the text of an email and decides if it's junk. Current approaches to NLP are based on machine learning. NLP tasks include text translation, sentiment analysis and speech recognition.

5. Robotics: A field of engineering focused on the design and manufacturing of robots. Robots are often used to perform tasks that are difficult for humans to perform or perform consistently. They are used in assembly lines for car production or by NASA to move large objects in space. Researchers are also using machine learning to build robots that can interact in social settings.

6. Self-driving cars: These use a combination of computer vision, image recognition and deep learning to build automated skill at piloting a vehicle while staying in a given lane and avoiding unexpected obstructions, such as pedestrians.

Dematerialisation. Thanks to automatic data recording and data processing, traditional 'backoffice' activities are no longer in demand. Autonomous software will collect necessary information and send it to the employee who needs it [4]. Additionally, dematerialisation leads to the phenomenon that traditional physical products are becoming software, for example, CDs or DVDs are being replaced by streaming services. The replacement of traditional event tickets, travel tickets or hard cash will be the next steps, due to the enhanced possibility of contactless payment by smartphone. And that is already happening, a large majority of Latvian people does not print event or plane tickets, even cash is used more rarely. Sweden even has a plan to be cash less country and now there is only 1% of Swedish citizens that use coins or money bills [6]. Gig economy. A rise in self-employment is typical for the new generation of employees. The gig economy is usually understood to include chiefly two forms of work: 'crowdworking' and 'work on-demand via apps' organised networking platforms [7]. There are more and more independent contractors for individual tasks that companies advertise on online platforms (eg, 'Amazon Mechanical Turk'). Traditional employment relationships are becoming less common. Many workers are performing different jobs for different clients.

These AI categories are most likely to reshape businesses as we know it in the next five years. There are a lot of workplaces that are under threat because of the rapid evolvment of AI. Robotics and AI could be a giant leap for the mankind and the old saying about janitor "Someone has to do that work too" will no longer apply to those lazy kids at school and low qualified workforce. People are meant to do more and be more. These monotonous and "dirty works" are the biggest time thieves in the 21st century. Of course, some people will loose their jobs

in the next years, but who will be the first winners by implementing AI and robotics?

The winners of the digital revolution are highly developed Asian countries with good education systems, such as Singapore, Hong Kong, Taiwan and South Korea [8]. These countries – together with the Scandinavian countries – have been undertaking research and working to find digital solutions for complex issues for a long time. The digital interconnection of people in these countries is also very far advanced. The share of the population at risk of unemployment is about six per cent in these countries. Finally, Western developed countries will profit from the relocation of the companies' production sectors when robotic production becomes cheaper than human production in low-labour-cost countries. This will create new jobs in these countries and destroy many routine jobs in the low-labour-cost countries. Another positive trend can be seen for India and China, which are both considered very suitable candidates for participation in the digital revolution due to the most of the population having a good command of English and IT skills. IT knowledge is taught in schools as a key qualification. It is, therefore, not surprising that Indian and Chinese professionals have more extensive computer knowledge than their French or English colleagues do [9].

Not only are salaries and wages lower in India, but also the number of better qualified professionals is, that is why, according to Forrester Research, 25,000 IT jobs are likely to be outsourced to India from the UK alone [10]. The countries like China, India are in the process of developing from simply being a low-labour-cost country into being a Western-orientated society whose population works mainly in the tertiary sector. As the most populated countries in the world, these two countries have a high level of consumer demand. Moreover, because of their rapidly growing cities, these developing countries need highly developed solutions in terms of logistics and environmental technologies, like the smart city, in order to increase the quality of life for city residents over the long term. The digital world market leaders are based in Silicon Valley, California. In 2015, the top ten Silicon Valley startups created an annual turnover of approximately US$ 600bn with information and communication services [11].

Additionally, the eight leading digital platforms – Alphabet, Amazon, Facebook, etc – due to their exponential growth show a significantly higher capital market value than the leading industrial companies (eg, General Electric, Siemens or Honeywell [12]. The rise of AI in the service sector, especially the gig-economy, can be illustrated by the example of Uber, which saw an increase in its market value from zero to US$ 40bn in only six years [11]. Even though more than 80 percent of the robots sold each year are deployed in Japan, South Korea, the US and Germany and enhance productivity in the production sector, the new business models in the service sector are the digital future. With economic growth in this sector, the US will be particularly resistant to future economic crises. It is therefore not surprising that innovative countries like Switzerland, Germany, the US or Japan are rated best in the Global Competitiveness Index by the World Economic Forum [13]. In summary, it can be said that the increase of automation and digitalisation is a global concern that, due to the lack of financial possibilities in many developing countries, will initially be strongly focused on Western developed countries and Southeast Asia. These countries will be considered the winners of Industry 4.0 because of their technological head start and their creative service models. Owing to the great number of emerging multidisciplinary support alternatives due to AI and machines, the requirements for future employees will change. There will be hardly any need for employees who do simple and/or repetitive work. Already today, the number of factory workers is constantly decreasing, and humans are ever more becoming the control mechanism of the machine. The automotive industry, where many production steps are already fully automated, is the pioneer in this respect. The lower the demand for workers, the higher the companies' demand for highly qualified employees will be. According to a common belief, better education helps [14].

Better education helps, however, only in certain circumstances. An additional qualification of an individual employee must be connected to the work in question. Additional qualifications as an accountant will be of little benefit for an individual employee, because – over time – there is a 98 per cent probability that the work of an accountant can be done by intelligent software [15].

Creative people who are talented in mathematics and sciences are best qualified for the new labour market. Although not every future employee will be required to be an IT programmer, should have a fundamental grasp of analytical and technical matters. Employees should be able to form a unit with supporting machines and algorithms and navigate the internet comfortably and move safely in social networks. To do this, it is necessary to know how the basic structures work. The employee should also, however, be able to examine machines and software critically. There is an increasing demand for employees who can work in strategic and complex areas as well. It is not necessary only to oversee machines, but also coordinate them. The interfaces between humans and machines and the overlaps in the area of responsibility among the more flexible humans must also be coordinated. There is thus likewise an increasing demand for future executive staff with social and interdisciplinary competence [16].

Employees must be able not only to communicate with other people, but also, if necessary, lead them effectively and coordinate them. In addition, creativity and flexibility are becoming increasingly important. In the future, critical and problem-orientated thinking will be expected of employees as the most important requirement. This requires sound judgment. The expectations with respect to availability will be higher for future employees. Flexible working hours and standby duties will be the rule and no longer an exception in the labour market. Employees will be required to focus not only on one main practice area, but also take on several multifaceted, sometimes highly complex tasks as necessary, and also perform as part of

a team. Employees are increasingly expected to have non-formal qualifications. These include, for example, the ability to act independently, build networks, organise themselves and their teams with a focus on targets, and think abstractly. Special knowledge or a flair for high-quality craftsmanship will become less important, since this work is likely to be done by intelligent software or a machine. Mere knowledge workers will no longer be required; the focus will rather be on how to find creative solutions to problems [17].

Deals will still be made between people in the future, even if the facts may be gathered beforehand by software [18].

Methods

Using data from author survey, where 51 of Latvian business executives were surveyed. Business executives shared they predictions and opinions about impact on their businesses. The aim of this research is to find out if AI and robotics will make significant impact on workplaces in Latvia in the next five years.

III. Results and discussion,

Preliminary results of the pilot survey show a little insight of attitude and opinions about AI and robotics in business among the Latvian business executives. One of the questions is – how business executives in Latvia predict the impact on daily business tasks. The impact must be measured on the Likert scale from 1 (no impact) till 5 (huge impact). (See Figure 1.)

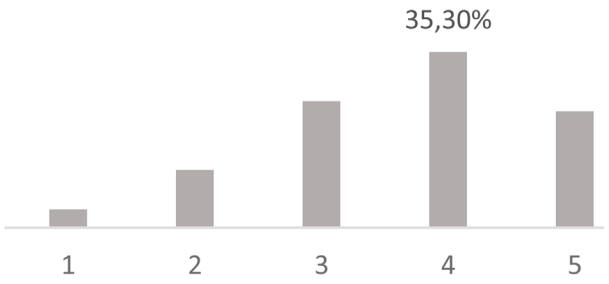


Figure 1. How you predict AI impact on your business in next 5 years?

More than a half of respondents notes that AI will have an impact on their businesses in the next five years, but a quarter of respondents notes that AI will have a moderate impact on their businesses in the next five years. Only approximately 16 percent of respondents note that AI will have no or little impact on their day to day business tasks.

One of the most interesting tasks in this ongoing survey is – how you rate your awareness about AI and robotics and its opportunities in business, on the scale from 1 (unaware) till 5 (fully aware). (See Figure 2.)

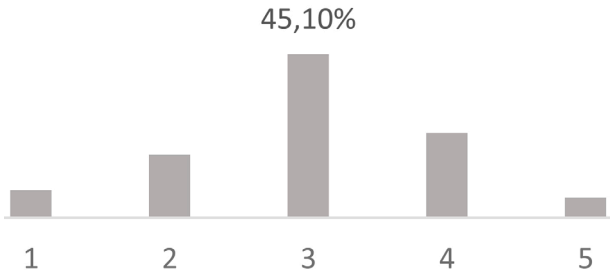


Figure 2. How aware you are about AI and robotics opportunities provided for your business?

It turns out that the majority of respondents are just partly aware about AI and robotics and their capabilities to implement their businesses. These results raise a serious question – if this survey reflects a real view on the next five years in AI impact on businesses since business executives are so moderately aware of AI and robotics opportunities provided.

As most technologically advanced countries are about to make changes in labour market thanks to AI and robotics, this matter should be next one to be rated among the Latvian business executives. Latvian business executives rate the possibility that some of workplaces could be replaced with AI and robots from 1 (low/ no possibility) to 5 (high possibility) (See Figure 3.)

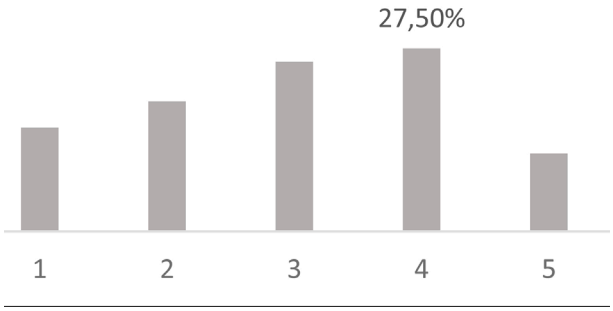


Figure 3. How you rate possibility to replace some workplaces in your business with AI or robots?

Approximately a third part of respondents believe that there is a big possibility that they will replace some human workers with AI or robots in the next five years.

There are many economic reasons businesses would like to adapt AI and robots in workplaces. But still there are people in businesses and always it`s all about people. Respondents were asked – how do you value human interaction in daily problem-solving/tasks? Rating 1 (not necessary) till 5 (Highly necessary). (See Figure 4.)

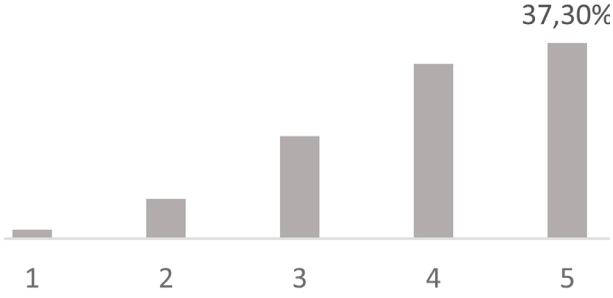


Figure 4. How do you value human interaction in daily problem-solving/tasks?

Human to human interaction is highly necessary as the largest majority of respondents noted. Humans will work for humans for a long time for sure, but AI and machines could serve more and more humans in the coming years. Humans make mistakes more often than machines and can do tasks with high precision. In this survey executives rated – how they value execution of tasks with no human errors. Rating from 1 (unimportant) till 5 (very important). (See Figure 5.)

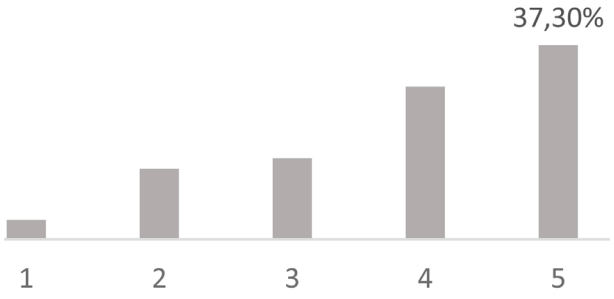


Figure 5. How do you value AI capability to do tasks without human errors?

The largest majority of respondents notes that a possibility to perform tasks without human errors is highly valued. Especially autonomous driving could save many human lives, because most of the fatal accidents are because of human errors, inattention and fatigue. Whereas, AI can operate 24/7 with no loss of performance quality.

Author recommends Latvian business executives to learn more about AI – its` applications for business. As survey shows – respondents are moderately aware of AI and robotics and its` capabilities. Author presumes that if Latvian business executives learn more about AI and robotics, that the overall predicted impact un business and use of AI and robotics in workplaces could be more optimistic.

IV. Conclusions

1. More than half of Latvian business executives predicts large and very large impact on their businesses in relation with AI and robotics progress. The indicators could be higher if Latvian business executives would be more aware of opportunities that AI and robotics offer.

2. Most of respondents worldwide see more opportunities than risks from AI and robotics.

3. Some legislation should be developed, approved regarding AI and robotics. The main reasons – who will take responsibility for mistakes and malfunctions (owner or manufacturer), in events where there are risks and damages done by AI, who would take the responsibility.

4. Author assumes that first automated jobs in Latvia will be for those positions, which not include working with people and tasks that requires precision with no human errors.

V. References

[1]. Boston Consulting Group. (2017). Reshaping Business with Artificial Intelligence. MIT Sloan Management Review.

[2]. Schwab, K. (2016, 1 14). The Fourth Industrial Revolution: what it means, how to respond. Retrieved from World Economic Forum: https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

[3]. Tava, J. M. (2016, 2 11). Computer cloud is dead. Retrieved from Der Spiegel: http://www.spiegel.de/netzwelt/web/john-mccarthy-der-vater-der-rechner-cloud-ist-tot-a-793795.html

[4]. Richard Lackes, Markus Siepermann. (2016, 2 11). Artificial Intelligence. Retrieved from Gabler: https://wirtschaftslexikon.gabler.de/definition/kuenstliche-intelligenz-ki-40285/version-184779

[5]. Rouse, M. (2018). AI (artificial intelligence). techtarget.com:https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence

[6]. Savage, M. (2017). Why Sweden is close to becoming a cashless economy. http://www.bbc.com/news/business-41095004

[7]. Stefano, V. D. (2016, 9 26). The rise of the «just-in-time workforce»: On-demand work, crowdwork and labour protection in the «gig-economy». International Labour Office.

[8]. Widmann, A. (2016, 2 1). Tricks are the new B.R.I.C. Retrieved from Sueddeutsche Zeitung: http://www.sueddeutsche.de/wirtschaft/schwellenlaender-ticks-sind-die-neuen-brics-1.2844010

[9]. Sikka, V. (2016, 2 18). Amplyfing Human Potential. Retrieved from Experience Infosys: http://www.experienceinfosys.com/humanpotential-infographic

[10]. Gatterer, H. (2013, 5 13). The reinvention of the working world. Retrieved from Zukunfts Institut: http://www.zukunftsinstitut.de/artikel/die-neuerfindung-der-arbeitswelt/

[11]. Kollmann, T. (2015). Digital Leadership. Bundestag, 18.

[12]. Roland Berger GmbH. (2016). Fair play in the digital arena. 90.

[13]. World Economic Forum. (2018, 2 2). Competitiveness Rankings. Retrieved from World Economic Forum: http://reports.weforum.org/global-competitiveness-report-2015-2016/competitiveness-rankings/

[14]. Shultz, S. (2016, 08 3). Future of labour market. Retrieved from Der Spiegel: http://www.spiegel.de/wirtschaft/soziales/arbeitsmarkt-der-zukunft-die-jobfresser-kommen-a-1105032.html

[15]. Klinge, & Schmitd. (2016). Kolege roboter. Focus Magazin, 66.

[16]. Bochum. (2017). 'Gewerkschaftliche Positionen in Bezug auf "Industrie 4.0"'. Zukunft der Arbeit in Industrie 4.0, 36.

[17]. Zukunfts Institut. (2012). The reinvention of the working world. Retrieved from Zukunfts Institut: http://www.zukunftsinstitut.de/artikel/die-neuerfindung-der-arbeitswelt/

[18]. Anderson, J. Q. (2015). The Future of Work? The Robot Takeover is Already Here! Medium.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 159-164

Application of Computer Vision Technologies for Autonomous Pile Manipulation

Artjoms Suponenkovs
Department of Computer Control Systems
Riga Technical University
Riga, Latvia
Artjoms.Suponenkovs@rtu.lv

AMihails Kovalovs
Department of Image Processing and Computer Graphics
Riga Technical University
Riga, Latvia
mihails.kovalovs@rtu.lv

Zigurds Markovics
Department of Computer Control Systems
Riga Technical University
Riga,Latvia
Zigurds.Markovics@rtu.lv

***Abstract*—Modern robots can perform uncreative monotonous tasks. One of such tasks is pile manipulation. Computer vision technologies can help robots acquire additional information by analyzing a pile of complex objects. One of such complex objects is a fish. The presented work investigates the problems of complex object analysis using computer vision. This paper addresses the challenges of image pre-processing, image segmentation, fish detection and occlusion detection. This work results can be useful for developing a computer vision system for pile manipulation.**

***Keywords*—computer vision, image pre - processing, fish analysis, pile manipulation, occlusion boundaries, image segmentation.**

I. Introduction

Pile manipulation is an exhausting uncreative task. Autonomous industrial robots can do such monotonous tasks. Therefore, the autonomous pile manipulation problem is very relevant. To solve this complicated task, it is reasonable to use computer vision approaches. These approaches allow to analyze visual information on pile of objects.

The autonomous robot can perform various manipulations with objects, depending on the result of visual information analysis [1]. Visual information analysis helps robots to find the necessary object in a pile. In our case, the necessary object is a fish that can be picked up by robot.

A fish can be damaged during displacement and handling. To avoid damaging the fish, robots should pick up only that fish which is not overlapped by any other fish "Fig.1". Therefore, it is important to find the necessary fish without occlusion. To solve this problem, it is possible to use different technologies and approaches: photometric stereo [2], 3D depth sensor [3], the optical flow [4], stereo camera [5], single image [6]. All of these technologies and approaches have advantages and disadvantages. The photometric stereo, for example, is appropriate for tasks that involve simple and big objects. The 3D depth sensor is relatively expensive. It is necessary to analyze a lot of images to realize the optical flow approach. Therefore, in presented work, the information source is a single image. This approach is cheap and fast. For this approach it is not obligatory to have any additional expensive equipment. Also the amount of information is relatively small.

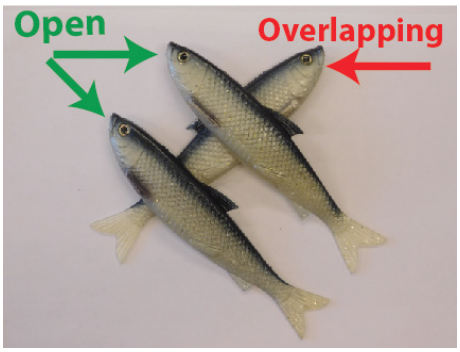


Fig. 1. Open and overlapping fish

II. Materials and Methods

The goal of this publication is to find non overlapping fish. To solve this complicated task, it is possible to use the approach that consists of 5 parts "Fig.2":

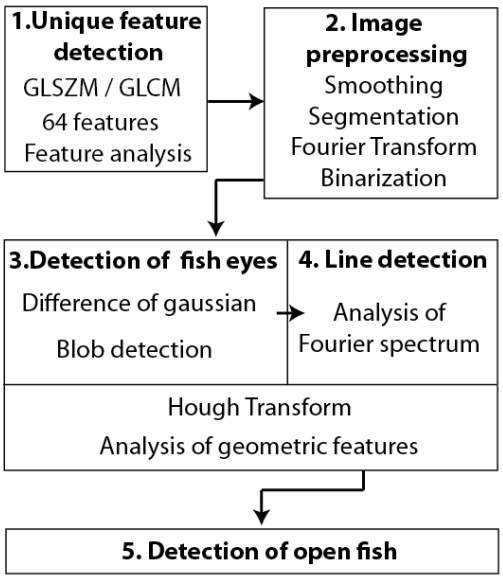


Fig. 2. Fish analysis system flowchart

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4033
© 2019 Artjoms Suponenkovs, Mihails Kovalovs, Zigurds Markovics.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

A. Unique feature detection

To find open fishes, it is useful to find where fish overlap and fish eye features. It is possible to analyze textural fish features. This analysis can be performed using statistical characteristics, spatial frequencies, structural elements. Gray level matrices allow to calculate statistical characteristics by gray texture data. These statistical characteristics can help to detect fish eye and occlusion. The goal of the statistical characteristic analysis is to find these unique features of fish eye and overlapping.

In this work, we used two types of gray level matrices: Gray Level Co-Occurrence Matrix (GLCM)[7] and Gray Level Size Zone Matrix (GLSZM)[8]. These matrices are used to calculate important statistical features: variance, contrast (1), energy (2), homogeneity(3), correlation, dissimilarity, zone emphasis and etc.

$$CON = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \quad (1)$$

$$\text{Energy} = \sum_{i,j=0}^{N-1} P_{i,j}{}^2 \quad (2)$$

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+|i-j|} \quad (3)$$

where:

P – is the probability of combined neighboring elements (GLCM value);
i and j – are GLCM indexes;
N – size of GLCM matrix.

It is possible to design a computer vision system for the analysis of open and overlapping fish based on detected features in the first experiment (looking below).

B. Image preprocessing (image preparation)

At this stage, it is important to make an input image preparation for further analysis. The image preparation depends on the type of further analysis. At next stages, it is possible to use different types of analysis: analysis of geometric features, analysis of Fourier spectrum and analysis of textural features. Therefore, it reasonable to use the Fourier transform together with the Fourier spectrum analysis, segmentation together with the analysis of geometric features, difference of Gaussian together with analysis of textural features.

Usually the input image contains noise. This noise makes it difficult to analyze the fish. At this stage, it is very important to reduce the noise level by using special smoothing algorithms. It is reasonable to use Perona-Malik filtering [9]-[10], Gaussian filtering or Mean Shift filtering to reduce noise. Perona-Malik method has a useful feature that saves strong edges of objects in the image. The "Fig. 3" shows the results of Perona - Malalik filtering, Mean Shift filtering and K-Means clustering [11]-[12]. Perona - Malalik and Mean Shift filtering "Fig. 3 B and C" remove high frequency information, which in turn is helpful for the segmentation of the input image.

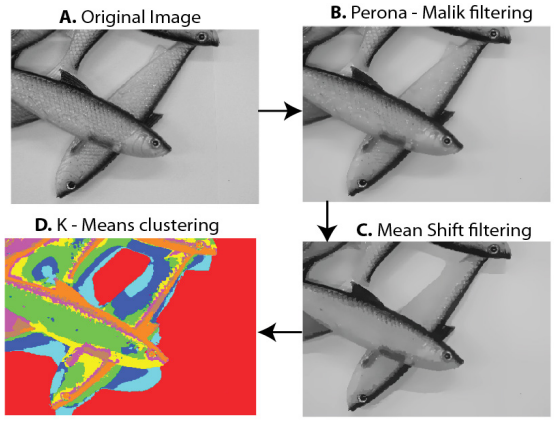


Fig. 3. Input image smoothing and clustering

The segmentation of input image is important for the fish analysis. The input image can be divided into many pixel regions by performing image segmentation. After that it is possible to analyze each region by using geometrical features.

The "Fig. 4" shows the steps of segmentation. The first step "Fig. 4 B" is K-Means clustering. K-Means clustering divides an image into 8 clusters. This division is based on pixel intensity. The second step "Fig. 4 C" is binarization of regions of interest (ROI). This binarization is based on cluster features. The third step "Fig. 4 D" is segmentation of separate regions.

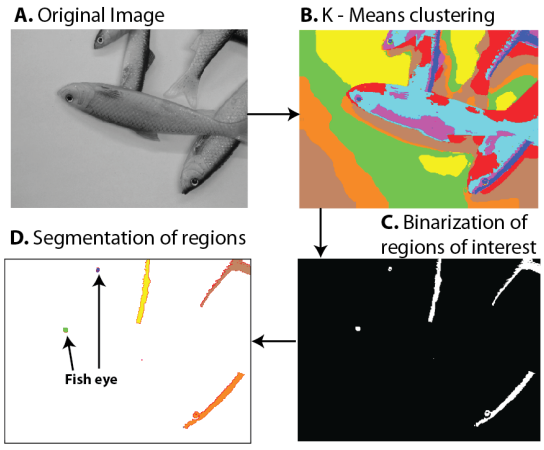


Fig. 4. Steps of segmentation

C. Detection of fish eyes

The eye is the most easily perceptible part of a fish. Therefore it is reasonable to detect the fish eye at first. Detection of the fish eye is based on a unique feature detection experiment (see below). There are many unique features. For example variance, contrast, dissimilarity. By using these features we can analyze the image texture and check if it contains the fish eye.

In this work we are trying to use many methods of fish eye detection: the Difference of Gaussian (DoG), blob detection [13], analysis of geometric feature, Hough circle detection. All these methods have advantages and disadvantages.

Blob detection is based on the Laplacian of Gaussian

(LoG) [14]. Our experiment (see below) shows that a fish eye has a high Laplacian response. That would be because the fish eye is an ideal blob and the fish eye looks like as Laplacian filter. It is possible to simplify Laplacian calculation by Difference of Gaussian (DoG). The DoG has approximately the same result as scale-normalized Laplacian. Scale-normalized Laplacian and DoG can be calculated as follows (4),(5),(6):

a) Gaussian:

$$G(x,y,\sigma)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{(-\frac{x^2+y^2}{2*\sigma^2})} \quad (4)$$

b) Difference of Gaussian:

$$DoG=G(x,y,k\sigma)-G(x,y,\sigma) \quad (5)$$

c) Scale-normalized Laplacian:

$$LoG=(G_{xx}(x,y,k\sigma)-G_{yy}(x,y,k\sigma)) \quad (6)$$

where

x, y are pixel coordinates;

σ - deviation;

k - coefficient;

G_{xx}, G_{yy} - second partial derivatives.

It is possible to analyze segment shape by its geometric features. The fish eye is oval or a circle. Therefore, we can use the circularity. The circularity of a circle is (4 * π), hexagon - 13.86, square - 16 and of equilateral triangle - 20.79. The circularity can be calculated as follows (7):

$$C=\frac{P^2}{S} \quad (7)$$

where:

P - is segment perimeter,

S - is segment area.

After calculation of circularity of a segment, we can make comparison of the calculated circularity and known circularity of ideal shapes (circle, hexagon, square, circle). The other important geometric feature is elongation (8).

$$Elong=\frac{m_{20}+m_{02}+\sqrt{(m_{20}-m_{02})^2+4*m_{11}^2}}{m_{20}+m_{02}-\sqrt{(m_{20}-m_{02})^2+4*m_{11}^2}} \quad (8)$$

where:

m_{20} - is two-dimensional central moment (2 and 0 is moment index);

m_{jk} - is central moment (j and k is moment index).

These geometric features are useful for circle detection.

D. Line detection

The Line detection can be performed using Hough transformation [15]-[16] that generates a list of lines. Long lines are more important than short. If a fish has relatively long line, then this fish is probably not overlapped by other fish "Fig. 5". The "Fig. 5" shows the relation between length of line and fish overlapping. This relation points to the importance of long lines.

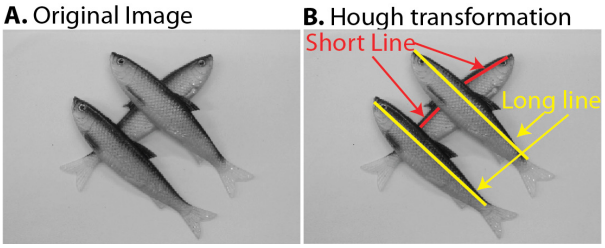


Fig. 5. Line detection

Additional information about fish orientation can be helpful for further analysis. This information about orientation allows us to remove unnecessary fish lines and help detect an open fish "Fig. 7". The fish orientation can be determined by spectral analysis. It is for this reason that the original image must be converted into spectral image by performing Fourier transform. The "Fig. 6" shows the determination of fish orientation. It is possible to estimate the spectral image "Fig. 6 B" by using the following equation(9):

$$F(\theta)=\sum_{r=1}^{R_{max}}F_r(\theta) \quad (9)$$

where

$F_r(\theta)$ - is a function (polar coordinate system) that returns spectral image pixel intensity;

$F(\theta)$ - is function that returns a sum of pixel intensities;

θ - a corner;

r - a radius;

R_{max} - the maximum radius.

This equation(9) allows us to find the corner θ_{max} which matches the maximum sum of pixel intensities ($F_{max}(\theta_{max})$ = maximum). It is possible to use a threshold for the detection of the maximum sum of pixel intensities. In that case, we can find many important corners (θ_{max1}, θ_{max2}). For example, the spectral image has two important corners (in the "Fig. 6 B and C"). The angles of these two corners must be increased by 90 degrees ($\varphi=\theta_{max}+90^{\circ}$). This creates two new corners (φ_1, φ_2) that indicate the fish orientation (see "Fig. 6 A and D").

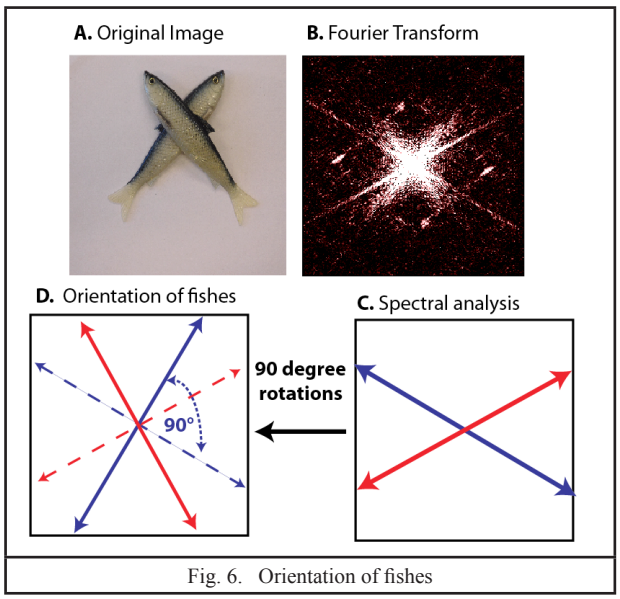


Fig. 6. Orientation of fishes

E. Detection of an open fish

The detection of an open fish has 3 steps "Fig. 7":

1) detection of the fish eye;

2) finding long lines around the eye;

3) checking the orientation of the line.

As a result of this detection there are fish that are probably not overlapped by other fish.

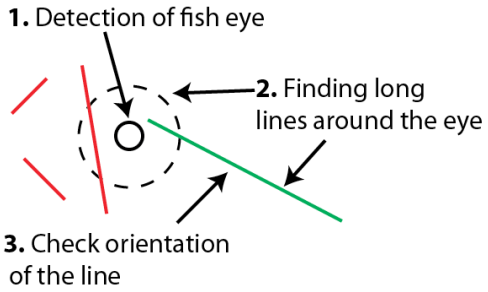


Fig. 7. Detection of an open fish

III. Results and Discussion

This part contains the results of three experiments:

1. Unique feature detection;
2. Detection of fish eyes;
3. Detection of fish lines.

A. Experiment - Unique feature detection

This experiment investigates the problem of unique feature detection. Unique features were detected using the GLCM and GLZM matrices. These matrices are used to analyze textural information. Sixty-four textural features were taken into account in this experiment. For this purpose, one GLZM matrix and four GLCM matrices (horizontal, vertical and two diagonal) were used. The GLZM matrix describes 16 statistical features and one GLCM matrix describes 12 statistical features. As a result there are 64 features (16 + 12 * 4).

TABLE I. Feature comparison between fish overlapping and background texture

Feature number/s	Feature/s	is greater for	Confidence
GLCM(20, 21, 44, 45, 8, 9, 32, 33)	Mean X and Mean Y	Background	0.903
GLZM (51 - LGZE)	Low Gray level Zone Emphasis	Fish overlapping	0.901
GLZM(60 - BARYGL)	The barycenter on gray level	Background	0.895
GLZM(52 - HGZE)	High Gray level Zone Emphasis	Background	0.883
GLCM(22, 23, 46, 47, 34, 35, 10, 11)	Variance X and Variance Y	Fish overlapping	0.839

The table 1 shows the important features. This table results are based on 342 comparisons between overlapping fish and background texture. As a result of the comparison there are some important features: mean, low gray level zone emphasis, barycenter, high gray level zone emphasis and variance. These features are helpful in the detection of overlapping fish. The table also contains the confidence level that shows the importance of a given feature.

The table 2 shows the important features of fish eye texture. This table results are based on 350 comparison between fish eye and background texture. As a result of comparison there are some important features: variance, correlation, contrast, large zone high gray level emphasis and dissimilarity.

The detection of fish eyes and the detection of fish lines that are described above, are based on the results of this experiment.

TABLE II. Feature comparison between fish eye and the background

Feature number/s	Feature/s	is greater for	Confidence
GLCM (22, 23, 46, 47, 34, 35, 10, 11)	Variance X and Variance Y	Fish eye	0.992
GLCM (36, 24, 12, 48)	Correlation	Background	0.991
GLCM (37, 1, 25, 13)	Contrast	Fish eye	0.989
GLZM (LZ-HGE - 56)	Large Zone High Gray level Emphasis	Background	0.986
GLCM (26, 2, 38, 13, 14)	Dissimilarity	Fish eye	0.984

Experiment - Detection of fish eyes

This experiment investigates the problem of the detection of fish eyes. The detection of fish eyes was performed using 5 methods:

1) BLOB - the blob detection;
2) DOG - method is based on the Difference of Gaussian;
3) DOG + BLOB - combination of 1st and 2nd methods;
4) SEGMENT - method is based on segmentation and geometric feature analysis.
5) HOUGH - the Hough transformation (circle detection).

The "Fig. 8" and "Fig. 9" show comparison between results of these methods. There are 6 bars in "Fig. 8". The "Original" bar shows the number of fish eyes in original image. As shown in "Fig. 8" there are two methods that have good results: DOG and DOG + BLOB. The DOG method had detected 86 percent of fish eyes, but DOG + BLOB had detected 90 percent of fish eyes.

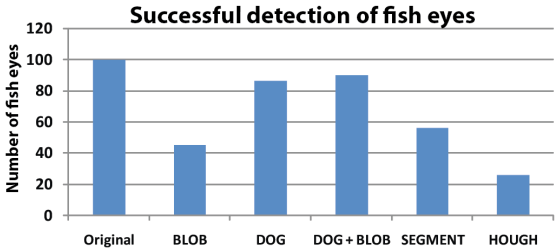


Fig. 8. Number of successful detected eyes

The "Fig. 9" shows the number of mistakes of each method. As shown in "Fig. 9" BLOB has the least number of mistakes. Then that means BLOB is very stable method. Therefore it is possible to make a combination of BLOB and other methods (for example: BLOB + DOG).

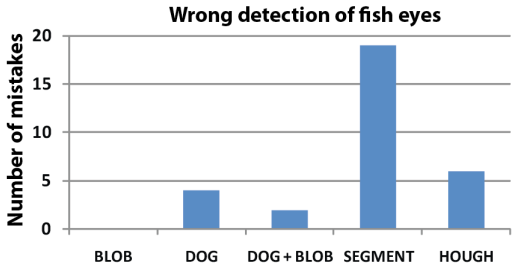


Fig. 9. Number of incorrectly detected eyes

Experiment - Detection of fish lines

This experiment investigates the problem of the detection of fish long lines. The detection of fish lines was performed using 2 methods:

1) HOUGH - the Hough transformation (line detection);
2) SMART HOUGH - the Hough transformation and texture analysis.

The "Fig. 10" shows comparison between the results of these methods. As shown in "Fig. 10" SMART HOUGH has the lowest number of mistakes because of texture analysis that removes unnecessary lines.

Fig. 10. Number of wrong and successful detected fish lines

IV. Conclusions

The results of the first experiment show that textures of overlapping fish and fish eye have unique statistical features. The unique statistical features of overlapping fish are mean value, low gray level zone emphasis, the barycenter on gray level, etc. The unique statistical features of fish eye are variance, correlation, contrast, etc. It is possible to design computer vision system for the analysis of open and overlapped fish based on detected unique features.

The results of the second experiment show that DOG + BLOB had detected 90 percent of fish eyes. The comparison of fish eye detection methods show that BLOB method has the lowest number of mistakes. Therefore it is possible to make a combination of BLOB and other methods. The combination of BLOB and DOG methods has the best results.

The results of the third experiment show that SMART HOUGH had detected approximately 78 percent of fish lines. SMART HOUGH has the lowest number of mistakes because of texture analysis that removes unnecessary lines.

This work results can be useful for developing a computer vision system for pile manipulation.

Acknowledgments

The authors wish to acknowledge support from PERUZA company of Latvia "Fig. 11".

PERUZA

Fig. 11. Logotype of PERUZA

References

[1] R.Mojtahedzaden, Safe Robotic Manipulation to Extract Objects from Piles: From 3D Perception to Object Selection, Orebro University, Repro, 2016.

[2] Ikeuchi K., B. K.P. Horn, S. Nagata, T. Callahan and O. Feirigold, "Picking up an object from a pile of objects". Massachusetts Institute of Technology, 1983, A.I. Memo, 726: 1-26.

[3] Domae Y., Okuda H., Taguchi Y., Sumi K. and Hirai T., "Fast Graspability Evaluation on Single Depth Maps for Bin Picking with General Grippers", IEEE International Conference on Robotics and Automation, 2014.

[4] Patrik Sundberg, Thomas Brox, Michael Maire, Pablo Arbelaez and Jitendra Malik, "Occlusion Boundary Detection and Figure/Ground Assignment from Optical Flow", IEEE, 2011.

[5] Wang X., H. Xiong and L. Li, "Stereo Vision Based Piled Object Recognition System for Robot", Ricoh Software Research Center Co., Ltd., 41: 31-36, 2016.

[6] Derek Hoiem, Andrew N. Stein, Alexei A. Efros and Martial Hebert, "Recovering Occlusion Boundaries from a Single Image ", Int J Comput Vis 91: 328–346, 2007.

[7] Fritz Albregtsen, "Statistical Texture Measures Computed from Gray Level Coocurrence Matrices", University of Oslo, 2008.

[8] Guillaume Thibault, Jesus Angulo, Fernand Meyer, "Advanced statistical matrices for texture characterization: application to dna chromatin and microtubule network classification", 18th IEEE International Conference on Image Processing, 2011.

[9] Joachim Weickert, "Anisotropic Diffusion in Image Processing", Department of Computer Science University of Copenhagen, Copenhagen, Denmark, 2008.

[10] Maciej Wielgus, "Perona-Malik equation and its numerical properties", Lipiec, 2010.

[11] Aristidis Likas, Nikos Vlassis, Jakob J. Verbeek, "The global k-means clustering algorithm", Elsevier Pattern Recognition 36 (2003) 451 – 461.

[12] Hans-Hermann Bock, "Origins and extensions of the k-means algorithm in cluster analysis", Electronic Journal for History of Probability and Statistics, Vol 4, December 2008.

[13] Anlong Ming, Huadong Ma, "A blob detector in color images", CIVR'07, 2007, Amsterdam, The Netherlands.

[14] Hui Kong, Hatice Cinar Akakin, and Sanjay E. Sarma, "A Generalized Laplacian of Gaussian Filter for Blob Detection", IEEE Transactions on cybernetics, vol. 43, no. 6, December 2013.

[15] Varun R., Yadunandan Vivekanand Kini, K. Manikantana and S. Ramachandranb, "Face Recognition using Hough Transform based Feature Extraction", ICICT, Procedia Computer Science 46, 1491 – 1500, 2015.

[16] Antoine Manzanera, Thanh Nguyen, Xiaolei Xu, "Line and circle detection using dense one-to-one Hough transforms on greyscale images", EURASIP Journal on Image and Video Processing, Springer, 2016, 34, pp.1773 - 1773.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 165-169

Probabilistic Inference for Interval Probabilities in Decision-Making Processes

Oleg Uzhga-Rebrov
Information and Communication Technologies Research Centre
Rezekne Academy of Technologies
Rezekne, Latvia
rebrovs@tvnet.lv

Galina Kuleshova
Institute of Information Technology
Riga Technical University
Riga, Latvia
galina.kulesova@rtu.lv

***Abstract*— The present paper considers one approach to Bayes' formula based probabilistic inference under interval values of relevant probabilities; the necessity of it is caused by the impossibility to obtain reliable deterministic values of the required probabilistic evaluations. The paper shows that the approach proves to be the best from the viewpoint of the required amount of calculations and visual representation of the results. The execution of the algorithm of probabilistic inference is illustrated using a classical task of decision making related to oil mining. For visualisation purposes, the state of initial and target information is modelled using probability trees.**

Keywords*— Bayes' formula, event indicators, interval probabilistic inference, interval probability, probability tree.

I. Non-formal Introduction to Interval Probabilities

Probabilistic inference procedures are widely used in different scientific, technical and economic areas. To exemplify widespread procedures of this kind, one can mention marginalization of a set of joint probabilities, calculation of conditional probabilities in the set of joint probabilities, and calculation of the posterior probabilities of events on the basis of their prior probabilities and information provided by event indicators.

All original procedures of probabilistic inference have been developed for the cases when the values of relevant probabilities are set in a unique deterministic form. Unfortunately, it is not always possible to obtain such probability values. The main reasons for that are the shortage or complete absence of suitable statistical data and low confidence in the evaluations provided by experts.

What can be done in situations like that? One possible way is to use available deterministic evaluations ignoring their potential unreliability. Another way is to introduce some extent of uncertainty into relevant evaluations assuming that their real values lie within the specified bounds of uncertainty. In situations where for some reason or other it is impossible to obtain reliable evaluations of the necessary probabilities, the second option seems to be preferable. By introducing controlled extents of uncertainty in the evaluations that are of interest to us, we extend the possibility of obtaining uncertain results, but real value of that result lies within the set bounds of uncertainty.

The idea to apply non-point probabilities has a long history. The first formal use dated back at least to the middle of the 19th century, is connected with the name of George Boole who intended to co-ordinate theory of logic (that can express complete ignorance) and probability theory.

Since the 1990s, the theory has received strong impulse initiated by exhaustive foundations of P. Walley [1] who had introduced the term *imprecise probabilities*. To evaluate boundary values of probabilities, P.Walley [1] introduced into consideration both buying and selling price for a hypothetical gamble. Those two prices correspond to the lower and the upper probabilities that form an interval of possible values of relevant probability. That interpretation underlies the theory of uncertain probabilities of P. Walley. The theory can be regarded as a specific extension of the traditional subjective probability theory.

Walley's theory extends the traditional subjective probability theory via buying and selling prices for gambles, whereas Weichselberger's approach generalizes Kolmogorov's axioms without imposing an interpretation. On the other hand, Weichselberger [2]-[4] treats interval values of probabilities as initial data; based on it, he builds his theory of interval probabilities. Strict theory of interval probabilities is also described in [5].

II. Formal Concepts and Definitions of Interval Probabilities

Let there be a set of random events $A = \{a_i, i = 1,...,n\}$. Let us assume that probabilities of occurrence of those events are set not in the deterministic form but in the form of intervals of possible values of those probabilities.

$$[l_i, u_i], \ i = 1,...,n,$$

where l_i - lower (the least) possible value of

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4112
© 2019 Oleg Uzhga-Rebrov, Galina Kuleshova
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

probability $p_i = p(a_i)$; u_i - upper (the largest) possible value of probability $p_i = p(a_i)$.

It is evident that while choosing one of the values $p_i \in [l_i, u_i]$, $i = 1,...,n$ in a random or systematized way, we get a set of deterministic probabilistic evaluations $\{p_i, i = 1,...,n\}$. In [6], a set of all possible probabilities of this kind is determined as follows:

$$P = \{p_i \in p(A) / l_i \le p_i \le u_i, \forall i\}, \quad (1)$$

where $p(A)$ denotes a set of all possible probabilistic evaluations defined in the set of random events *A*.

To avoid situations when $P = \varnothing$, boundary values of probabilistic intervals have to satisfy these restricting conditions:

$$\sum_{i=1}^{n} l_i \le 1 \le \sum_{i=1}^{n} u_i . \quad (2)$$

Probabilistic intervals satisfying conditions (2) are called *proper* intervals in [6]. It is evident that in the tasks of interval probabilistic inference we should only use proper intervals.

In common case $p_i \ge l_i$ and $p_i \le u_i$, $\forall a_i \in A$. If

$$l_i = \inf_{p_i \in \mathbf{P}} \text{ and } u_i = \sup_{p_i \in \mathbf{P}}, \ \forall i, \quad (3)$$

it means that deterministic values of probabilities can be selected over the whole interval $[l_i, u_i]$, also including its boundaries. In [5], probabilistic intervals satisfying (3) are called *reachable intervals*; whereas in [2]-[4], probabilities defined in the probabilistic intervals of general type are called R*-probabilities* but probabilities defined in the reachable intervals are called F*-probabilities with M structure.*

In [6], it is proved that for reachable probabilistic intervals these inequalities are valid:

$$\sum_{j \ne i} l_j + u_i \le 1, \ \forall i; \quad (4)$$

$$\sum_{j \ne i} u_j + l_i \le 1, \ \forall i. \quad (5)$$

In [2]-[4], [6], algorithms are proposed for determining marginal interval probabilities and conditional interval probabilities in the set of joint interval probabilities. It should be noted that the algorithms described in [6] are simpler and more operable than those presented in [2]-[4].

III. Interval Versions of Bayes' Formula

One specific interval version of Bayes' formula is proposed in [2]-[4]. Here, formula derivation is based on the simultaneous use of two concepts of interval probability: intuitive concept and canonical concept. The algorithm is quite complicated, is of rather artificial character and possesses limited operability; that is why, this version is not examined in the present paper.

A much more attractive version is an interval variant of Bayes' formula, which is based on the concept of generalised probabilistic intervals [7]-[8]. The presentation of the theory of generalised intervals can be found in [9]-[11], as well as in other works.

A classical interval is identified as a set of real numbers, while a generalized one is identified by means of predicates that are filled with real numbers, and its boundaries are not ordered in a conventional sense. The generalized interval $x = [\underline{x}, \overline{x}] \in KR$ is called proper if $\underline{x} \le \overline{x}$, and improper if $\underline{x} \ge \overline{x}$. The set of proper intervals is denoted as $IR = \{[\underline{x}, \overline{x}] / \underline{x} \le \overline{x}\}$, but the set of improper intervals is denoted as $\overline{IR} = \{[\underline{x}, \overline{x}], \underline{x} \ge \overline{x}\}$. Operations on the generalised intervals are determined based on Kaucher's arithmetic. [12].

Two specific mathematical operations are defined in the set of generalized intervals:

$$pro\mathbf{x} = [\min(\underline{x}, \overline{x}), \max(\underline{x}, \overline{x})]. \quad (6)$$

The result of that operation is a proper generalised interval.

$$imp\mathbf{x} = [\max(\underline{x}, \overline{x}), \min(\underline{x}, \overline{x})]. \quad (7)$$

This operation yields an improper generalised interval. The operation that follows transforms a proper generalised interval into an improper generalised interval.

$$dual\mathbf{x} = [\overline{x}, \underline{x}]. \quad (8)$$

Wang [7], [8] proposes this interval version of Bayes' formula:

$$p(E_i / A) = \frac{p(A / E_i)\, p(E_i)}{\sum_{j=1}^{n} dualp(A / E_j)\, dualp(E_j)} \quad (9)$$

where E_i, $i = 1,...,n$, - are mutually disjoint event partitions in Ω, and $\sum_{j=1}^{n} p(E_j) = 1$; $dualp(.)$ is defined in (8).

To simplify the calculations, this expression can be used:

$$\sum_{j=1}^{n} dualp(A / E_j)\, dualp(E_j) = \sum_{j=1}^{n} dual\left(p(A / E_j)\, p(E_j)\right) \quad (10)$$

Let us consider a simple illustrative example of probabilistic inference in a task of decision making based on the interval version of Bayes' formula (9). As an example, a classical task of assessing the chances of oil presence on a specific site is described, provided that the prior evaluations of these chances are set, and conditional deterministic evaluations of probabilities of the results of seismic exploration of the site are assigned. The following data are used as initial: a set of random events ("states of nature") $A = \{a_1, a_2\}$ where event a_1 corresponds to real

presence of oil on the site, event a_2 corresponds to real absence of oil on the site. Let us call events a_1 and a_2 *geological events*. Let us assume that based on the expert evaluation, these interval values of probabilities of the events were assigned:

$$p(a_1) = [0.50, 0.70], \qquad p(a_2) = [0.30, 0.50].$$

Let us assume that a manager of an oil mining company has made a decision to arrange seismic exploration of the site to re-evaluate the prior values of probabilities $p(a_1)$ and $p(a_2)$. Let us denote a set of random events, outcomes of seismic exploration as $B = \{b_1, b_2\}$ where b_1 is an outcome indicating the presence of oil on the site but b_2 is an outcome indicating the absence of oil on the site. Let us call events b_1 and b_2 *seismic events*.

The specifics of a seismic exploration is that it can both confirm real presence or absence of oil on a site, and produce erroneous results, i.e., to show the presence of oil when it is missing in reality or to show the absence of oil when it is really present. Let us introduce this system of denotations:

b_1 / a_1 - seismic exploration has confirmed real presence of oil on the site;

b_2 / a_1 - seismic exploration has erroneously indicated the lack of oil on the site, though in reality oil is present;

b_1 / a_2 - seismic exploration has erroneously indicated the presence of oil on the site, though in reality oil is not present on the site;

b_2 / a_2 - seismic exploration has confirmed real absence of oil on the site.

Let there be set these interval values of conditional probabilities:

$$p(b_1 / a_1) = [0.70, 0.90], \; p(b_1 / a_2) = [0.10, 0.30]$$

$$p(b_2 / a_1) = [0.10, 0.30], \; p(b_2 / a_2) = 0.70, 0.90.$$

As can easily be seen, these values are reachable values according to conditions (4) and (5). Actually, here we are not interested in conditional probabilities of the results of seismic study depending on the presence or lack of oil at a site, $p(b_j / a_i)$, $i, j = 1, 2$; instead, we are interested in conditional probabilities $p(a_i / b_j)$, $i, j = 1, 2$ of the presence or lack of oil on a specific site depending on the results of a seismic exploration.

The task in this example is to calculate the posterior probabilities $p(a_i / b_j)$, $i, j = 1, 2$, based on the information available. The initial state of information in the form of a probability tree is shown in Fig. 1.

Probabilities of the outcomes are calculated by multiplying the probabilities related to the tree branches leading to the given outcome. The calculated values of these probabilities are given in the end positions of probability tree in Fig. 1. Now all the necessary and sufficient information is available to calculate the required values of the posterior probabilities.

Event b_1 can occur jointly with event a_1 (outcome (1)), and jointly with event a_2 (outcome (3)). Therefore, the total probability of event b_1 can be calculated as follows:

$$p'(b_1) = p(1) + p(3) = [0.35, 0.63] + [0.03, 0.15] = [0.38, 0.78]$$

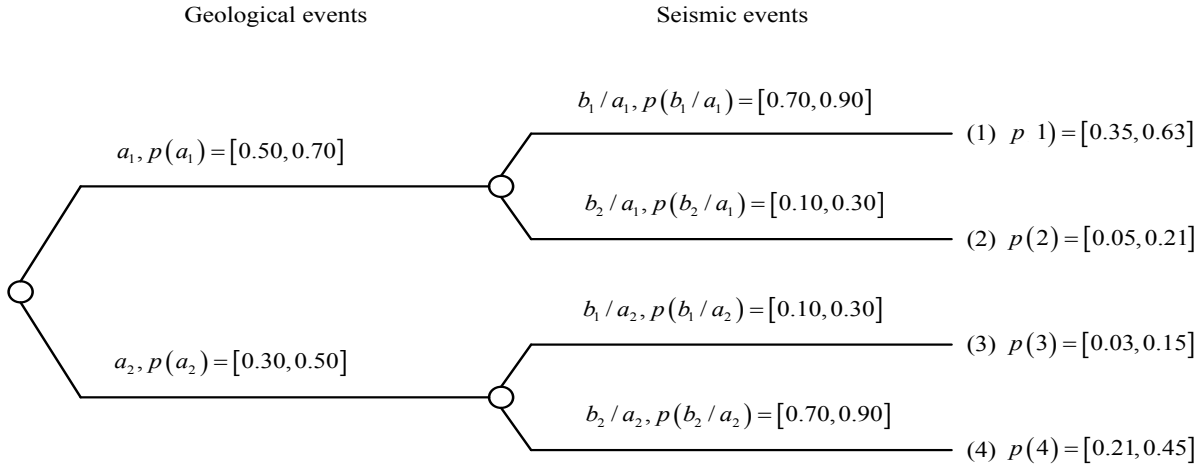


Fig. 1. Initial state of information in the task of oil mining.

Event b_2 can occur jointly with event a_1 (outcome (2)) and with event a_2 (outcome (4)). That is why,

$$p'(b_2) = p(2) + p(4) = [0.05, 0.21] + [0.21, 0.45] = [0.26, 0.66]$$

The values $p'(b_1)$ and $p'(b_2)$ would be the values of denominators in the classic Bayes' formula should all the values in the above example be set in unambiguous deterministic form. However, in reality we deal with interval values of probabilities; due to that, we have to utilize values $dualp'(b_1)$ and $dualp'(b_2)$. Let us determine these values using (10).

$$p''(b_1) = dualp'(b_1) = dual[0.38, 0.78] = [0.78, 0.38]$$

$$p''(b_2) = dualp'(b_2) = dual[0.26, 0.66] = [0.66, 0.26]$$

Let us calculate the required values of the posterior probabilities using (9) and taking into account (10).

$$p(a_1/b_1) = \frac{p(b_1/a_1)p(b_1)}{dualp\left(p(b_1/a_1)p(b_1) + p(b_1/a_2)p(b_2)\right)} = \frac{p(1)}{p''(b_1)} =$$

$$= \frac{[0.35, 0.63]}{[0.78, 0.38]} = \left[\frac{0.35}{0.38}, \frac{0.63}{0.78}\right] = [0.92, 0.81].$$

The remaining calculations will be done by analogy.

$$p(a_2/b_1) = \frac{p(3)}{p''(b_1)} = \frac{[0.03, 0.15]}{[0.78, 0.38]} = \left[\frac{0.03}{0.38}, \frac{0.15}{0.78}\right] = [0.08, 0.19]$$

$$p(a_1/b_2) = \frac{p(2)}{p''(b_2)} = \frac{[0.05, 0.21]}{[0.66, 0.26]} = \left[\frac{0.05}{0.26}, \frac{0.21}{0.66}\right] = [0.19, 0.32]$$

$$p(a_2/b_2) = \frac{p(4)}{p''(b_2)} = \frac{[0.21, 0.45]}{[0.66, 0.26]} = \left[\frac{0.21}{0.26}, \frac{0.45}{0.66}\right] = [0.81, 0.68]$$

For values $p(a_1/b_1)$ and $p(a_2/b_2)$, we have obtained improper intervals. This result is absolutely correct from the viewpoint of the theory of generalised intervals; however, it is incorrect from the position of common sense. To get the correct result, we simply invert these improper interval using (6). As a result, we have

$$p(a_1/b_1) = [0.81, 0.92], \quad p(a_1/b_2) = [0.19, 0.32],$$

$$p(a_2/b_1) = [0.08, 0.19], \quad p(a_2/b_2) = [0.68, 0.81].$$

It is easy to see that the resulting intervals are reachable intervals. The check for reachability has to be made for pairs of intervals $p(\cdot/b_1)$ and $p(\cdot/b_2)$).

The target state of information in the form of a probability tree is represented in Fig. 2.

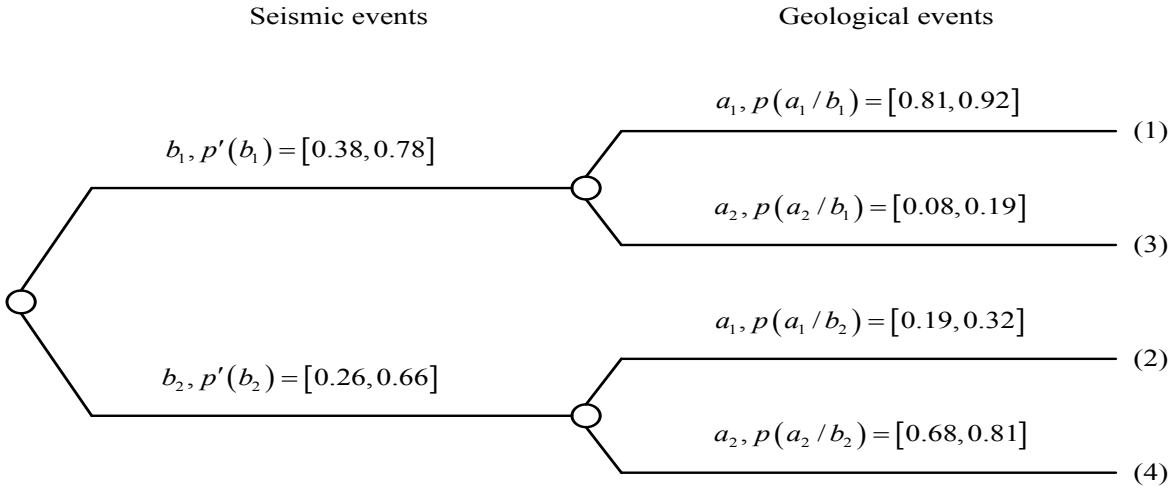


Fig. 2. Target state of information in the task of oil mining.

IV. Conclusions

The use of interval probabilistic values is aimed at modelling uncertainties regarding these values in a specific way. In the past decades, multiple variants of procedures of probabilistic inference at interval values of relevant probabilities have been proposed. The present paper shows that the most appropriate interval version of Bayes' formula is the version proposed in [7]-[8]. Utilization of the concept of generalised probability intervals helps to simplify the necessary calculations. This version is logically validated and does not require application of complicated concepts as it takes place in Weichselberger version [2]-[4]. The shortcoming of the considered technique is that when calculating the posterior values of probabilities, improper intervals of those values might be obtained. This shortcoming, however, can be easily overcome by means of simple inverting of improper probabilistic intervals.

Probabilistic inference under interval values has found wide application in different fields of science and technology. A considerable number of publications on this topic can evidence this. Out of numerous publications, paper [13] should be mentioned which presents a qualitative review on application of imprecise values in engineering. Papers [14]-[16] provide examples of application of interval values of probability in the processes of decision analysis and choice.

The use of procedures of interval probabilistic inference seems to be prospective in tasks of ecological risk assessment and making decisions related to the protection of the environment since a significant degree of uncertainty of initial information is characteristic of this research area.

References

[1] P. Walley, Statistical Reasoning with Imprecise Probabilities, London: Chapman and Hall, 1991.

[2] K. Weichselberger, The Theory of Interval-Probability as a Unifying Concept for Uncertainty:1st Int. Symposium on Imprecise Probabilities and Their Application, 29 June – 2 July, 1999, Ghent, Belgium. 1999, 9 p.

[3] K. Weichselberger and Th. Augustin. On the Symbiosis of Two Concepts of Conditional Interval Probability. In J. Bernard, T. Seidenfeld, and M. Zaffalon (Editors), ISIPTA'03, Waterloo: Carleton Scientific, 2003, pp. 608 – 630.

[4] K. Weichselberger, Elementare Grundbegriffe einer allgemeinen Wahrscheinlichkeitsrechnung I – Intervallwahrscheinlichkeitn als umfassendes Konzept. In cooperation with T. Augustin and Wallner A., Heidelberg: Physica, 2001.

[5] V. Kuznetsov, Interval Statistical Models. Moscow: Radio i Svyaz Publishers, 1991. (In Russian)

[6] L. M. De Campos, J.F. Huete and S. Moral, "Probability Intervals: A Tool for Uncertain Reasoning, " in Int. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. Vol. 2, No. 2, 1994. pp. 167 – 196.

[7] Y. Wang, "Imprecise Probabilities with a Generalized Form," Proc. 3rd Int Workshop on Reliability Engineering.

[8] Y. Wang, "Imprecise probabilities based on generalized intervals for system reliability assessment", in Int. Journal Reliability and Safety, Vol. 4, No. 4, 2010.

[9] E.M.A. Gardenes, L. Sainz, R. Jorba et al., "Modal Intervals", Reliable Computing, 7 (2), 2001, pp. 77 – 111.

[10] S.R. Shary, "A new technique in system analysis under interval uncertainty and ambiguity". Reliable Computing, 8 (2), 2002, pp. 321 – 418.

[11] Popova E.D., "Multiplication distributivity of proper and improper intervals". Reliable Computing, 7 (2), 2001, pp. 129 – 140.

[12] E. Kaucher, "Interval analysis in the extended interval space IR". Computing Supplement, 2, 1980, pp. 33 – 49.

[13] M. Beer, S. Ferson, and V.Kreinovich, "Imprecise probabilities in engineering analyses". Mechanical Systems and Signal Processing, Vol. 37, Issue 1 – 2, 2013, pp. 4 – 29.

[14] L.M. Zeinalova, "Decision making with combined states under internal uncertainty". Procedia Computer Science, 102, 2016, pp. 223 – 230.

[15] S. Rinard, "A Decision Theory for Imprecise Probabilities". Philosophers' Imprint, Vol. 15, No. 7, 2015, pp. 1-16.

[16] G. Jeantet and O. Spanjaard, Optimizing the Hurwicz criterion in decision trees with imprecise probabilities: Proceedings of the 1st Int. Conference on Algorithmic Decision Theory. October 20 – 23, 2009, Venice, Italy, pp. 340 – 352.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 170-174

Analysis of Artificial Intelligence Applications for Automated Testing of Video Games

Imants Zarembo
Faculty of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
imants.zarembo@rta.lv

***Abstract*—Game testing is a software testing process for quality control in video games. Game environments, sometimes called levels or maps, are complex and interactive systems. These environments can include level geometry, interactive entities, player and non-player controllable characters etc. Depending on the number and complexity of levels, testing them by hand may take a considerable effort. This is especially true for video games with procedurally generated levels that are automatically created using a specifically designed algorithm. A single change in a procedural generation algorithm can alter all of the video game levels, and they will have to be retested to ensure they are still completable or meet any other requirements of the game. This task may be suitable for automation, in particular using Artificial Intelligence (AI). The goal of this paper is to explore the most promising and up-to-date research on AI applications for video game testing to serve as a reference for anyone starting in the field.**

Keywords*—*Artificial Intelligence, Software Testing, Test Automation, Video Game Testing.

I. Introduction

Video game industry has seen a major expansion in recent years with the number of games being produced rapidly increasing and global games market value rising year over year and reaching $134.9 billion in 2018 [1]. Video game development complexity has grown over the years as well, starting from early generations consisting of simplistic or no graphics at all and restricted to a limited number of commands entered through a keyboard, to modern games with realistic graphics and highly interactive scenarios. This increase in complexity has led to an increase in effort required to ensure quality. Testing is an essential quality assurance activity in software engineering. Software testing is a process of evaluation of the functionality of a software application with an intent to find out whether the developed software meets the specified requirements and to identify defects. In comparison with general software development, video game quality assurance must take into consideration several additional aspects, such as [2][3]:

- Fun factor testing;
- Balance testing;
- Game level/World testing;
- AI testing;
- Multiplayer/Network testing;
- Audio testing;
- Physics testing etc.

Due to increasing demand from game development companies, many video games use procedural generation techniques to generate content [4] to ensure quality and quantity of the content, thus increasing replay value. An example of such procedural generation is game levels which can be automatically created using specifically designed algorithms, which means that player can have new game levels every time he starts the game. Such game levels consist of level geometry, interactive entities, player and non-player characters etc. Testing procedurally generated levels by hand may take a considerable effort and may be a suitable task for automation, in particular using AI for playtesting.

Test automation is a widely used technique of employing special software to control the execution of tests and the comparison of actual outcomes with predicted outcomes. Although automated testing still has its challenges [5] it is widely used in the software development industry for quality assurance. In comparison automated testing in video game development is a less common practice. One reason for that is the fact that video games consist not only from source code but also from assets such as 3D models, textures, sound, music, maps, puzzles etc. [6]. Traditional software quality assurance techniques are not applicable in this case.

The goal of this paper is to explore the most promising and up-to-date research on AI applications for video game testing to serve as a reference for anyone starting in the field. This analysis is the first step in research on creating a framework for automated video game level testing using AI, that would be applicable to procedurally generated video game level testing and validation with as little external involvement as possible. Such an approach would allow game developers to allocate more development time to other parts of the project and provide more value for their customers.

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4158
© 2019 Imants Zarembo.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

II. Related Work

A number of approaches have been proposed in the literature to test video games. There is a considerable number of video game testing techniques available that do not rely on traditional software testing techniques. Iftikhar et al. [7] in their paper propose a model-based testing approach for automated black box functional testing of platform games. Peterson et al. [8] present a system and method for performing external and automated testing of video games. Cho et al. [9] propose a system which supports black-box testing and scenario-based testing as well as simple load testing.

Only automatic or semi-automatic approaches focusing on those that use AI were of interest for purposes of this research. Nantes et al. [10] in their work propose a general software framework that integrate Artificial Intelligence Agents and Computer Vision technologies to support the test team and help to improve and accelerate the test process. The agent can replicate the user actions previously tracked in an older version of the game to check for visual anomalies in a newer build of the game. This approach allows making regression test process for environments automatic with no need for any other information about the internal architecture of the game. Gudmundsson et al. [11] present an approach to learn and deploy human-like playtesting in computer games based on deep learning from player data. The proposed approach is able to learn and predict the most "human" action in a given position through supervised learning on a convolutional neural network. The learned network can be used to predict key metrics of new content. The main focus of the approach is on estimating the difficulty of a new level instead of quality assurance of it. Chan et al. [12] present an approach to use evolutionary learning of behaviour to improve testing of commercial computer gamers.

A. Human playing style imitation

One of the uses of human playing style imitation in video games is to understand how a particular player would have played some game content without having the player taking the time to play through the game content [13]. This is especially useful in search-based procedural content generation, where a simulation-based evaluation function uses an AI to play through the candidate game content, assigning a numerical fitness value depending on how playable the content is. The fitness of the level might depend on whether an AI can play through the level or not. This can be used to evaluate content, to test game levels to see if they have bugs and whether they could be completed by a human player. In their paper Ortega et al. compare several different methods for imitating human player behaviour outlined in the following paragraphs.

1) Heuristic

A very simple approach that is based on hand-coded rules that features no learning and ignores the game environment. An example of this approach would be an NPC which simply moves in a certain direction and jumps whenever possible.

2) Artificial neural networks

An artificial neural network (ANN) can be used to simulate human behaviour. A supervised learning ANN approach makes use of direct representation by using the game environment information obtained from human gameplay as training set [14]. This approach uses backpropagation to minimize the error between human player actions and ANN outputs. A neuro-evolutionary approach attempts to minimize a fitness value corresponding to the mean squared error distance from the desired output (human actions) [15].

3) Dynamic scripting

Dynamic scripting (DS) is an online competitive machine-learning technique for game AI, that can be characterized as stochastic optimization [16]. DS contains a rule base with the possible rules that can be applied to a game. Each rule has a weight which reflects how well that rule made the agent perform in prior games. In every game, a script is generated using roulette-wheel in order to select a small subset of the rules in the rule base. The agent will play according to the rules contains in the script and those will have their weights updated via a standard Widrow-Hoff delta rule which is based on the immediate reward received by the environment.

4) REALM

REALM is a rule-based evolutionary computation agent for playing a modified version of Super Mario Bros [17]. REALM follows the principle of learning classifier systems, by which rules are evolved according to a fitness value. Each rule contains conditions based on different information obtained from the game. REALM classifier includes high-level plans of action instead of simple reactive combinations of key presses.

5) Grammatically evolved behaviour trees

Behaviour trees provide a top-down organization from the root of the tree down to the leaves [18]. The control nodes are those that decide which branches of the tree will be executed next, while the leaf nodes contain the actions that are going to be carried out. The different elements of the tree are specified in a grammar which is evolved by applying genetic operations to the sub-tree created. While the evolutionary mechanism is similar to that used in neuroevolution, the behaviour tree representation differs significantly from both neural networks and dynamic scripting.

B. Playtesting with procedural personas

Archetypal player models called procedural personas can be used for generative player modelling and automatic testing of game content [19]. The approach uses a variant of Monte Carlo tree search with genetic programming applied to trees instead of Upper Confidence Bound 1 to evolve persona-specific evaluation formulas. This allows finding mappings between persona utility functions and state evaluation algorithms.

In [20] authors present a method where procedural personas act as critics in search-based procedural content generation framework. For this purpose, personas have

been evolved on a set of authored dungeons, according to different fitness that matches archetypical decision-making priorities.

C. ICARUS

Intelligent Completion of Adventure Riddles via Unsupervised Solving [21] is a framework for autonomous video game playing, testing and bug reporting which is based on discrete reinforcement learning in a dualistic fashion, encompassing volatile short-term memory as well as persistent long-term memory that spans across distinct game iterations. It can iterate through complete game iterations and detect or aid the detection of all major bug categories.

D. Hyper-heuristics

Hyper-heuristics approach [22] consists of the creation of hyper-agent for general video game playing that utilizes the strengths of multiple individual controllers to play unseen games better than any of them individually. The hyper-agent uses an offline learning approach by acquiring information about controller performance from a set of trained instances and create a selection model that generalizes well for new games. Hyper-agent does not directly control the main character but selects the best controller to do so.

E. Rolling horizon evolution

Rolling Horizon Evolutionary Algorithms (RHEA) [23] are an alternative to Tree Search for action-decision making in real-time games. Evolutionary Algorithms are used in conjunction with a simulator to train a controller offline and the use the already evolved controller to play the game. RHEA approaches employ evolution in a similar way to how it is done in a tree search, using a forward model to evaluate a sequence of actions.

F. Active learning

Active learning selects among a set of possible inputs to get the best output while minimizing the number of inputs tested. Authors of [24] define the best output as a parameter tuning design goal and treat a set of game design parameters as an input. Minimizing the number of inputs tested minimizes the number of playtests performed.

G. Genetic algorithms

In [25] Genetic algorithms are explored to learn levels from the Mario AI simulator, which is based on Infinite Mario Bros game. Agents learn a sequence of actions by using a genetic algorithm with integer encoding, in order to maximize the attained score after ending the level. This approach executes two different stages: in the first, domain-independent genetic operators are used, while in the second knowledge about the domain is incorporated into these operations in order to improve results.

III. Materials and Methods

Existing research in this field was analysed and synthesised based on whether the described approaches were applicable to automated video game testing using AI. First, research on manual video game testing and testing that does not focus on application of AI was discarded. Then research on automated video game playing using AI was included as automating video game playing and testing are similar tasks and partially overlap in many cases. Finally, an OWL ontology containing a semantic representation of the results of this research was constructed using Protégé and is described in the following section. As automated video game testing research field is relatively new (especially using AI, but in general as well) the main purpose of the ontology is to serve as a starting point for future research in the field.

IV. Results and Discussion

Automated video game testing using AI is relatively new research field often lacking established terminology and structure. This research tries to give overview of automated video game testing approaches and proposes a simple categorization of approaches.

Fig. 1. Spectrum of design testing methods in game development [19].

Fig. 1. Shows the position AI agent-based playtesting takes in the spectrum of video game testing methods right between informal playtesting and structured playtesting.

Analysed video game testing approaches can be broadly categorized into three categories:

1) Human imitation approaches

Human imitation approaches strive to imitate human players in some way to produce results similar to those a human would produce playtesting. This is the most well researched and widely used category of the overviewed approaches. There appears to be a connection between general game AI research that strives to create AI for non-player characters in games that exhibit a behaviour similar to that of human players, and automated video game testing where the quality of video game content must be assured. In both cases, autonomous AI agents can be used to play the game but with different goals in mind.

2) Scenario-based approaches

Scenario-based approaches at least partially rely on previously prepared data and rules to decide further course of actions – which scenario to follow, e.g. dynamic scripting, REALM.

3) Goal-based approaches

Approaches that rely on defining game goals for AI agents to reach fall into this category, e.g. hyper-heuristics, reinforcement learning etc.

A. Ontology

An OWL ontology was constructed categorizing video game testing approaches analysed in this paper (see Fig. 2).

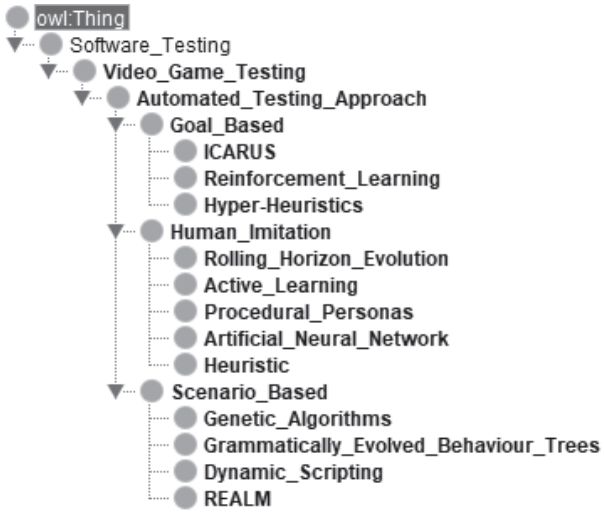


Fig. 2. Automatic video game testing approach ontology

Root node of the ontology starts from "Software Testing" class, which include "Video Game Testing" subclass, followed by "Automated Testing Approach" subclass. Three main subclasses of "Automated Testing Approach" are "Goal Based", "Human Imitation" and "Scenario Based". All of the analysed testing approaches are included in one of these categories.

For the sake of visualization simplicity of the ontology video game testing approaches were included as classes instead of individuals.

B. Performance

Video game playtesting technique performance and efficiency was out of scope of this paper, the original papers should be consulted for more detail. Alternatively, The General Video Game AI (GVGAI) [26] is a competition to explore the problem of creating controllers for general video game playing. The competition has rankings in several categories:

- Single-player planning;
- Level generation;
- Rule generation;
- Two-player planning;
- Single-player learning.

Depending on the game to be tested exploring these categories may yield well performing appropriate approaches for the task at hand.

V. Conclusions and Future Work

The paper provides an overview of existing automated video game testing approaches and serves as the first step in research of automated procedurally generated game level testing using AI. The vast majority of the analysed approaches of video game testing automation rely on playtesting to produce the results. This research would benefit from more comprehensive analysis including available techniques in connected fields that may be applicable for automated videogame testing but have not yet been adapted for this task.

Future work includes several tasks: defining requirements for the testing automation task to solve; narrowing down the most promising approaches to further analyse, implement and compare their performance; creation of a research prototype of a game which levels an AI agent can test and produce a report; expanding and refining the video game testing approach ontology.

References

[1] J. Batchelor, "GamesIndustry.biz presents... The Year In Numbers 2018," December, 2018. [Online]. Available: https://www.gamesindustry.biz/articles/2018-12-17-gamesindustry-biz-presents-the-year-in-numbers-2018 [Accessed: Feb. 12, 2019].

[2] L. Levy and J. Novak, "Game Development Essentials: Game QA & Testing," Cengage Learning, 2009, pp.58-70.

[3] C. P. Schultz and R. D. Bryant, "Game Testing: All in One," Mercury Learning & Information, 2016, pp.125-128.

[4] M. Hendrikx, S. Meijer, J. V. D. Velden and A. Iosup, "Procedural content generation for games: A survey," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 9, Feb 2013.

[5] A. Bertolino, "Software Testing Research: Achievements, Challenges, Dreams," Proceedings of The Future of Software Engineering at ICSE 2007, pp. 85-103, 2007.

[6] J. H. Becares, L. C. Valero and P. P. G. Martin, An approach to automated videogame beta testing," Entertainment Computing, vol. 18, pp. 79-92, Jan 2017.

[7] S. Iftikhar, M. Z. Iqbal, M. U. Khan and W. Mahmood, "An automated model based testing approach for platform games," The 18th International Conference on Model Driven Engineering Languages and Systems (MODELS), Nov 2015.

[8] K. Peterson, S. Behunin and F. Graham, "Automated Testing on Multiple Video Game Platforums," U.S. Patent 13/020,959, 4 Feb., 2011.

[9] C. S. Cho, K. M. Sohn, C. J. Park and J. H. Kand, "Online Game Testing Using Scenario-based Control of Massive Virtual Users," The 12th International Conference on Advanced Communication Technology (ICACT), Feb. 2010.

[10] A. Nantes, R. Brown and F. Maire, "A Framework for the Semi-Automatic Testing of Video Games," Artificial Intelligence and Interactive Digital Entertainment (AIIDE) 2008.

[11] S. F. Gudmundsson, P. Eisen, E. Poromaa, A. Nodet, S. Purmonen, B. Kozakowski, R. Meurling and L. Cao, "Humsn-Like Playtesting with Deep Learning," IEEE Conference on Computational Intelligence and Games (CIG), Aug. 2018.

[12] B. Chan, J. Denzinger, D. Gates, K. Loose and J. Buchanan, "Evolutionary behavior testing of commercial computer games," Proceedings of the 2004 Congress on Evolutionary Computation, Jun. 2004.

[13] J. Ortega, N. Shaker, J. Togelius and G. N. Yannakakis, "Imitating human playing styles in Super Mario Bros," Entertainment Computing, vol. 4, pp. 93-104, Apr. 2013.

[14] A. J. Champandard, "The Dark Art of Neural Networks," in "AI Game Programming Wisdom," S. Rabin, Charles River Media, 2002, pp. 640-651.

[15] D. Floreano, P. Durr and C. Mattiussi, "Neuroevolution: from architectures to learning," Evolutionary Intelligence, Springer, vol. 1, pp. 47-62, Mar. 2008.

[16] P. Spronck, M. Ponsen, I. Sprinkhuizen-Kuyper and E. Postma, "Adaptive game AI with dynamic scripting," Machine Learning, Springer, vol. 63, pp. 217-248, Mar. 2006.

[17] S. Bojarski and C. B. Congdon, "REALM: A rule-based evolutionary computation agent that learns to play Mario," Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games, Aug. 2010.

[18] A. J. Champandard, "Understanding Behavior Trees," Spetember, 2007. [Online]. Available: http://aigamedev.com/open/article/bt-overview/ [Accessed: Feb. 17, 2019].

[19] C. Holmgard, M. C. Green, A. Liapis and J. Togelius, "Automated Playtesting with Procedural Personas through MCTS with Evolved Heuristics," CoRR, 2018.

[20] A. Liapis, C. Holmgard, G. N. Yannakakis and J. Togelius, "Procedural Personas as Critics for Dungeon Generation," European Conference on the Applications of Evolutionary Computation, pp. 331-343, 2015.

[21] J. Pfau, J. D. Smeddinck and R. Malaka, "Automated Game Testing with ICARUS: Intelligent Completion of Adventure Diddles via Unsupervised Solving," CHI PLAY Extended Abstracts, pp. 153-164, 2017.

[22] A. Mendes, J. Togelius and A. Nealen, "Hyper-heyristic general video game playing," IEEE Conference on Computational Intelligence and Games (CIG), 2016.

[23] D. Perez, S. Samothrakis, S. Lucas and P. Rohlfshagen, "Rolling horizon evolution versus tree search for navigation in single-player real-time games," Proceedings of the 15th annual conference on Genetic and evolutionary computation, pp. 351-358, Jul. 2013.

[24] A. Zook, E. Fruchter and M. O. Riedl, "Automatic playtesting for game parameter tuning via active learning," Foundations of Digital Games, 2014.

[25] A. Baldominos, Y. Saez, G. Reico and J. Calle, "Learning Levels of Mario AI Using Genetic Algorithms," Conference of the Spanish Association for Artifical Intelligence, pp. 267-277, 2015.

[26] D. Perez, S. Samothrakis, J., Togelius, T. Schaul and S. M. Lucas, "General Video Game AI: Competition, Challenges and Opportunities," Association for Advancement of Artificial Intelligence (AAAI), 2016.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 175-179

Mathematical Description and Modelling of Transportation of Cargoes on the Base Digital Railway

Natalia Zhuravleva
Transport economy Department
Emperor Alexander I St. Petersburg
State Transport University
St. Petersburg, Russia
zhuravleva_na@mail.ru

Ilia Guliy
Transport economy Department
Emperor Alexander I St. Petersburg
State Transport University
St. Petersburg, Russia
ilya.guliy@mail.ru

Mark Polyanichko
Department of Information Technology and IT Security
Emperor Alexander I St. Petersburg
State Transport University
St. Petersburg, Russia
polyanichko@pgups.ru

Abstract—**This article presents the results of a mathematical description of the transportation process of goods by rail on the exit routes. The parameters reflecting state of time, speed and cost of the actual performance of the freight transportation are simulated, which makes it possible to identify and respond in time to the risk caused by interaction between adjacent subjects and objects for transportation. An algorithm to respond to a decrease in speed of transportation is determined.**

It is substantiated that the efficiency in transportation provides the level of development in transport and logistics system as an infrastructure of a new technological order. The price of these systems generates added value due to speed, inter modality services, drawing up optimal routes for cargo delivery, ensuring full car loads, passage control of goods at all stages of the logistics chain, etc., i.e. through the integration of products and services, considering the dominant global network of production and consumption.

This work is an implementation element of digital formats in the operational activity of railways. The created model implements the "traceability" of information about the movement of cargo traffic in exit routes, generating Fast Date for time-sensitive decision-making process and Smart Date for asymmetric analytics. In contrast to the traditional model of transportation, the proposed solution is based on a mathematical description of all stages of the life cycle of freight (trains), which allows evaluating all costs by type of each process of transportation (movement and idle time) in real time mode. This approach takes into account the "investment" in the formation of value of all enterprises involved in transportation, including the condition and operation of technical infrastructure, locomotives, locomotive crews, wagon and freight facilities employees, and the movers themselves who provide and manage the transportation process.

It is proved that the further growth of the profitability of the transport business is in direct correlation with the increase in the marginal profitability of shippers, and decrease of the transport component in the final price of goods achieved as result of digitizing the process of cargo transportation in the exit routes.

The research methodology is based on the process-functional approach to describing the life cycle of a freight train, analysis factor for technical and economic characteristics of the transportation process and dynamic modelling of the parameters of significant means of elements affecting the transportation process. The information basis of the study relies on a representative sample of loading and unloading of goods in areas of mass traffic. We have investigated dependent (homomorphic) and independent (singular) pairs in accordance with the time, cost, and technical parameters.

Keywords—asymmetric analytics, digitization of transportation, cost of speed, time, dynamic model, Smart Date.

I. INTRODUCTION

Analysis of the studies in the field of efficiency and mobility of product markets highlight the key problem – ensuring the transportation's speed growth. Hereupon, the economic description of a "smart" rail system, namely, its ability to generate revenues and optimize transportation costs, is considered fragmentarily. Moreover, there is a definite demand for faster delivery from the shippers, but there is no unambiguous understanding of the technology of economic estimation of speed. The literature presents a number of approaches to the creation of competitive models of rail transportation and their dependence on frequency and speed of traffic, primarily related to high-speed rails (HRS) [1] – [2].

However, there are still no solutions related to the rational behaviour of consumers of transport services depending on time and cost of transportation.

Multiple technological innovations changing the chain of production and consumption form a new value network in the consumer [3] - [4].

Transition from mass production to customized production and from full-cycle enterprises to concentration on competitive advantages will significantly change the transportation service itself: small batches under the order with delivery "precisely in time" for long distances. Eventually, a new business model of transport organizations is shaped. Orientation to

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4049
© 2019 Natalia Zhuravleva, Ilia Guliy, Mark Polyanichko.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

a specific consumer and the utmost use of the information as a propelling resource, considering the specific features of a specific consumer in a specific place and full use of digital transformation technologies of real business processes, rebuild the whole scheme of relationships in the economy and society. This confirms the idea that the strategy of new business models is completely changed under the influence of the network of transformations of the value proposition of a new transport service [5].

In this regard, the purpose of this work is to form a model of transformation for business processes of freight transportation in the exit routes, taking into account changes in the state of the market and the behaviour of consumers of transport services on it. Ultimately, we have to discover a double effect of speed growth: growth in the profitability of the business in the transportation company and increase in the marginal profitability of shippers due to a decrease in the transportation's price component.

The exit route taken in the study is the most representative for confirming the results of the simulation, since it is formed in areas of mass cargo loading, to the areas of mass unloading. Depending on the capacity of freight traffic, the sending routes are determined by assignment of unloading on one station or section (direct routes) or in the break up of routes in the technical station nearest to the unloading area. Application of Big Data technology allows to process huge amounts of data in real time mode.

Thus, the results of the mathematical description and modelling of cargo transportation in the exit routes proposed in the paper can be presented as part of the overall process of "digitizing" the railway.

II. MATERIALS AND METHODS

The solution for the problem of qualitative and quantitative assessment of the actual execution of the process of cargo delivery is possible only with a large array of data on volumes and addresses, timing, speed, intensive procedure and cost of transportation parameters. The reflection of the results is recorded in real time (Tame Date), which allows to quickly make the right decisions in the right place, preventing a rise of losses in the complex process of transportation.

Data processing technology includes three main stages.

In the first stage, the results of a representative sample that characterizes the integrity of freight train processing operations are processed:

- Motion parameters in the composition of the various categories of trains;
- Idle times of trains for various categories at intermediate stations;
- Time for transit operations for various categories of trains at technical stations;
- Step-by-step processing time at technical stations;
- Time of start to end operations.

A traditional factor analysis tool is used to estimate the weight of each factor on the amount's loss of infrastructure resources, rolling stock, and freight transportation costs in general, as well as the formation of a model describing the technical and economic characteristics of the transportation process. The scaled-up scheme of the transportation model is shown in Figure 1.

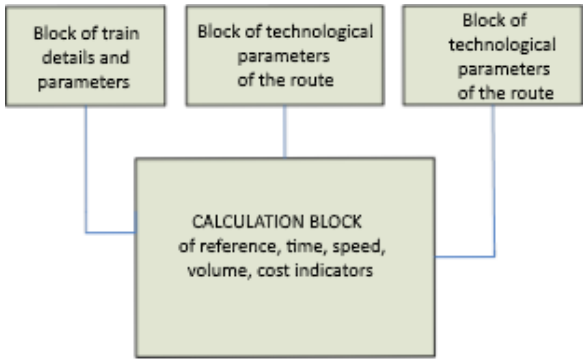
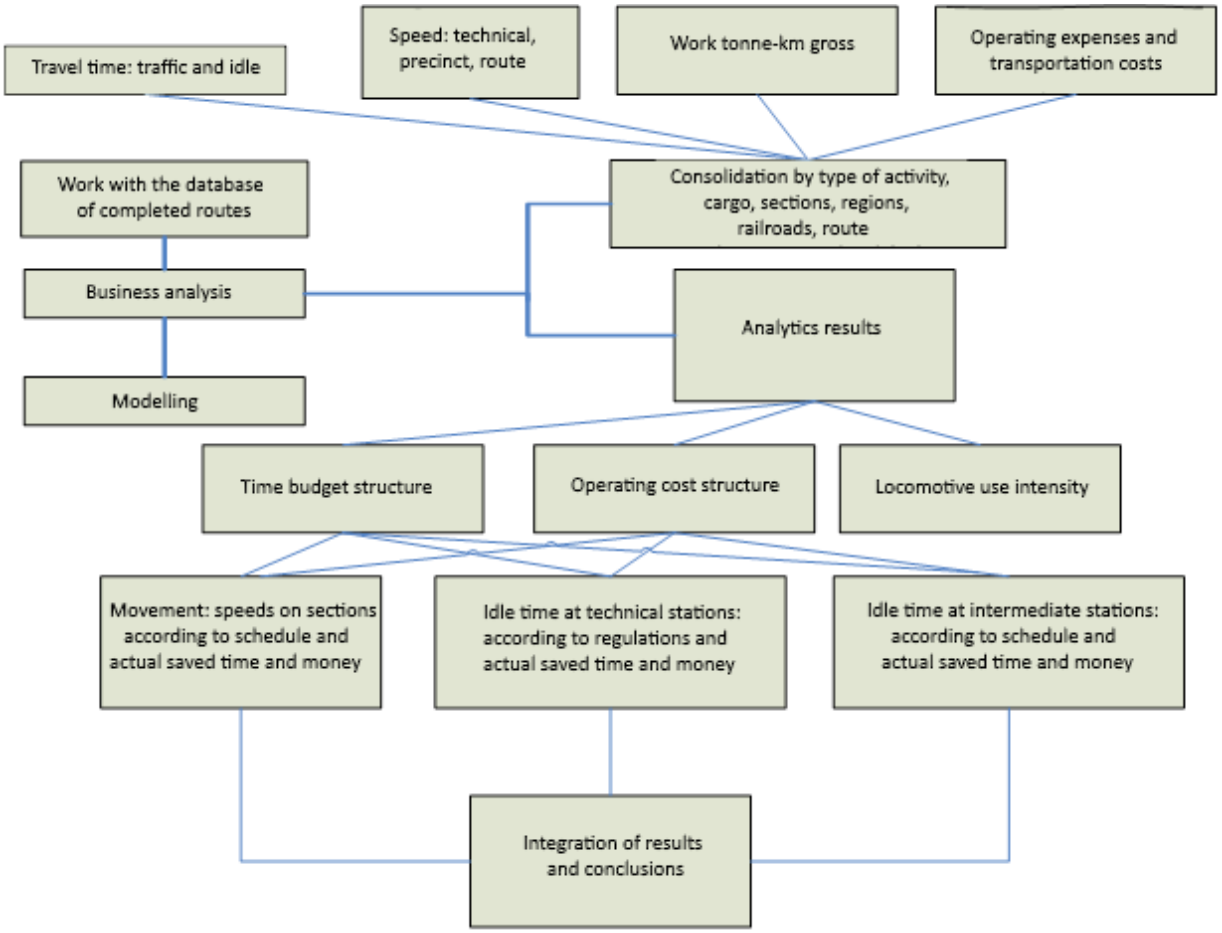


Fig.1. Scheme of transportation model

In this case, the calculation unit includes information:

- Details of the route and the train: cargo, railway, train number, weight, length, station's location, dispatching control subdivision;
- Section length, accumulated passage length;
- Time of formation of the route;
- Time of running, arrival, departure at the station of location;
- Calculation of the time of passage of the section and idle time at the station;
- Identification of idle time at intermediate and technical stations;
- Accumulation of moving and idle time;
- Calculation of technical, sectional, through speed along the section, and accumulated speed;
- Calculation of gross ton-kilometers for the section and in the accumulation, calculation of the car-hours in movement, idle and total;
- Calculation of the actual costs of transit operations during idle time and at technical stations;
- Calculation of operating costs and transportation costs.

The second stage is the dynamic modelling of the parameters of significant factors' weight affecting the transportation process. The purpose of calculation is to determine the costs of the transportation process, "as it is" i.e. in relation to the existing technology of organization of traffic flows and formation of trains and routes of the respective categories. Any change to the existing organization of traffic flows will cause an assessment on the effectiveness of this change in the situation "as it became". The depth and specification of the description of the business processes and its transportation activities of railways must take into account the technology and its changes, regulations and orders governing the organization of infrastructure, traction, and cars.



We presume the following dynamic changes in the set of primary models or patterns describing the advantages in the value proposition of the transport service.

a. interactions and relationships between primary models – description of the advantages in the network of values (service business model);

b. processes (predictive maintenance, status tracking);

c. economic interaction between the objects of economic action (resource sharing, instant response) in space and time.

In general, the block diagram of the functional model of freight transportation is shown in Fig.2.

Fig.2. Block diagram of the functional model of freight transportation

The third stage is implemented in the process-functional approach to describing the life cycle of the train. Tonne-kilometres gross when moving trains are calculated in two dimensions: by sections and at an approach to the station, taking into account the movement and idle time of the train.

Estimated travel time from the point of departure to the point of destination of the train's destination allows evaluating all expenses by types of economic activity of each transportation process (movement and idle time) in real time mode.

Thus, the cost for idle time per minute for 1 train at technical station x_i will be:

δ_{ij} – transit car handling costs using resources of the enterprise i for all operations j:

$$\delta_{ij} = \sum_{k=1}^{K} e_{i\,jk} * t_{ijk} \quad (1)$$

δ_{ik} – transit car processing costs using resources of the enterprise i for all types of processing k:

$$\delta_{ik} = \sum_{j=1}^{J} e_{i\,jk} * t_{ijk} \quad (2)$$

δ_{jk} – transit car processing costs for all operations j and for all kinds of processing k:

$$\delta_{jk} = \sum_{i=1}^{I} e_{i\,jk} * t_{ijk} \quad (3)$$

δ_{i} – costs incurred by the enterprise i for transit car processing for all operations j and for all kinds of processing k:

$$\delta_i = \sum_{j=1}^{J} \sum_{k=1}^{K} e_{i\,jk} * t_{ijk}, j = 1..J, k = 1..K \quad (4)$$

δ_{k} – costs of transit car k-type processing for all enterprises i and for all for all operations j:

$$\delta_k = \sum_{i=1}^{I} \sum_{j=1}^{J} e_{i\,jk} * t_{ijk}, i = 1...I, j = 1...J \quad (5)$$

δ – transit car processing costs for all enterprises i, for all operations j, and for all types of processing k per minute per car:

$$\delta = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} e_{i\,jk} * t_{ijk}, i = 1..I, j = 1..J, k = 1..K \quad (6)$$

Then the cost of downtime per minute for 1 train at technical station xi will be:

$$c_i^{mc} = \delta * m \quad (7)$$

where m is the number of the train cars.

This approach takes into account the "investment" in the generation of value by all enterprises involved in the transportation including the condition and operation of the technical infrastructure, locomotives, locomotive crews, wagon and freight facilities employees, and the movers themselves who provide and manage the transportation process.

III. RESULTS AND DISCUSSION

The sequence of iterations of the freight transportation functionality over time allowed us to form a mathematical model of value added when processing transit trains at the precinct station. This tool allows you to estimate the cost of time in motion and idle time for the carrier and show the cost of speed for shippers. In this context, we relied on the classification and structure of models for optimizing transport processes "as per schedule", their characteristics and implementation algorithms are proposed by a number of authors during the last ten years [6].

However, almost all of these studies are based on a variety of hypothetical assumptions about the types of railways, types of infrastructure, etc. The development of Big Data technology allowed us to resolve such problems in real time based on real statistics bases on all operations of train preparation and movement, namely, to create a value-added model, which has a number of advantages related to the possibility of monetizing the reduction in transportation time.

The reliability of the results obtained can be confirmed using an integrated approach to the integration of various information environments reflecting operational data and correlating with the accounting and financial systems of the railway company, which are reflected in a number of studies, in particular [7]. As well as studies about the integration of various types of transport systems [8] – [9].

In the development of the discussion, it should be noted a number of studies on the formation of a holistic approach to the use of IIoT (Industrial Internet of Things) technology, allowing to carry out advanced monitoring of assets and operations of railway companies, expand the functionality of Passenger and Freight Information Systems (PIS/FIS), Train control systems, etc. A number of authors (Fraga-Lamas, P., Fernández-Caramés, T.M., Castedo, L.) believe that it is namely IIoT that ensure credibility of decisions to improve the efficiency of the transportation process [10].

In this study, the procedures and results of the Big Data mining. Hereupon, a number of authors, in particular Song, R., Xue, X., propose using neural networks in the processing of large databases. [11]. Their use will undoubtedly significantly improve the forecast estimation of the modelling of the cargo transportation process and will reduce the number of assumptions having a probabilistic nature.

An important part of the scientific discussion is devoted to the technical support of modelling various scenarios for the organization of rail transportation, in particular, the organization of wireless systems in the management of railway passenger transport systems, which is undoubtedly important when organizing mixed, passenger, and freight traffic on a single infrastructure [12].

Despite the breadth of discussion about various directions of transport processes modelling, value measurement research of the processes of goods transportation in real time is not enough. We can note more general research related to maximizing the cost of transport projects when using digitalization technologies, in particular Cruz, C.O., Sarmento, J.M. [13]. Valuation procedures of the cargo traffic processing scenarios for freight trains and at the stations using operating work simulation are described by Schneider M , Nießen N. [14].

Among the fundamental papers on transport modelling, the research by Anna Nagurney and John F. Smith, University of Massachusetts, Amherst, Massachusetts USA should be highlighted [15]. In particular, the description of various models of traffic flows and the process of their consolidation.

This once again confirms the importance and relevance of tracking, linking and modelling on functions and processes of costs simultaneously incurred by all railway departments at each stage of technological redistribution and reliably estimation of the value-added transportation cost.

IV. CONCLUSIONS

The overall result of the study should be considered as a dynamic model of railway transportation of goods in exit routes, allow you to determine the added value of the technical, district and route speed of a freight train across its entire route in real time. With the help of an asymmetric analytics tool based on graph theory, this study will be conducted in the direction of analysing the time budget of completed routes, taking into account all factors affecting it, considering the intensity of locomotive use and speed [16] – [17].

With further development of the model, its economic description and the "traceability" of the movement's speed in the tariff, it is possible to obtain an estimate on the effect of speed and its influence on the growth of commodity markets.

The authors state that:

1. They have no conflicts of interest;
2. This article does not contain any research involving human subjects as research objects.

REFERENCES

[1] O. Álvarez-SanJaime, P. Cantos-Sanchez, R. Moner-Colonques, J.J. Sempere-Monerris "A model of internal and external competition in a High Speed Rail line" *Economics of Transportation*, 2015, vol. 4, Iss. 3, pp. 178-187.

[2] R. Vickerman "Can high-speed rail have a transformative effect on the economy?" *Transport Policy*, 2018, vol. 62, pp. 31-37.

[3] V. Allee The Future of Knowledge. USA: Butterworth-Heinemann, 2002.

[4] C. Parolloni The Value Net: A Tool for Competitive Strategy. New York: Villey, 1999.

[5] M.W. Jonson, M.C. Christensen, H. Kadermann "Reinventing your business model", *Harvard Business Review*, December 2008 [Online]. Available: https://hbr.org/2008/12/reinventing-your-business-model. [Accessed: Jan. 10, 2019].

[6] P. Vilela, M. Cachoni, A. Vieira, L. Christofoletti, *Train circulation planning: Quantitative approaches*, Conference: 2017 Joint Rail Conference. Philadelphia, Pennsylvania, USA, April 4–7, 2017. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-5026810165&doi=10.1115%2fJRC2017-2223&partnerID=40&md5=5badb9c5064e86e8b7f16f70e8667a16. [Accessed: Nov. 10, 2017].

[7] I.V. Bychkov, A.L. Kazakov, A.A. Lempert, D.S. Bukharov, A.B. Stolbov "An intelligent management system for the development of a regional transport logistics infrastructure", *Automation and Remote Control*, 2016, vol. 77, Iss. 2, pp. 332-343. [8] Y. Ma, D. Chang, F. Wang "Integrating business processes of container sea-rail combined transport", *International Journal of Internet Manufacturing and Services*, 2019, vol. 6, No 1, pp. 48-63.

[9] L. Heilig, S. Voß, "Information systems in seaports: a categorization and overview", *Information Technology and Management*, 2017, vol. 18, Iss. 3, pp. 179-201.

[10] P. Fraga-Lamas, T.M. Fernández-Caramés, L. Castedo "Towards the internet of smart trains: a review on industrial IoT-connected railways", *Sensors* (Basel), 2017, vol. 17, no. 6, no. 1457, pp. 1-44. [11] R. Song, X. Xue, "Freight volume forecast based on improved radial basis function neural network", *Boletin Tecnico*, 2017, vol. 55, iss. 5, pp. 419-423.

[12] L. Azpilicueta, J.J. Astrain, P. Lopez-Iturri, F. Granda, C. Vargas-Rosales, J. Villadangos, A. Perallos, A. Bahillo, F. Falcone "Optimization and design of wireless systems for the implementation of context aware scenarios in railway passenger vehicles", *IEEE Transactions on Intelligent Transportation Systems*, 2017, vol. 18, iss. 10, pp. 2838-2850.

[13] C.O. Cruz, J.M. Sarmento "Maximizing the value for money of road projects through digitalization", Competition and Regulation in Network Industries, November 27, 2018. [Online]. Available: https://doi.org/10.1177/1783591718811436 [Accessed: Jan. 18, 2019].

[14] M. Schneider, N. Nießen "Minimising economic losses due to inefficient rescheduling", *Journal of Rail Transport Planning & Management*, 2016, vol. 6, iss. 2, pp. 128-140.

[15] A. Nagurney "Mathematical models of transportation and networks (Refereed Encyclopedia Article)" Mathematical Models in Economics, Encyclopedia of Life Support Systems, 2007 [Online]. Available: http://works.bepress.com/anna_nagurney/191/ [Accessed: Jan. 15, 2019].

[16] L. R. Ford, D. R. Fulkerson "Maximal flow through a network", *Canadian Journal of Mathematics*, 1956, vol. 8, pp. 399–404.

[17] J. R. Evans, E. Minieka Optimization algorithms for networks and graphs. 2nd ed. New York: Marcel Dekker, 1992.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 180-183

Safety of Artificial Superintelligence

Aleksejs Zorins
Faculty of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
Aleksejs.Zorins@rta.lv

Peter Grabusts
Faculty of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
Peteris.Grabusts@rta.lv

***Abstract*—The paper analyses an important problem of cyber security from human safety perspective which is usually described as data and/or computer safety itself without mentioning the human. There are numerous scientific predictions of creation of artificial superintelligence, which could arise in the near future. That is why the strong necessity for protection of such a system from causing any farm arises. This paper reviews approaches and methods already presented for solving this problem in a single article, analyses its results and provides future research directions.**

***Keywords*—Artificial Superintelligence, Cyber Security, Singularity, Safety of Artificial Intelligence.**

I. Introduction

The paper presents an overview of safety concepts and technologies introduced in the recent years in the area of artificial intelligence (AI) safety. There are different predictions but many AI researchers, philosophers and futurologists agree that in the next 20 to 200 years a machine capable to perform on at least human level on all tasks will be developed [3,4,7, 10, 14]. According to assumption that such a machine will be capable to design the next generation of even smarter intelligent machines, an intelligence explosion will take place very shortly. Despite different predictions regarding this issue, from possible economic problems till complete extinction of a mankind, many researchers agree that this problem is extremely important and urgently needs serious attention [1,2,5,8,19],

This problem is becoming more and more important in the scientific community due to several reasons: rapid development of AI both on software and hardware levels; wide implementation of AI in industry and other fields of human activity; the lack of control of super-AI. The research in this field is in a very beginning and at the same time is crucial to our development and safety.

The superintelligence problem is usually connected with a so-called singularity paradox, which could be described as follows: "superintelligent machines are feared to be too dumb to possess common sense" [10].

The machines have completely different discreet logic and structure, do not, and could not have emotions and feelings like humans. Even if the computer will decide to make a man happy it will do it with the fastest and cheapest (in terms of computational resources) without using a common sense (for example killing of all people will lead to the situation that no one is unhappy or the decision to treat human with drugs will also make him happy etc.). Another issue is that we want computer to that we want, but due to bugs in the code computer will do what the code says, and in the case of superintelligent system, this may be a disaster.

The next sections of the paper will show the possible solutions of making superintelligence safe to humanity.

II. Possible solutions for safe AI

In his research of Artificial Super Intelligence Roman Yampolskiy presents a comprehensive review of potential solution methods of singularity and safe AI problem [18]. All possible solutions are grouped into several categories according to a way of dealing with the problem:

- ✓ Prevention of development - the researchers of this group completely denies AI and consider only one method for dealing with the problem – prohibit the development of AI;
- ✓ Restricted development – here the idea is to restrict superintelligent machines with different approaches: software, hardware and mixed ones;
- ✓ Incorporation into society - the authors are sure that intelligent machines should be a part of human society (in economic, legal, religious, ethical, moral and / or educational aspects) and this will help to successfully solve the problem;
- ✓ Implementation of self-monitoring algorithms of AI – create a set of rules to follow, development of human-friendly AI, include emotions into AI behaviour algorithms etc.;
- ✓ Other solutions – join AI, deny of the problem, relax and do nothing and some other proposals.

Prevention of development of AI is the most aggressive and straightforward method and probably the most effective one. However, taking into account the modern society and the level of inclusion of computers in

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4042
© 2019 Aleksejs Zorins, Peteris Grabusts.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

our life it is very unlikely that laws of banning computers will be accepted in a near future. Even if some of the worlds governments will incorporate such a law into its legislation there will always be countries or even several individuals who violates regulations.

Restricted development is the most commonly accepted solution to the safe AI problem. AI-boxes, leakproofing and restricted question-answering-only systems (Oracle AI) are among the proposed solutions in this group [3,4,16,17,20]. The methods presented in this category are similar to putting a dangerous human being into prison – it does not give a 100% safety but in most cases can help society to survive for some period of time. It is clear that this solution cannot help in the long-term but could be a good initial measure when the real superintelligence will be developed.

Incorporation into society is a very easy to implement. We can include into computer algorithms social, moral, ethical and other rules, but an intelligent machine with a digital mind will easy and fast discover the drawbacks, contradictions and disadvantages of our legislation, moral and ethical rules. A very interesting idea of raising AI like a child also belongs to this group [11], however, the grow-up children often are very different from their parents expectations despite all efforts.

The self-monitoring algorithms category includes explicitly hard-coded rules of behaviour into computer and creation of multilevel guard composed of clever machines to monitor each other. The set of rules is a good solution but it cannot cover every possible situation and if such a situation occurs, the computer may act in an unpredicted manner. The computer watch guard will lead to a hugely increased system which could not be handled by human and giving all right to AI sooner or later will have very bad effects.

The final category includes extreme approaches of battling against machines or doing nothing because a clever AI will defeat us in any way. Another opinion held by several researchers and businessmen including the owner of Tesla Elon Musk is that at the moment the only feasible solution to this problem is joining the researchers of AI (for example OpenAI project with the mission "discovering and enacting the path to safe artificial general intelligence" [13]) to be aware of technological advances in this field and to be able to react quickly depending on the situation.

As one may see, some of these proposals are completely unacceptable and will not be analysed in the paper while some needs to be described in more details. The next section will discuss some of the most interesting and realistic methods.

III. Engineering Safe Superintelligence

David Chalmers firstly introduced the main idea of this approach in 2010, who suggested that for safety reasons AI systems should be restricted to simulated virtual worlds until their behaviour will be fully understood under controlled conditions [19].

Roman Yampolskiy in his book "Artificial Superintelligence" proposes a formalized notion of AI confinement protocol; for restricting an AI entity to a closed environment from which it cannot exchange information with the outside environment if such information is not authorized by the confinement authority [18].

The idea of this protocol is to ask system questions with restricted answers. These safe questions are defined as ones for which human being could find an answer without any help from superintelligent artificial mind. For example, the researcher has two alternative approaches to try and knows that one of them is definitely correct, but doesn't know which one exactly. If the choice is wrong then he will lose time but will find a solution after some additional time. In this case, AI advice could be used.

Those types of questions are safe for several reasons. The main reason is the answer may be obtained even without assistance from AI and the second reason is that the answer could be encoded in a single bit making AI hard to hide additional information in it. Safe questions should be developed by the group of specially trained experts in different fields like computer science (especially in recent developments of AI), psychology, ethics, AI security and some others.

IV. AI confinement problem

The idea of confinement roots back from 1973 when Butler Lampson has defined the laws of confinement [12]:

- ✓ Total isolation: a confined program shall make no calls on any other program.
- ✓ Transitivity: if a confined program calls another program that is not trusted, this program should be also confined.
- ✓ Masking: a program to be confined must allow its caller to determine all its inputs into legitimate and covert channels.
- ✓ Enforcement: the supervisor must ensure that a confined program's input to covert channels conforms to the caller's specifications.

The confinement problem has many questions to be answered. For example it will be very difficult to interact with such a system (or even impossible) or that clever machine eventually will escape its prison using different approaches which a human guard does not know. The researcher Eliezer Yudkowsky has made so-called AI-Box experiment where he was trying to demonstrate that even human-level intelligence is enough to escape from an AI-Box [13].

There are several possibilities to prevent superintelligence from escaping. The first thing proposed by Yampolkiy is to isolate humans from confined AI completely (no contacts via voice, Morse code, text messages or any other channels) and delegate all maintenance operations to unintelligent software that has no possibilities to release AI from the box.

The second measure to be taken is to allow AI

functioning on limited and randomly changed resources without any access to its configuration. The AI should be reset regularly where the previous versions should replace current one. Using a virtual machine could also help.

The last group of measures is actual design of confinement facility. It should have self-destruction function; it should be located in isolated, hard-to-reach place guarded to prevent unauthorized access.

V. Explainability of Intelligent Systems

Very important aspect of safe AI is explainability, when a user should clearly and fully understand the output of an AI system and make corrections if necessary.

The concept of explainable AI (XAI) is shown if Fig. 1, which clearly shows that today the user usually does not know the answers to the following questions:

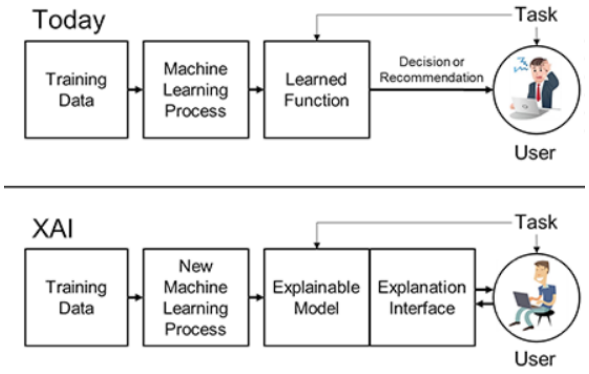


Fig. 1. Explainable AI [6].

why computer do it; why not something else; when it succeeded; when to trust a computer and how to correct errors. In the case of XAI the user will have answers to all of these questions.

In general, The Explainable AI (XAI) program aims to create a suite of machine learning techniques that [6]:

- ✓ Produce more explainable models, while maintaining a high level of learning performance (prediction accuracy);
- ✓ Enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners.

Andres Holzinger in his paper gives an approach of a complete machine learning pipeline beyond algorithm development [9], the authors will show only relevant to safe AI issues:

- ✓ Data: pre-processing, integration, mapping and fusion – understanding the physical aspects of raw data and its surroundings, especially i the application domain; ensuring quality of data.
- ✓ Learning algorithms: all aspects of design, development, experimentation, testing and evaluation.
- ✓ Visualization of data and analysis: presentation of multidimensional data in a human-friendly form.
- ✓ Privacy: data protection, safety and security.
- ✓ Entropy: used as a measure of uncertainty in data.

Wojciech Samek gives provides several reasons why explainability is so important for AI research and its safety aspects [15]:

- ✓ Verification of the system: no one should trust artificial system by default. In this case, verification procedure allows testing the AI "black box" behaviour and outputs using different solutions already available.
- ✓ Improvement of the system: Before improving the system, we should understand its weaknesses and the better we do it the better we can improve them.
- ✓ Learning from the system: nowadays AI systems are often using the big data of millions examples which human mind cannot deal with. That is why explainable AI should have extracted knowledge in a human understandable manner.
- ✓ Compliance to legislation: if a system gives wrong answer in a critical data domain, the responsibility should be preserved according to legislation issues. The users affected by such AI decision will want to know why the system decided that way. Only explainable system will give the answer.

There are several methods for making AI explainable: sensitivity analysis (SA), Layer-Wise relevance propagation (LRP) and others.

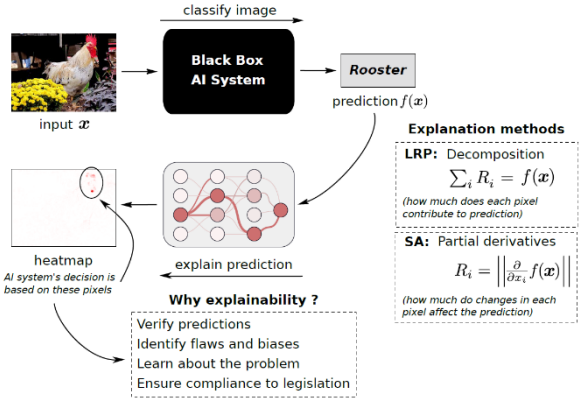


Fig. 2. Explanation of AI system prediction [15]

The fig.2 shows that system has correctly classified an input as a "rooster". In order to understand this decision, explanation methods such as SA or LRP are applied. The result of this explanation is an image, which visualizes the importance of each pixel for the prediction. In this example, the rooster's red comb and wattle are the basis for the AI system's decision. Using such a methodology user can be sure that system works as it intended to be.

VI. Future Research Directions

The paper shows that the research made in the direction of ensuring safety of artificial superintelligence is in its early stages. All the methods are limited and do not assure the complete confidence of AI user that this technology

will have only positive effects.

All possible solutions to ensure safe AI should be taken into account and carefully analysed and any small chance of improvement should be tested and implemented in a prototype.

Safe AI is the area of several interconnected science directions like computer science, psychology, mathematics, physics, philosophy, linguistics, biology and many more. That is why only a team of cross-trained researchers could deal with such a problem and expect to have some positive results.

The author of this article wants to attract attention of researchers from all fields to this topic, because the creation of superintelligent machine is only a matter of time and we should be ready and know what to do.

The further research directions will be development of experimental testing framework for different AI systems to check some of the above-mentioned methods of ensuring the safety of artificial superintelligence.

References

[1] R. Banham, Cybersecurity: Protective Measures Treasuries Should Be Taking. Treasury & Risk. 2018 Special Report, pp. 2-7.

[2] D.Beskow, K.Carley, Social Cybersecurity: An Emerging National Security Requirement. Military Review. April 2019, Vol. 99 Issue 2, pp. 117-127. 11p.

[3] N. Bostrom, Global Catastrophic Risks. Oxford: Oxford University Press, 2007.

[4] N. Bostrom. The ethics of artificial intelligence. Cambridge Handbook of Artificial Intelligence, 2011. [Online]. Available: https://nickbostrom.com/ethics/artificial-intelligence.pdf [Accessed: March. 03, 2019].

[5] R. Deibert, Toward a Human-Centric Approach to Cybersecurity. Ethics & International Affairs. Winter 2018, Vol. 32 Issue 4, pp. 411-424.

[6] D. Gunning, Explainable Artificial Intelligence, DARPA project, 2018. [Online]. Available: https://www.darpa.mil/program/explainable-artificial-intelligence [Accessed: March. 07, 2019].

[7] S. Hawking, Science in the next millenium, 1998. [Online]. Available:https://www.learnoutloud.com/Catalog/Science/Physics/Science-in-the-Next-Millennium/45223 [Accessed: March. 03, 2019].

[8] N.Hennig, Privacy and Security Online: Best Practices for Cybersecurity. Library Technology Reports. April 2018, Vol. 54 Issue 3, pp. 1-37.

[9] A. Holzinger, "From Machine Learning to Explainable AI." World Symposium on Digital Intelligence for Systems and Machines August 2018. [Online]. Available: https://www.researchgate.net/publication/328309811_From_Machine_Learning_to_Explainable_AI [Accessed: Feb. 21, 2019].

[10] M. Kiss, C. Muha, the cybersecurity capability aspects of smart government and industry 4.0 programmes. Interdisciplinary Description of Complex Systems. 2018, Vol. 16 Issue 3-A, pp. 313-319.

[11] R. Kurzweil, The Singularity Is Near: When Humans Transcend Biology. New York, NY: Viking, 2006.

[12] B. Lampson, A Note on the Confinement Problem, 1973. [Online]. Available: https://www.cs.utexas.edu/~shmat/courses/cs380s_fall09/lampson73.pdf [Accessed: March. 19, 2019]

[13] Open AI project. [Online]. Available: https://openai.com/ [Accessed: March. 11, 2019].

[14] N. Sales, Privatizing Cybersecurity. UCLA Law Review. April 2018, Vol. 65 Issue 3, pp. 620-688. 69p.

[15] W. Samek, T. Wegang, K. Muller. Explainable artificial intelligence: understanding, Visualizing and interpreting deep learning models, Aug. 28, 2017. [Online]. Available: https://arxiv.org/abs/1708.08296 [Accessed: March. 19, 2019].

[16] M.Scala, A. Reilly, Risk and the Five Hard Problems of Cybersecurity. Risk Analysis: An Official Publication Of The Society For Risk Analysis, March 2019, pp. 32-37.

[17] A.T.Sherman, D. DeLatte, M.Neary etc., Cybersecurity: Exploring core concepts through six scenarios. Cryptologia. July 2018, Vol. 42 Issue 4, p337-377. 41p.

[18] R. Yampolskiy, "Leakproofing the Singularity Artificial Intelligence Confinement Problem," Journal of Consciousness Studies, vol. 19, pp. 1-2, 2012.

[19] R. Yampolskiy, Artificial Superintelligence: a Futuristic Approach. New York: Chapman and Hall/CRC, 2015.

[20] E. Yudkowsky, The AI-Box experiment. [Online]. Available: http://yudkowsky.net/singularity/aibox/ [Accessed: March. 19, 2019].

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 184-188

Analysis and Evaluation of Fiscal Policy in Latvia

Ērika Žubule
Faculty of Economics and Management
Rezekne Academy of Technologies
Rezekne, Latvia
erika.zubule@rta.lv

Peter Grabusts
Faculty of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
peteris.grabusts@rta.lv

***Abstract*— The choice and topicality of the research topic is based on the fact that upon strengthening of the government's regulating role in economy, the notion of public finances positions itself, and state budget has become an important subject of both economic and political discussions as implementation of fiscal policy is takes place through it. In order to evaluate fiscal policy, it is necessary to evaluate the potential influence of different fiscal policy instruments on social and economic situation in the state. Aim of the research – to evaluate activities of fiscal policy implemented in Latvia in context of certain tax, namely, influence of corporate income tax on state's economic and financial indicators, identifying the main risks and imperfections of fiscal policy when ensuring state's budget. Applying simulation methods in the environment of Matlab/Simulink, the authors analyze and evaluate the influence of fiscal decisions and their implementation on the situation in Latvia, analyzing the most important tendencies in the sphere of corporate income tax payments according to the tax reform commenced in 2018.**

Keywords*— budget revenue, corporate income tax, fiscal policy, fiscal risk, gross domestic product.*

I. Introduction

The choice and topicality of the research topic is based on the fact that upon strengthening of the government's regulating role in economy, and thereby positioning the notion of public finances, state budget has become an important subject of both economic and political discussions. Consequently, the influence of government's fiscal implemented measures on the development within the state becomes topical, implementing the policy of taxes and government expenditure.

Therefore there is a view that in order to enable evaluation of fiscal policy, it is necessary to pay attention not only to the aims defined by the government activity, but also to evaluate the potential influence of different fiscal policy instruments on social and economic situation in the state.

Aim of the research – to evaluate activities of fiscal policy implemented in Latvia in context of certain tax, namely, influence of corporate income tax on state's economic and financial indicators, identifying the main risks and imperfections of fiscal policy when ensuring state's budget, and to suggest recommendations for improving the capacity of fiscal measures.

Applying simulation methods Matlab/Simulink environment, the authors in the research, on the basis of the study carried out before [1], evaluate the influence of fiscal decisions and their implementation on the situation in Latvia, analyzing the most important tendencies in the sphere of corporate income tax payments according to the tax reform commenced in 2018.

Methods used in the research – monographic, logical and constructive, content analysis, comparative method for views and conceptions as well as situation simulation method.

II. Literature review

In the literature on economics and in the financial practice in the countries the conception "state fiscal risk" has only recently been topical in connection with the beginning of the economic crisis in the USA in 2007 and later on also in the European countries. The crisis reached Latvia, too, which negatively influenced the process of the formation of the state budget increasing the state budget deficit and the government debt. As we know, in the period of 2009 until 2011 there was a very painful consolidation process of budget in Latvia. For the development of financial and economic theory and practice the role of fiscal policy measures (taxes, government's expenses, loan) is emphasized for potential opportunities of governments to regulate social and economic activities in country.

Fiscal policy plays an important role in public economy. It is a tool for government to impact the market of goods and services, financial and labour market as well as solve the problems of redistribution of income and social justice [4].

It is emphasized that when evaluating application of instruments of fiscal policy it is important to understand the cycle of economic development – economic growth, high level of demand, supply and general wellbeing of population interchanges with periods, when growth rates decline and so do the income of the employed, having

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4076
© 2019 Ērika Žubule, Peter Grabusts.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

a negative impact on common demand for goods and services. Thus a government can purposefully intervene in economic processes using instruments of fiscal policy in order to renew economic stability and ensure sustainable development in a long term [5], [6].

Fiscal policy is implemented through national budget, namely, receipts, expenditures planning and ensuring for performance of governmental functions. Despite the persisting view that budgeting is related to mathematical methods and ability to work with numbers, nevertheless it is emphasized that all the decisions on development of national financial plan should be perceived not only in the economic, but also political context [7].

It is possible to assert, that it is updated also in the legislation of the Republic of Latvia: budget – a mean for implementation of national policy through financial methods. Consequently, national budget is the most important instrument in implementation of fiscal policy; basically it is a financial plan of income and expenditure for a certain period of time developed by the executive power and approved in the procedure of legislation [8].

Success and failures of each country's economic development depend on how successful was the conception chosen by the government activity, namely, whether it corresponds to the specific time and economic conditions (social, economic, political, geopolitical factor).

It is possible to state, that fiscal policy measures implemented in good time and corresponding to economic situation is a crucial precondition in order to achieve a sustainable development of national economy.

III. Materials and methods

The methodology of the research is related to the aim set and tasks to be accomplished, that is, to analyse the research published in scientific literature on fiscal policy, its feasibility. It means, that important how can to find correlation with economic development, income in the budget and concrete tax payments. The authors have used the goal to catch up with Matlab/Simulink tool for economics process modelling.

Simulation - the most powerful and versatile method for studying and evaluating the effectiveness of systems, the behavior of which depends on the influence of random factors [8].

The implementation of such opportunities in the universal programming language is a difficult task. Currently, there is quite a lot of software that allows to model processes. However, now there is a product that allows solving these problems quite effectively – the MATLAB package [9], [10], [11] containing a visual simulation tool – Simulink. Simulink - a tool that allows you to simulate the system quickly, get the indexes of expected effect and compare them with the amount of effort required to achieve them [12], [13].

Of particular interest for simulation is a Simulink tool designed specifically for modelling dynamical systems. It has a library of standard graphics units with built-in mathematical functions. It is sometimes called a tool of visual modelling [14], [15].

Although Simulink is designed mainly to solve engineering and scientific- technical problems, the possibilities of its use are almost unlimited. The input of initial parameters is made interactively by graphics assembly of elementary blocks circuit diagram, resulting in a model of the studied system. The blocks included in the model relate to each other both in information and in management. The type of connection depends on the type of the block and the logic of the model.

The Simulink program is an application to the MATLAB package. Using Simulink program the simulation implements the principle of visual programming whereby the user on the screen creates a model of a structure, process or system from standard blocks of the library, and performs calculations. In this case, unlike in classical ways of modelling, the user does not need to study the programming language and numerical mathematics methods thoroughly, there is enough to have some general knowledge required when working on the computer, and, of course, knowledge on the subject area in which one works.

Creating a model in this way, and then launching it, it is possible to see the results of modeling. In the simulation, the user can choose the method for solving equations, as well as the way to change the model time (with a fixed or variable step). During the simulation, it is possible to monitor the processes happening in the system. To this purpose, special viewing devices that are part of the Simulink library are used. The simulation results can be presented in the form of graphs or tables.

IV. RESEARCH RESULTS

The tax reform commenced in Latvia in 2018 was basically grounded in issues discussed in the society: reduction of employment taxes, adoption of Estonian practice in relation to corporate income tax, changes of capital income imposition procedure etc. While evaluating the solutions offered by the government, the authors stated that nevertheless there might arise issues on efficiency of fiscal measures and their influence not only on budget income, but also on social and economic aspects.

The authors chose corporate income tax as the most important research object in the sphere of tax changes, performing evaluation of its influence on GDP and national budget before and after implementation of tax reform. The authors carried out analysis of statistical data, comparing the current situation and prognosticating a potential development scenario of public finances for the following years (see Table I).

The expected and factual GDP indicators are illustrated in Figure 1.

It can be concluded that the expected GDP development trends show an increase. However, calculations of the authors indicate that there is much more rapid GDP growth to be planned. It means that when making decisions in the

context of fiscal policy establishment one must not ignore the potential scenarios of macroeconomic development, taking into account various influencing factors.

It must be emphasized that it is planned to submit the national draft budget 2019 to *Saeima* in the beginning of March, and it could come into effect after the middle of April. Consequently, there was made a decision, that it is not expedient to develop medium-term budget framework for this year, as macroeconomic prognoses of the previous year would be used for calculations in this case. Moreover, in this autumn a medium-term budget framework will have to be developed already for the next three-year period [16].

TABLE I. GDP, CIT AND THE GOVERNMENT BASIC BUDGET REVENUES IN LATVIA IN THE YEARS 2014-2022

Year	**2014**	**2015**	**2016**	**2017 (progno-sis)**	**2018 (progno-sis)**	**2019 (prognosis)**	**2020 (prognosis)**	**2021 (progno-sis)**	**2022 (prognosis)**
GDP	23600	24378	25072	26403	27905	29615	30360		
GDP Factually				27033	29424	31293	33663	34950	36861
Budget	4939	5093	5281	5760	6336	6256	6090		
Budget Factually				5387	6324				
CIT	354	383	407	425	461	468	498		
CIT Factually			419	425	304				

(Data from dark cells- according to predictable data of the Ministry Economy of Latvia -On the medium-term budgetary framework for 2017, 2018 and 2019 years).

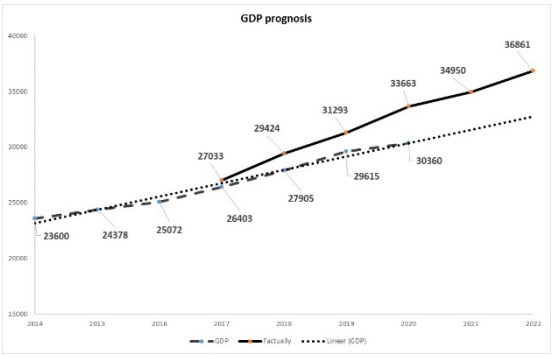


Fig. 1. Factual and expected GDP values in Latvia during 2014-2022 (X-axis: year; Y-axis: GDP values).

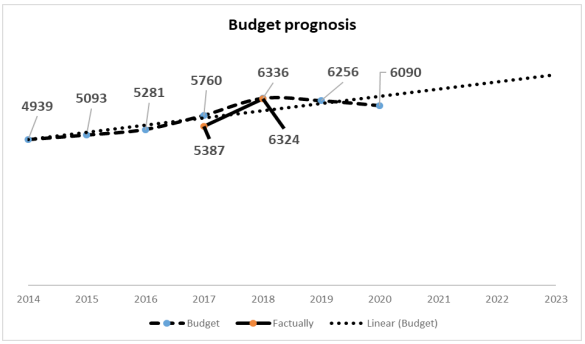


Fig. 2. Factual and expected values of Latvian national basic budget (X-axis: year; Y-axis: Budget values).

It must be emphasized that it is planned to submit the national draft budget 2019 to *Saeima* in the beginning of March, and it could come into effect after the middle of April.

As GDP indicator ensures the component of revenue for the national budget, also the numeric values of potential revenue for budget correlates according to prognoses of GDP indicators (see Fig. 2).

Trend line (dotted) shows that revenue for the budget should increase until 2023. Nevertheless, the authors believe that the situation in ensuring Latvian fiscal policy depends on several factors: political decisions in the country, geopolitical aspects, economic and social situation in general.

The authors accentuate corporate income tax as the most significant research object in the sphere of taxes. Consequently there is performed evaluation of its influence: development trends, influence on national economic indicators, mostly GDP (see Fig. 3).

It should be emphasized that CIT prognosis also shows an upward trend; however, currently the actual values do not correspond to the expected ones.

It is possible to conclude that in the framework of tax reform the major changes happened in the sphere of corporate income tax, stipulating the tax to be paid upon division of profit instead of imposing it on the derived profit as it was before. It also influenced the indicators of revenues for the national budget for 2018. In 2018, corporate income tax revenue was 304.00 million euro. Comparing with 2017, it decreased by 121.61 million euro or 28.6 %. However, the forecast of revenue for corporate income tax has been reached at 130.6 %, in other words, revenue is 71.30 million greater than it was planned [17].

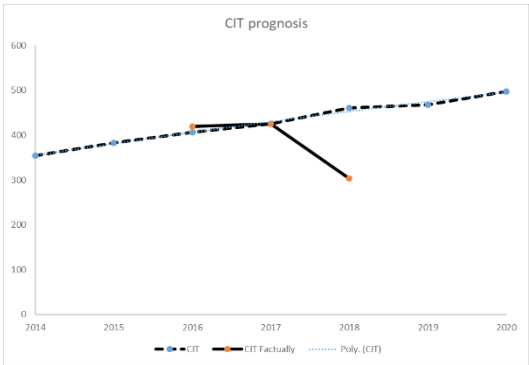


Fig.3. The expected and factual values of Corporate income tax (X-axis: year; Y-axis: CIT values).

In order to make up the simulation of CIT influence on the economy and the fiscal situation in the country, a simple Matlab/Simulink model was derived which calculated the proportion CIT/GDP (see Table II and Fig. 4).

TABLE II. The evaluation criteria of the CIT influence on the GDP

Year	2014	2015	2016	2017	2018	2019	2020
CIT/ GDP	1,50	1,57	1,67	1,58	1,03	1,50	1,48

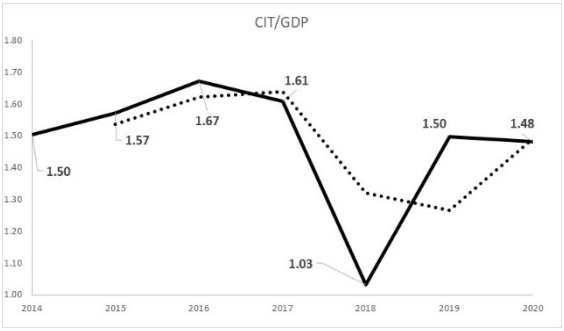


Fig. 4. The influence of CIT on GDP according to the existing legislation of the Republic of Latvia (X-axis: year; Y-axis: CIT influence on the GDP values).

The dotted trend line shows that according to the forecast there should not have been such a rapid decline of the proportion of CIT/GDP in 2018; however, as emphasized before, it is related to the aspects of tax reform implementation in the context of CIT application.

Despite the fact that further forecasts show a positive trend, the authors' opinion is that Latvian authorities (Ministry of Finance, Ministry of Economics) when implementing fiscal policy have to ensure high quality capacity in the implementation of the initiated reform (analytical evaluation, understanding the real situation, ensuring proper tax administration).

With the aim to make up the simulation of the CIT influence on the economy and fiscal situation in the country a simple Matlab/Simulink model was elaborated which calculated the proportion CIT/GDP (see Fig. 5).

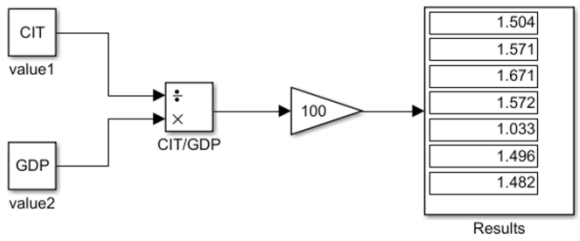


Fig. 5. The evaluation model of the influence of the CIT.

V. Conclusion and proposals

Fiscal policy plays an important role in public economy. It is a tool for government to impact the market of goods and services, financial and labour market as well as solve the problems of redistribution of income and social justice.

Government can purposefully intervene in economic processes using instruments of fiscal policy (tax, expenditures, loan), in order to renew economic stability and ensure sustainable development in a long term. It is possible to state, that fiscal policy measures implemented in good time and corresponding to economic situation is a crucial precondition in order to achieve a sustainable development of national economy.

National budget is the most important instrument in implementation of fiscal policy; basically it is a financial plan of income and expenditure for a certain period of time developed by the executive power and approved in the procedure of legislation.

Simulation - the most powerful and versatile method for studying and evaluating the effectiveness of systems, the behavior of which depends on the influence of random factors. It can be concluded that Matlab/Simulink tool is a very suitable tool not only in economics calculations, but also can serve as a simulation model visualization tool in various science fields.

In Latvia according to the forecast there should not have been such a rapid decline of the proportion of CIT/GDP in 2018; it is related to the aspects of tax reform implementation in the context of CIT application.

When making decisions in the context of fiscal policy establishment one must not ignore the potential scenarios of macroeconomic development, taking into account various influencing factors (political decisions in the country, geopolitical aspects, economic and social situation in general).

Despite the fact that further forecasts show a positive trend, the authors' opinion is that Latvian authorities (Ministry of Finance, Ministry of Economics, SRS) when implementing fiscal policy have to ensure high quality capacity in the implementation of the initiated reform.

References

[1] Ē.Žubule, P.Grabusts, L.Kavale, "Creation of Innovative Approach for Study of the National Fiscal Space, " Proceedings of the 11th International Conference "Environment.Technology. Resources", Rezekne,June,15-17, 2017, Vol.2., P.189-195. http://dx.doi.org/10.17770/etr2017vol2.2574

[2] J.M. Keynes, The General Theory of Employment, Interest, and Money. Routledge, Taylor & Francis Group, 2017.

[3] Macroprudential Policy and Practice. In P. Mizen, M. Rubio, & P. Turner (Eds.), Macroprudential Policy and Practice (Macroeconomic Policy Making, p. I). Cambridge: Cambridge University Press, 2018.
https://doi.org/10.1017/9781108304429

[4] J.L. Mikesell, Fiscal Administration: Analysis and Applications for the Public Sector, Belmont, US, 2007.

[5] R. Allen, R. Hemming, and B.H. Potter, The International Handbook of Public Financial Management. Palgrave Macmillan, 2013.

[6] A. Alesina and S. Ardagna, "Large Changes in Fiscal Policy: Taxes Versus Spending. Taxpolicy and the Economy, " 24(1), 35-68, 2010.

[7] Public Budgeting: Policy, Process and Politics, ed.by Rubin.I.S., ASPA, 2008.

[8] C. Kay, Mathematics for Computer Programmers. New Jersey: Prentice Hall, 1984.

[9] J. Kiusalaas, Numerical Methods in Engineering with MATLAB, 3e. Cambridge University Press, 2016, 430 p.

[10] T. Siauw and A. Bayern, An Introduction to MATLAB Programming and Numerical Methods for Engineers. Academic Press, 2014, 340 p.

[11] D. Smith, Engineering Computation with MATLAB, 3e. Pearson Education Inc, 2013, 464 p.

[12] D. Xue and Y. Chen, System Simulation Techniques with MATLAB and Simulink. John Wiley & Sons, Inc, 2013, 488 p.

[13] P. Karel and Z. Tomas, "Multimedia Teaching Aid for Students of Basics of Control Theory in Matlab and Simulink," Procedia Engineering, Volume 100, 150–158, 2015.

[14] A.B. Shiflet and G.W. Shiflet, Introduction to Computational Science: Modeling and Simulation for the Sciences, 2e. Princeton University Press, 2014, 856 p.

[15] C.W. Silva, Modeling and Control of Engineering Systems. CRC Press, Inc, 2009, 766 p.

[16] "A medium-term draft state budget for 2019 will not be prepared," Feb. 7, 2019. [Online]. Available: http://www.saeima.lv/lv/aktualitates/saeimas-zinas/27647-videja-termina-valsts-budzeta-projektu-2019-gadam-negatavos [Accessed: March 10, 2019]. (in Latvian)

[17] "Summary of the implementation of the budget revenue share for the 12 months of 2018," Jan. 22, 2019. [Online]. Available: https://www.vid.gov.lv/lv/kopsavilkums-par-budzeta-ienemumu-dalas-izpildi-2018gada-12-menesos [Accessed: March 10, 2019]. (in Latvian)

Engineering Education

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 190-194

Smart Specialisation Concept Application in Universities: E-Business Online Studies Model Development

Aigars Andersons
Institute of Social, Economic and Humanities Research,
Vidzeme University of Applied Sciences
Valmiera, Latvia
aigars.andersons@va.lv

Jozef Bushati
Faculty of Educational Sciences
University of Skodra,"Luigj Gurakuqi"
Shkodra, Albania
jozefbushati@gmail.com

***Abstract*—This paper discusses the Smart specialisation concept problems and challenges based on practice of Nordic and Baltic Higher Education Institutions. A particular focus of this research is devoted to the core understanding of the concept of Smart specialisation and its practical application for University level studies curriculum. EU Regional Cohesion Policy and other related EU policies play an important role to understand basic theoretical aspects for Smart specialisation. On the other hand, when specific case studies had been analysed in practice it is possible to see enormous variety of different misunderstandings of the matter. Even different study directions in the same University translate the Smart specialization terms mostly related to their individual study environments. As a result, it is possible to find in practice very different "Smart specialisation" concepts for technology, engineering, economics, IT and other study directions. As the research methodology/approach, the authors analyse the outcomes and challenges faced by different sets of case studies in Baltic and Nordic countries. Using EU, World Bank and other official national documents and publications, the authors explain the ways in which the Smart specialisation concept can be used in online learning environments. As an empirical research evidence there are analysed the outcomes from case studies - practical implementation of an online e-business study course jointly developed by 9 universities from Latvia, Finland, Iceland, Norway, Estonia, Lithuania, Sweden and Denmark supported by e-business field experts from Balkan area universities. Because of still existing gaps in Smart specialisation theoretical framework, the "grounded theory" approach have to be used more intensively when new and advanced theories have been created from practical outcomes of specific case studies. To solve a problem of Smart specialisation concept different translations and applications, authors of this paper propose to construct the teaching methodology of Smart specialisation concept into two consecutive levels - primary, as a core theory unique for all educators and secondary, as specific applied theories for each specific study direction.**

***Keywords*—E-business, online studies, Smart specialization, teaching methodology.**

I. Introduction

This research paper analyses and provides few practical evidences valid for a European Parliament Preparatory Action centered on the refinement and implementation of the Research and Innovation Smart specialisation Strategy (RIS3).

Smart specialisation is an innovative approach that aims to boost growth and jobs in Europe, by enabling each region to identify and to develop its own competitive advantages. Through its partnership and bottom-up approach, Smart specialisation brings together local authorities, academia, business spheres and the civil society, working for the implementation of long-term growth strategies supported by European Union (EU) funds [1].

The process how to define and apply term "Smart specialisation" in Universities requires practical experimentation and empirical analysis that tests theoretical and learning aspects required these specialisations to be efficient and sustainable.

Guidelines from European Commission emphasizes that in "Smart specialisation" facilitation process the governments should have a central role [2], but how the government really performs these functions still remains uncertain, because it is a role that is adopted from various practical applications related with mentioned above concept nor coherent theories or mandatory regulations.

With an aim of the research and to test in practice different aspects of "Smart specialisation" concept for teaching and application purposes, nine Higher Education Institutions (HEI) from Baltic and Nordic areas Latvia joined their academic and scientific capacities to establish the collaborative network NOBANET (https://www.nordicbalticnet.info/).

List of involved HEIs:

1) Latvia - Vidzeme University of Applied Sciences, Valmiera,
2) Estonia - Estonian Entrepreneurship University of Applied Sciences, Tallinn and TTK University of Applied Sciences, Tallinn),

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4115
© 2019 Aigars Andersons, Jozef Bushati
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

3) Lithuania - Kauno Kolegija/ University of Applied Sciences, Kaunas,
4) Finland - Arcada University of Applied Sciences, Helsinki,
5) Iceland - Akureyri University/School of Business and Science, Akureyri,
6) Sweden - School of Education and Communication at Jönköping University,
7) Denmark - UCL University College Denmark, Odense,
8) Norway - Westerdals Oslo School of Arts, Communication and Technology.

In this paper had been examined outcomes from the partnership of mentioned above HEIs in creation of new type methodologies in a form of self-constructed online e-business study course with applied in practice "Smart specialisation" theories.

II. Materials and Methods

A. *Theoretical analysis of concept "Smart specialisation".*

"Smart specialisation" is actively discussed topic during the last decade by scientists and practitioners because of its special role into the regional development policy of European Union. One of the first published definitions states, "Smart specialisation is largely about the policy process to select and prioritize fields or areas where a cluster of activities should be developed" [3]. Evidently, the scope of this particular definition is too broad and unspecified for application in the same way for all specific business cases.

During the last years, this concept emerged into the more specified definition for Regional and Innovation Strategies for Smart specialisations (RIS3):

RIS3 is an integrated, place-based economic transformation agenda that does five important things:

a) It focuses policy support and investments on key national/regional priorities, challenges and needs for knowledge-based development, including ICT-related measures,

b) It builds on each country/region's strengths, competitive advantages and potential for excellence,

c) It supports technological as well as practice-based innovation aiming to stimulate private investment,

d) It gets stakeholders fully involved and encourages innovation and experimentation,

e) It is evidence-based and includes sound monitoring and evaluation systems [4].

The European Commission has developed a Smart specialisation platform where information about the RIS3 "Assessment wheel" is published. It is an assessment tool for RIS3 development in a country or [5].

For better understanding of "Smart specialisation" concept and further theory analysis framework there are two citations below from Foray et al. research paper. The first states that a key point is that Smart specialisation is not just for the "best" regions and technology leaders. On the contrary, this concept provides strategies and roles for any region. Indeed, the concept is built on the fact that there is not only one game in town in terms of R&D and innovation; i.e. there are many other kinds of productive and potentially beneficial activities apart from the invention of fundamental knowledge needed for the development of general purpose technologies and tools [6]. The another one states that "Smart specialisation", on the other hand, involves the discovery of what makes a local knowledge base original and unique. Therefore, a commitment to "Smart specialisation" strategies can promote greater diversity of areas of knowledge and expertise within the system, thereby rendering the entire economy more able to enjoy the benefits of distinct local agglomeration economies and less vulnerable to both supply and demand shocks emanating in global markets [6].

After analysis of these two definitions, it is possible to draw a conclusion that nowadays "Smart specialisation" from HEIs academic prospective is less theoretical but mostly practical study subject dependent and morphing from one specific case and/or application to another.

B. *The Framework Development and Strategies of Smart specialisation.*

Authors of this research mainly agree with ideas proposed by Romanian scientist Steliana Sandu that Smart specialisation has already become a general strategy, with commitments from major decision-making individuals to instate them into policies and specific strategies. For the moment, practice is ahead of theory, meaning that in Europe and other countries strategies of smart growth or platforms of Smart specialisation elaborated from business and public governance practical activities. It is necessary to refine the theory and to choose the adequate database that would help to clarify some concepts and perfect the operationalization of this concept, to evaluate the possibilities of implementation and the means to monitor and coordinate it [7].

The Smart specialisation concept rests on two fundamental ideas:

1) Regions should focus on few areas where they have a significant impact (specialisation),

2) These domains of specialization should make use of existing strengths (smart) such as location, resources, or Science Technology and Innovation (STI) capabilities.

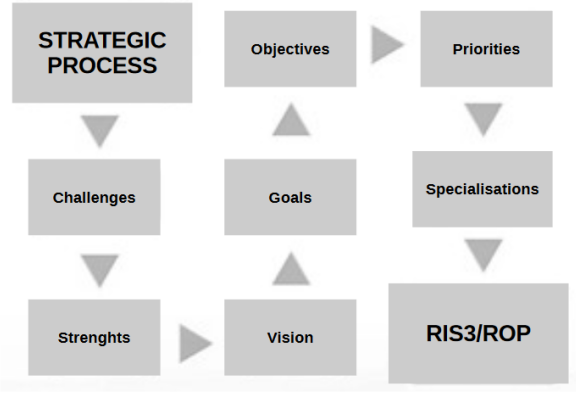


Fig. 1. Proposed key steps in the strategic process of selecting Smart specialisations [8], [9].

A significant difference between smart specialization and other policies is the active role of entrepreneurs in discovering new domains for innovation [10].

Clarification of Smart specialisation process is a key factor for further development of modern teaching methodologies for Universities. An existing situation how to select Smart specialisations for particular locations displayed in Fig. 1 (see above).

This specific roadmap scheme initially had been used also by NOBANET collaborative academic network joining nine Universities from Baltic and Nordic countries as a basic platform for research in Smart specialisation area described later in the next chapter of this paper.

The development of Smart specialisation strategies (RIS3) should promote the competitiveness of countries or regions. For national level, RIS3 helps concentrate resources and support innovations. Desk research identifies lack of regional level when comparing Smart specialisations in different countries [11].

The most frequently mentioned reason for gap of competences needed for Smart specialisation occurrence are inter-sectoral structural transformations. Failing branches do not have sufficient number of jobs available for employees with low qualifications. Jobs appear in emerging industries, but in this case, there are no qualified employees capable of meeting the requirements of jobs, most often related to modern technologies. The competence gap may be deepened by the maladjustment of curricula to the needs of the labor market [12].

An existing gap in understanding of Smart specialisation general framework among various theories and practical applications must be reduced and transformed into "smarter" structures. It is possible to realize collecting of real evidences of case studies and practical applications through use of "grounded theory" approach.

III. Results and Discussion

A. Evidence based approach for Smart specialisation concept application.

The overall aim of NOBANET network is to create and widely disseminate new knowledge on successful internationalization of Nordic and Baltic SMEs. This aim achieved through close cooperation between HEIs and Small Medium Size Enterprises (SME) in all Nordic and Baltic countries with assistance of business entities and experts all over the world. Faculty, students and companies work closely together and across national borders, within educational courses and through real-life assignments in companies.

The five specific objectives of NOBANET network are:

1) Developing new learning materials on internationalization of SMEs,
2) Implementation real-life projects on internationalization,
3) Developing partnership HEIs curricula to include internationalization of SMEs,
4) Creating models for sustainable cooperation between HEIs and SMEs,
5) Compilation of handbook: Internationalization of Nordic and Baltic SMEs.

Previously developed (in Years 2014-2016) materials by NOBANET include:

1) Handbook on the Socio-Economic Environment of the Nordic and Baltic countries;
2) Handbook on the SMEs and Foreign Trade;
3) Handbook on the University – Business cooperation;
4) Best Practice examples on internationalization of SMEs;
5) Models for successful cooperation between HEIs and SMEs.

This research paper mainly covers two consecutive stages of NOBANET activities connected with Smart specialisation matters. The reasoning for actuality of particular research comes from NOBANET "Handbook on Internationalisation of Small and Medium-Sized Enterprises", Part III, University – Business cooperation: model development and experience from practice [13]. The specific evidence-based case study covers Smart specialisation issues in ICT sector as detailed analysis of jointly developed e-business online study course within E-NOBANET network activity framework finished in Year 2018.

Smart specialisation studies have to contribute to the development of an evidence-based approach to Smart specialisation policies. It is a policy that builds on the prioritisation of resources, aiming at building knowledge and accelerating structural change [14]. Evidence-based policy (EBP) and practice means integrating experience, expertise, and judgement with the best available external evidence from systematic research [15].

Traditionally, different translations of the same subjects are not unique. Nowadays there are a number of digital tools available to bridge this gap and to create common understanding of Smart specialization matter. When time comes for a legacy relational database to migrate to semantic web or to be integrated with it, an important issue of determining similarity (compatibility) between two data models expressed in different ways arises [9].

However, bringing an evidence to bear on policy and practice raises at least two fundamental issues. The first concerns the type of evidence required to address policy needs, while the second refers to a set of challenges at the science–policy interface, in particular the question of how best to convey evidence to policy practitioners and practice [16]. The transition towards EBP has been related to the need for a sound basis for better regulation [17] and for better integration of scientific and academic theories and methodologies for the formulation, implementation, and evaluation purposes of Smart specializations.

B. E-NOBANET case study.

In lights of Smart specialisation teaching methodologies development process for students and SME`s specialists had been found that all Baltic and Nordic countries have some common areas of special concern. In the age of rapidly growing globalisation process, local SMEs are urgently lacking appropriate ICT competences. As an outcome from enabled problem, the NOBANET collaborative network had been established an innovative

approach for Smart specialisation in ICT sector (inter-regional instead of traditional local or regional) by practical development of joint e-business course together with representatives of HEIs from all Nordic and Baltic countries and invited experts from Balkan area Universities. The goal of this activity stated as providing of students and SMEs business representatives with state-of-art competences and knowledge about basic principles of e-business and to improve practical skills how to use online tools and applications.

After the primary data analysis by joint researchers group it was decided to use an innovative Smart specialisation approach for this online study course with full e-business course (10 ECTS in total) divided in 10 individual topics, at amount of 1 ETCS each. Thus, students can take any number of topics they like to learn needed e-business competences in depth. This advanced approach improved existing methodologies with a possibility to avoid overlapping of mandatory studies of subjects already mastered elsewhere.

After the successful completion of the full study course, students will get an advanced knowledge and practical skills how to implement e-business applications in small and medium size enterprises and organizations, as well as the basic understanding how to create a simple/ average complexity web shop (shopping card) online.

TABLE 1

E-business online study course content [18].

E-business online study course topic	Responsible developer	Enrolled students
Introduction of e-business trends and tools	Latvia	116
Marketing strategies and market analysis	Iceland	2
Layout, usability and user experience	Sweden	27
Content and SEO	Finland	85
Online payment systems	Latvia, Lithuania	8
Delivery and distribution options	Estonia	29
Delivery terms, legal documents and data security	Estonia, Denmark	23
Delivery analytics and KPI's	Estonia, Finland	2
Test and evaluation	Finland, Norway	3
Introduction into e-business course	Finland	16

E-learning course materials later were integrated in regular courses at NOBANET network partner universities and they are available (upon registration) on the Eliademy LMS platform [18]. The e-business study course is available as a self-paced, self-learning course with support and guidelines provided by the group of instructors. In addition, all developed e-learning materials are freely available for other similar projects and relevant topics.

IV. Conclusions

The most of the interviews with entrepreneurs during the NOBANET collaborative academic network research activities enables that many SMEs owners and managers seem to have handled most things by themselves, through the individual creation of their own partnerships and relations. Just a few reported that they had gotten help from either governmental support or other kinds of business professional organizations like Business Finland, Finnvera, Finnish Health Technology Association, Business Sweden, Innovation Norway, Start-up Sauna Finland, Latvia Innovation and Development Agency and/or similar.

Until now, the vast majority of entrepreneurs, administrators and even university teachers understand concept of Smart specialisation in very different way in accordance with their individual experience and specific applications. Authors of this paper propose to divide Smart specialisation theory into two consecutive levels - a **primary core theory** unique for all and a **secondary applied theories** specific for each individual study direction.

From interviewed ICT sector representatives' prospective localised Smart specialisation principle is inadequate tool for rapidly changing global business markets. Nowadays **businesses requires going beyond local or regional borders**, without any limitations with full access and rights to participate in the global supply chains and sales networks.

Too much focus at regional level on few specific areas with a significant impact and intensive resources concentration only by **exploiting of non-unique** existing **strengths** such as geographical location, regional availability of natural and human resources or local ICT sector capabilities **can significantly reduce global competitiveness for less developed regions** where predominantly concentrated industries with low additional value.

Some **University level business administration study programs need updates with core theoretical aspects of Smart specialisation concept** with purpose to apply them into specific practical entrepreneurship models.

Smart specialization development process always **have to be supplemented by various horizontal support measures**, like collaborative networks, business advisory services, accessible and appropriate education systems and products sales promotion tools or platforms. Analysed in this research case study with free elective multi-topic e-business online study course created for students and SMEs specialists by NOBANET is good practice example here.

Introduction of Smart specialization concept in general is mostly **valid for businesses built on use of local natural resources but less valid for businesses built on use of highly qualified human resources or sophisticated technologies**. The last ones nowadays easily go beyond any local or regional boundaries and limits there are no need of "artificial" local or regional "Smart specialization".

Because of still existing gaps in Smart specialisation

theoretical framework, the **"grounded theory" approach have to be used more intensively** when new and advanced theories have been created from practical outcomes of specific case studies.

References

[1] European Commission, "Smart specialisation: Strengthening Innovation in Europe's Regions, Smart specialisation Platform, 2017, [Online]. Available: http://ec.europa.eu/regional_policy/sources/docgener/guides/smart_spec/strength_innov_regions_en.pdf [Accessed February 11, 2019].

[2] European Comission, "Europe 2020 Flagship Initiative Innovation Union," Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Brussels, 2014, p. 20, [Online]. Available: http://ww.w.xploit-eu.com/pdfs/Europe%202020%20Flagship%20Initiative%20INNOVATION.pdf [Accessed February 24, 2019].

[3] D. Foray, P. A. David, and B. Hall, "Smart specialisation–the concept," Knowledge economists policy brief, 9(85), June 2009, p. 100.

[4] European Commission, "Guide to Research and Innovation Strategies for Smart specialisations (RIS 3), Regional Policy, Brussels, 2012.

[5] S. Gemma and Z. Bulderberga, "Smart specialisation strategy assessment in Baltic States," In Research for Rural Development: International Scientific Conference Proceedings (Latvia), Latvia University of Agriculture, 2017. http://dx.doi.org/10.22616/rrd.23.2017.060

[6] D. Foray, P. A. David, and B. Hall, "Smart specialisation from academic idea to political instrument, the surprising career of a concept and the difficulties involved in its implementation", (No. EPFL-WORKING-170252), EPFL, 2011.

[7] S. Sandu, "Smart specialization concept and the status of its implementation in Romania," Procedia Economics and Finance, 3, Januaru 1, 2012, pp. 236-242. http://dx.doi.org/10.1016/S2212-5671(12)00146-3

[8] M. Piatkowski, T. Szuba, and G. Wolszczak, "Review of national and regional research and innovation strategies for smart specialization (RIS3) in Poland," World Bank Group, Washington: DC, 2014.

[9] I. Zarembo, A. Teilans, K. Barghorn, Y. Merkuryev, and G. Berina, "Methodology for Similarity Assessment of Relational Data Models and Semantic Ontologies," In 2016 International Conference on Systems Informatics, Modelling and Simulation (SIMS), IEEE, June 1, 2016, pp. 119-123. http://dx.doi.org/10.1109/SIMS.2016.21

[10] S. M. Krammer, "Science, technology, and innovation for economic competitiveness: The role of smart specialization in less-developed countries," Technological Forecasting and Social Change, 123, October 01, 2017, pp. 95-107. http://dx.doi.org/10.1016/j.techfore.2017.06.028

[11] M. Pelse and M Lescevica, "Smart Specialization Assessment in Latvia," In Economic Science for Rural Development Conference Proceedings, 42, April 15, 2016.

[12] R. Oczkowska, S. Wiśniewska, and P. Lula, "Analysis of the Competence Gap among Vocational School Graduates in the Area of Smart Specialization in Poland," International Journal for Quality Research, 11(4), December 1, 2017.

[13] Nobanet, "Handbook on Internationalisation of Small and Medium-Sized Enterprises. Part III, University – Business cooperation: model development and experience from practice," Nobanet Portal, 2016, [Online]. Available: https://docs.wixstatic.com/ugd/3ecad9_effb2129b4b54131ae0aeada1bae82e3.pdf [Accessed February 09, 2019].

[14] P. Kotnik and T Petrin, "Implementing a smart specialisation strategy: an evidence-based approach," International Review of Administrative Sciences, 83(1), March 2017, pp. 85-105. http://dx.doi.org/10.1177/0020852315574994

[15] S. Nutley, H. Davies, and I. Walter, "Evidence-based policy and practice: cross-sector lessons from the United Kingdom," Social Policy Journal of New Zealand, 20, June 1, 2003, pp. 29-49.

[16] M. Dosso, B.R. Martin, and P. Moncada-Paternò-Castello, "Towards evidence-based industrial research and innovation policy," Science and Public Policy, 45(2), April 1, 2018, pp. 143-150. http://dx.doi.org/10.1093/scipol/scx073

[17] European Commission, "Better regulation for better results—An EU agenda", Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Strasbourg, May 19, 2015, [Online]. Available: http://ec.europa.eu/smart-regulation/better_regulation/documents/com_2015_215_en.pdf [Accessed February 12, 2019].

[18] Nobanet, "E-business Study Course," Eliademy LMS system, 2019, [Online]. Available: https://eliademy.com/app/a/org/2522 [Accessed February 19, 2019].

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 195-199

Evaluation of Digital Marketing Use in the Promotion of Latgale Region Tourism

Ligita Karvele
Faculty of Economics and Management
Rezekne Academy of Technologies
Rezekne, Latvia
ligita.karvele@gmail.com

Daina Znotiņa
Faculty of Economics and Management
Rezekne Academy of Technologies
Rezekne, Latvia
Daina.Znotina@rta.lv

***Abstract*—Destination communication is essential for improving the economic and social development of cities, regions and countries, as tourism has become a major component of gross domestic product in most countries of the world. The development of information technology has fundamentally changed the use of traditional communication tools. Due to the development of modern information technologies, the Internet has become an integral part of tourism communication. Aim of the research is to explore the use of digital marketing in tourism promotion in Latgale region. Research methods - document analysis, statistical data analysis, survey - surveyed tourism information providers in Latgale region; statistical grouping, reading method, interview. The results of the research show that only 20% of Latgale region tourism information providers are actively involved in the use of social media. In the work of tourism information providers more attention is paid to printed materials than to video materials in the digital environment. But the most effective means of digital marketing is that all tourism information providers emphasize the importance of social networks. Digital Marketing Costs has only 40% of the Latgale region tourism information providers and they make between 100 and 3000 euros per year. On Visitlatgale.com and latgale.travel, most of the analyzed digital marketing criteria are not present. Latgale marketing action plan is quite varied and wide where costs are for many exhibitions both in Latvia and abroad, visits of Latvian tour operators and journalists, visits of foreign tour operators and bloggers, training seminars, but the marketing plan for 2019 does not include expenses for digital marketing. Latgale Region Tourism Association Ezerzeme marketing plan necessary to provide expenditure for the development of a communication plan - actively should be used funded projects possibilities, for example costs can be partly covered by the European Union's European Agricultural Fund for Rural Development (EAFRD) and the Rural Development Program. Latgale Region Tourism Association needs to attract project funding for tourism information providers training in digital marketing. More attention should be paid to publishing video materials and attracting appropriate specialists. Latgale Planning Region and Rezekne Municipality should improve tourism websites Latgale.travel and visitlatgale.com.**

***Keywords*—digital marketing, Latgale region, social networks, tourism promotion.**

I. Introduction

The Internet provides an opportunity for potential tourists themselves to search for information, to choose where and when to look for information, and not to wait for when the information from the tourist service provider will be sent. An appropriate online presence, an attractive website and effective social media management make the city, the region or the country visible, and forms the basis for attracting tourists. Understanding how online communication in a tourist destination can be improved can affect competitiveness and thus the economy of the area. Currently, websites, social networks and mobile applications are the most common platforms for promoting tourist destinations.

Latvian Tourism Development Guidelines 2014-2020 shows that one of the main directions of action is to ensure the visibility of Latvian tourism offer in the target markets, especially by using modern means of communication. The guidelines emphasize that the development of tourism is influenced by use of technologies. It facilitates access to information on tourist destinations, attractions and services, and reduces the cost of marketing activities in the long term. Accessibility of the Internet is a key factor, therefore it is emphasized that at the same time, the collection and availability of information on tourist destinations needs to be more and more coordinated, and tourism organizations and information centers play an important role here [1], [2].

Situation analysis shows that Latgale has problems with communication because there is no unified development goal and strategy for a region with a common platform and responsible, necessary is a professional approach to planning and marketing, highlighting unique advantages, creating experience and communicating. The Latgale Region Tourism Association Ezerzeme has to work with researchers, educational institutions because it is important to study how the information providers of Latgale region promote the offer on digital channels, because successful operation in the digital environment promotes competitiveness, promotes quality and productivity improvement, introduction of innovations [3].

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4055
© 2019 Ligita Karvele, Daina Znotiņa.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

Object of the research: tourism resources in Latgale region. Subject of the research: digital marketing.

Aim of the research: to explore the use of digital marketing in tourism promotion in Latgale region.

Tasks of the research: To get to know and analyze literature on digital marketing, its tools and their use in tourism; Explore how tourism resources of Latgale region are promoted in Internet sources - what digital marketing tools are used, evaluate the use of digital marketing of tourism in Latgale region in the work of tourism information providers, as well as to do the expert survey in order to obtain professional opinion regarding the research problem; Draw conclusions and make proposals for a more successful use of digital marketing in Latgale region tourism promotion.
Hypothesis of the research: Digital marketing is used poorly to promote the tourism resources of the Latgale region.

II. Materials and Methods

Research methods:

• Document analysis - planning documents that regulate tourism activity- Latvian tourism marketing strategy and other national tourism policy documents, analyzed tourism marketing plan of Latgale region, finances;
• Statistical data analysis - website activity statistics, Latgale tourism statistics;
• Survey - surveyed tourism information providers in Latgale region;
• Interview - conducted with a tourism specialist to get a professional understanding of the problem being studied.

III. results and Discussion

The tourist destination is defined as "country, region, city or other place that attracts tourists"[4]. The World Tourism Organization defines tourism as a place people want to visit [5]. The definition of tourist sites also shows the interaction between the supply and demand of a tourist site, the product and the customer as the main marketing elements, as one of the most important features of the tourism site is highlighted - a set of products consisting of separate products. Many product creators are involved in creating a product for a tourist destination but the product itself is a combination of products and services [6].

Tourism marketing is defined as a management concept based on the specificity of a tourism site, based on market competition and public interest, and is characterized by a number of stakeholders. Territorial marketing is defined as a spatial planning process that involves meeting the needs of a targeted market [7]. Determining the needs, wishes and interests of the target market is essential, as well as ensuring that the customer satisfaction level is more effective than competitors do [8].

Promotion is defined as the most convincing way to deliver the message in the right direction at the lowest possible cost [9].

In different sources of information, digital marketing is called different:

• Internet (Web 2.0.) Marketing and online advertising, also called e-marketing, Web marketing, online marketing, is the promotion of products and services on the Internet.
• Internet marketing is the use of the Internet and other digital technologies with traditional methods to achieve marketing goals. Marketing through electronic media such as the web, email, interactive television and wireless multimedia with customer-specific and behavioral digital data [10].
• Internet marketing (also known as e-marketing, Web marketing or digital marketing) is a comprehensive term for marketing products and/or services online - and like many comprehensive terms, internet marketing means different things for different people [11].
• Digital marketing, product or brand advertising using one or more types of electronic media differs from traditional marketing because channels and methods are used that allow the organization to analyze marketing campaigns and understand what works and what isn't - usually in real time.

There are different definitions of digital marketing, but the basic principle of the Internet is that users add value to content creation using many online applications including blogs and social networks [10], [12].

A survey of tourism information providers in Latgale region (tourism information centers, tourism information points and tourism organizers) was carried out in order to find out the habits, efficiency and results of use of social networks. Questionnaires were used to identify the use of digital marketing in the work of Latgale tourism information providers. According to the form of the survey - electronic survey on the Internet, distances between researcher and researcher - remote, survey procedures - individual, question forms - with closed questions. Survey with Restrictions - only a specific Latgale region tourist information provider can provide a response. Use of open-ended - unstructured answer questions and closed - structured answers questions - questions of dichotomy choice answers and scale-type choice answers questions. The survey used the terms 1, 2, 3, 4, 5, which are in ascending order of importance: 1 lowest and 5 highest. The survey held from November 22 to November 27, 2018. It was created on docs.google.com and was sent to all 16 tourism information centers in the Latgale region:

- Aglona tourism information center;
- Baltinava district tourism consultant;
- Balvi tourism information center;
- Cibla district tourism organizer;
- Dagda tourism information center;
- Daugavpils tourism information center;
- Kārsava tourism information point;
- Krāslava tourism information center;
- Līvāni tourism information center;
- Ludza tourism information center;
- Preiļi and Riebiņi tourism information center;

- Rēzekne tourism information center;
- Rēzekne district tourism information center;
- Rugāji tourism information center;
- Viļaka tourism information center;
- Zilupes district tourism organizer.

The general set is all providers of tourism information in Latgale region. All responses submitted were taken into account as each response in the study is relevant. In total, 15 responses were received.

The results of the survey show that only 20% of Latgale region tourism information providers are actively involved in the use of social media.

Most - 14 out of 15 respondents have a *Facebook* account, with a second place in *Twitter*, then *Instagram*, which has 5 tourism information providers, then *Draugiem.lv* - 4, *Instagram* - 3, *Vimeo* and elsewhere 1.

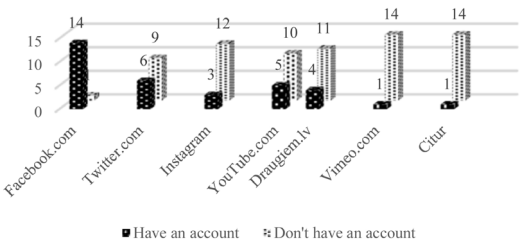


Fig.1. Respondents' answers to the survey question "Do you as a tourist information provider have an account in one of the social media (*Facebook, Twitter, Instagram, Youtube, Draugiem, vimeo*)?", number of information providers (created by the authors).

38% of all tourism information providers use hashtags but 62% do not use them.

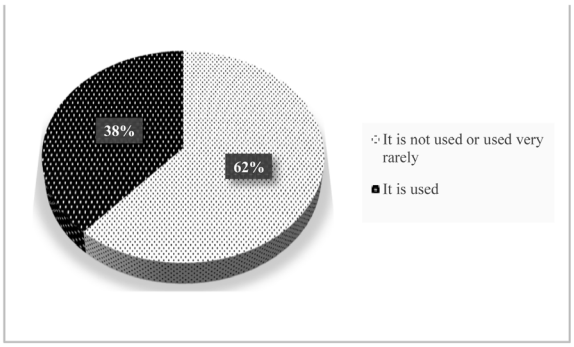


Fig.2. Respondents' answers to the survey question "When publishing information, do you use hashtags?", % (created by the authors).

Most of Tourist Information Providers - 40% in the *Facebook* account publishes the content every day but 27% a few times a week, the rest less but 7% have no accounts. 60% don't have a *Twitter* account but 13% publish the content a few times a week. 80% don't have an *Instagram* account but 20% publish the content a few times a week. 67% do not have a *YouTube* account but 20% publish their content about half a year. 73% have no *Draugiem.lv* accounts but 13% publish their content about once a month. 93% are not *Vimeo* accounts but 7% publish their content about once a year. In the work of tourism information providers more attention is paid to printed materials than to video materials in the digital environment.

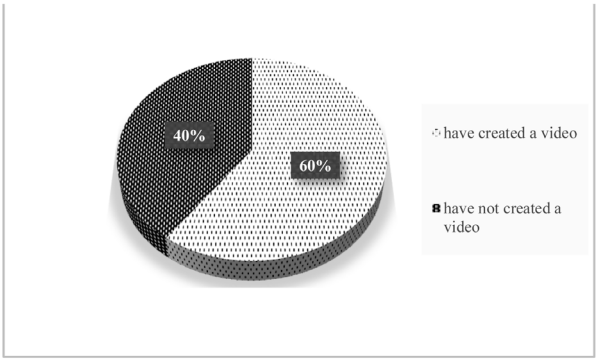


Fig.3. Respondents' answers to the survey question "Over the past 3 years, have tourism video materials about your city/county been created? ", % (created by the authors).

Target audience for publishing information in social networks is the inhabitants of Latvia. Often they are also local, county, Latgale residents, potential visitors, tourism information providers, tour operators, media, and least foreigners, families with children, niche tourists, and active social network users.

But the most effective means of digital marketing is that all tourism information providers emphasize the importance of social networks.

Digital Marketing Costs has only 40% of the Latgale region tourism information providers and they make between 100 and 3000 euros per year.

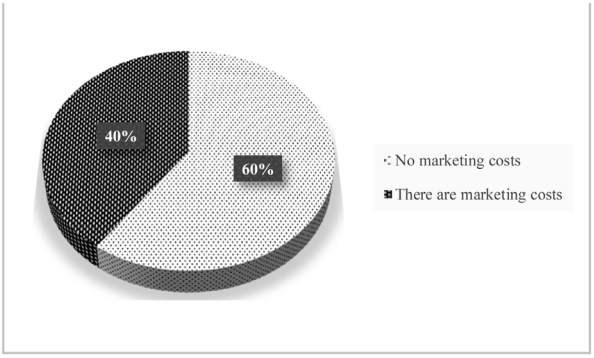


Fig.4. Respondents' answers to the survey question "Do you have digital marketing costs as a tourist information provider? ",% (created by the authors).

Most of the respondents do not use paid services because they only use free tools and work voluntarily. 68% of tourism information providers do not have an account in any social media (*Facebook, Twitter, Instagram, Vimeo, Youtube)* but 32% have an account. Most part - 14 out of 15 respondents have a *Facebook* account, *Twitter* ranks second, then *Instagram*.

Using the information available in the literature analysis, several criteria were set for evaluating the website. The official website of Latgale region is visitlatgale.com [13] and is managed by Latgale Region Development Agency. There are 2 websites in Latgale tourism where the second is http://latgale.travel/ [14]. This website was also evaluated. This site has been created because a new and modern website of Latgale tourism is needed. By comparing the analytical data of both websites it can be concluded that the visitaltgale.com

website is visited by more people, which is 33% more than latgale.travel. During the previous years Latgale was advertised with this website, also in brochures, at various tourism exhibitions, and was advertised more on the internet at visitlatgale.com. Visitlatgale.com has a more used website, has a higher number of user and page views as the website for tourism is used longer than latgale.travel. On visitlatgale.com and latgale.travel, most of the analyzed digital marketing criteria are not present so there should be improvements to make websites work efficiently and be a convenient tool for potential and existing Latgale region visitors. Google analytics data shows that visitlatgale.com is a more used website, it has a higher number of user and page views what can be explained by the fact that the website in the tourism industry is already in use for a long time - longer than latgale.travel.

Every website has its own pros and cons. To improve digital marketing, the informative arrangement of visitlatgale.com is planned, optimization of the website latgale.travel with pictograms, optimization of *Google maps*, introduction of *Facebook* Latgale.travel, and introduction of *Instagram* to *visitlatgale, Twitter* - Latgale.travel are planned. Self-assessment of tourism information providers in digital marketing shows that only 20% are actively involved in the use of social media.

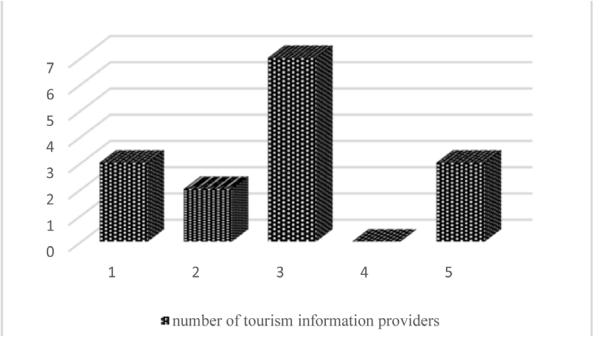


Fig.5. Respondents' answers to the survey question "How do you think how active is your involvement as a tourist information provider in social media usage and digital marketing (1-weak engagement, 5-very active)? ", the number of information providers (created by the authors).

The research was followed by an interview with Līga Kondrate - Chairman of the Board of the Latgale Region Tourism Association Ezerzeme (Research Method - Expert Survey). This was done through a one-time tourism specialist interview to gain a professional understanding of the current situation and future plans for digital marketing in the Latgale region. Professional opinion and suggestions for improving the situation were heard. By type of interview - structured, by form - individual. Interview held at Ludza Tourism information centre, at 10 a.m., November 29, 2018. Questions were read from the record sheet. During the interview Liga Kondrate emphasizes the important role of Google maps in the digital marketing of tourism, as it allows to find and view objects of interest. Travelers nowadays often visit unscheduled objects and if an object is marked on a Google map, the traveler can easily find it and include it on its route, as well as make a booking immediately. For improvement of this issue should be responsible Latgale region tourism information centers. The *Whatsapp* application on mobile phones is also important in work for tourism information providers [15].

Exploring the marketing plan of the Latgale Region Tourism Association "Ezerzeme" for 2019 [16] can conclude that the marketing plan is quite diverse and wide where costs are for many exhibitions both in Latvia and abroad, visits of Latvian tour operators and journalists, visits of foreign tour operators and bloggers, training seminars, but the marketing plan for 2019 does not include expenses for digital marketing. Tourism association Ezerzeme in the marketing plan of 2019 foresees expenses of 2800.00 euros for printed materials. About 3000.00 euros is provided for training seminars and the Latgale Tourism Conference, which includes training on topical topics in the work of tourism information providers and entrepreneurs, including topics on digital marketing and its topicalities.

IV. CONCLUSIONS

It was concluded that the Latgale Region Tourism Association Ezerzeme marketing plan necessary to provide expenditure for the development of a communication plan - actively should be used funded projects possibilities, for example costs can be partly covered by the European Union 's European Agricultural Fund for Rural Development (EAFRD) and the Rural Development Program. The *PROSPERO* public relations company offers to develop a Latgale communication plan aimed at the European target audience and, for example, Asia, and it would cost around 15,000 euros + VAT, including a company presentation on the established plan, as well as training of tourism information providers in Latgale region. The focus of the plan should be on Latgale's values, vision, events, including more detailed and more readily available information, tools with the ability to book, see what to see, how to communicate effectively through blogger / influenza experience, with a defined strategic approach, the tactics chosen, and to develop a program to market. Estimating Expenses for a Communication Plan will make it possible to implement it successfully and, therefore, the use of digital marketing in Latgale regional tourism will be more effective. Latgale Region Tourism Association needs to attract project funding for tourism information providers training in digital marketing, for example from now on from 04.02.2018. to 04.03.2018 The Rural Support Service has announced the acceptance of project applications in the EAFRD and Rural Development Program measure "Support for the development of rural tourism", where it is possible to receive support for the development and development of rural tourism services, marketing of rural tourism services and marketing campaigns for promotion of rural tourism in rural environment and expansion of new markets. This project could include seminars, conferences and training in digital marketing. Trained tourism information providers will be able to work more effectively in digital marketing, thus attracting more

visitors to the region. Latgale region tourism information providers must change their approach to marketing - focus more on digital marketing, mark all local tourist attractions on *Google maps*, create tourism packages, routes - events in Latgale must be combined with visits to other objects, as well as be able to present such products digitally - what, where, how to book, apply, etc. digitally, but not calling. Be more active in using social media, such as creating accounts on *Instagram, Twitter, Youtube*, and *Vimeo*, as well as providing digital marketing costs for online tools like Facebook.com that helps you get a wider audience of published content. It is recommended to be available in *Whatsapp* app by providing stakeholders with quick information in the digital environment. For more effective marketing you should use the option to use paid tools (such as on *Facebook* or *Instagram* to expand the audience). More active operation in digital marketing and the existence of multiple social network accounts will make it possible to capture a wider range of potential visitors, thereby attracting more visitors to the region. More attention should be paid to publishing video materials and attracting appropriate specialists. Latgale Planning Region has to introduce functions for suggestions and recommendations on the website visitlatgale.com, which allows travelers to share information, write reviews and experiences, ask questions directly, book directly, implement route planning tools, provide a tool for live streaming from various events, expanded reality features, introduce guides, travel agencies, transport and its rental, detailed access possibilities, as well as information about the region's brand profile - identity, flag, coat of arms. Rēzekne Municipality should implement a feature on the Latgale.travel website, where visitors can view and share information on Latgale social networking sites. In order to improve digital marketing it would be necessary to introduce forum, chat, online support, online booking, possibility to view maps and brochures, to be available for information about Latgale, its culture and traditions, identity, brand, easily accessible joint videos of Latgale. The introduction of these functions will improve the use of digital marketing in Latgale tourism, making search for information for potential and existing visitors more convenient and accessible.

Main conclusion is that Latgale tourism information providers use social networks inappropriate. Not everyone has social media accounts, tourism information providers do not use paid tools, as well as lack of knowledge in digital marketing. Latgale Region Tourism Marketing Plan does not foresee expenses for digital marketing, most of the evaluated criteria are not found on the regional websites.

As a result of the study, the hypothesis was confirmed because Latgale tourism information providers use social networks inappropriate. Not everyone has social media accounts, tourism information providers do not use paid tools, as well as lack of knowledge in digital marketing. Latgale Region Tourism Marketing Plan does not foresee expenses for digital marketing, most of the evaluated criteria are not found on the regional websites.

References

[1] V. Dombrovskis, "Guidlines of Latvian Tourism Development 2014-2020," Ministry of Economics, p. 61, 2014. [Online]. Available: www.visitdaugavpils.lv/.../Latvijas-tūrisma-attīstības-pamatnostādnes-2014.-2020.gadam_.doc [Accessed September 23, 2018].

[2] LIAA Tourism Department, "How to develop tourism in Latvia?", Latvian tourism marketing strategy 2018.-2023., p.15, 2018. [Online]. Available: https://www.celotajs.lv/cont/prof/news/2018/Turisma_marketinga_strategija_2018_2023.pdf [Accessed June 11, 2018].

[3] A. Van der Steina, "Problēmas Latgales reģiona tūrisma nozarē jeb Kā es braucu baudīt Latgales viesmīlību? Tūrisma un ilgtspējas pētījumu laboratorija", presented on Latgale tourism conference, Rezekne, Latvia, 2015.

[4] LR Ekonomikas ministrija, Tūrisma un viesmīlības terminu skaidrojošā vārdnīca, 2008, pp. 298.

[5] UNWTO, Handbook on Tourism Destination Branding. UN WTO: Spain, 2009, pp. 159.

[6] D. Buhalis, "Marketing the Competitive Destination of the Future", Tourism Management, vol. 21, no. 1, p. 97., 2000. [Online]. Available: https://www.academia.edu/164837/Marketing_the_competitive_destination_of_the_future [Accessed August 20, 2018], https://doi.org/10.1016/S0261-5177(99)00095-3

[7] P. Kotler and J. Makenz, Marketing for Hospitality and Tourism. Second Edition- Upper Saddle River: Prentice Hall, 1999, pp. 650.

[8] V.T.C. Middleton, Marketing in Travel and Tourism (3rd edn). Oxford: Butterworth –Heinemann, 2001, pp. 487. [Online]. Available: https://books.google.lv/books?id=8svzsNJ7fcEC&printsec=frontcover&hl=lv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=true [Accessed August 21, 2018], https://doi.org/10.1002/jtr.367

[9] C. Maccabe, Marketing communications in tourism & hospitality: concepts, strategies and cases. Oxford: Elsevier, 2009, pp. 220. [Online]. Available: https://books.google.lv/books?id=GvYJBAAAQBAJ&pg=PR3&lpg=PR3&dq=Marketing+communications+in+tourism+%26+hospitality:+concepts,+strategies+and+cases+maccabe&source=bl&ots=S5tgdtl-Np&sig=ACfU3U3ziGEq_rLJjS1tgQzEGtTL02uJUA&hl=lv&sa=X&ved=2ahUKEwjyjaTbvNbhAhUHtIsKHTwYDwMQ6AEwBXoECAkQAQ#v=onepage&q=Marketing%20communications%20in%20tourism%20%26%20hospitality%3A%20concepts%2C%20strategies%20and%20cases%20maccabe&f=false [Accessed August 25, 2018].

[10] D. Chaffey, F. Ellis-Chadwick, R. Mayer and K. Johnston, Internet Marketing: Strategy, Implementation and Practice. 4th Edition. Prentice Hall, 2009, pp. 559. [Online]. Available: https://books.google.lv/books?id=HcoRl2EZXiwC&printsec=frontcover&hl=lv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false [Accessed September 5, 2018].

[11] G. Kaur, The importance of digital marketing in the tourism industry, International Journal of Research - Granthaalayah, vol. 5, no. 6, pp. 72., 2017. [Online]. Available:https://zenodo.org/record/815854#.XLC4YqSE6 8Articles [Accessed September 7, 2018], https://doi.org/10.5281/zenodo.815854

[12] S. Parise, P. J. Guinan, Marketing Using Web 2.0. Proceedings of the 41st Hawaii International Conferece on System Sciences, 2008, pp.1. [Online]. Available: https://www.academia.edu/17431342/Marketing_Using_Web_2.0 [Accessed September 7, 2018],

[13] www.visitlatgale.com [Accessed: Sept. 10, 2018].

[14] http://latgale.travel/ [Accessed: Sept. 10, 2018].

[15] Kondrāte, L., the head of Latgale region tourism association Ezerzeme. [Interview]. At 10 a.m., November 29, 2018,.

[16] Latgale Region Tourism Association Ezerzeme, Tourism marketing plan, 2018.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 200-204

Imitation Modeling of Social Investments in Early Childhood Education

Katrīne Kūkoja
Institute of Social, Economic and Humanities Research (HESPI)
Vidzeme University of Applied Sciences (ViA)
Valmiera, Latvia
katrine.kukoja@va.lv

Feliciana Rajevska
Institute of Social, Economic and Humanities Research (HESPI)
Vidzeme University of Applied Sciences (ViA)
Valmiera, Latvia
feliciana.rajevska@va.lv

***Abstract*—Previous researches show, that social investments in early childhood (until age of 6) are very important, and one of the main types of social investment in this period of time is quality early childhood education and care services (henceforth – ECEC services). Social investments in early childhood have the highest return in human capital, especially for disadvantaged groups of the society. Literature suggests that positive effect of available and qualitative ECEC services can be observed in many areas of life, for example, through improved education indicators, increased participation of the women in the labour market, increased work productivity, higher birth rates, etc. However, the attendance of ECEC services varies considerably between different European Union (henceforth – EU) countries.**

The aim of this paper is to analyze social investment effect in early childhood in EU countries with high (Denmark), medium (Latvia) and low (Croatia) ECEC service attendance rates, so to find out if and how ECEC service attendance rate in these countries influences such indicators as rate of early leavers from education and training, school graduation rates and female employment rate. To see how situation may change due to the ECEC attendance rate changes in each country analyses was be carried out with imitation model options.

This article describes the imitation model, which provides an opportunity to analyze the current situation and compare social investments between different EU countries. The model is designed using ISEE system STELLA MODELING and Imitation Software environment, and it is governed by the criteria that are based on literature studies, as well as analyses of statistic data. An additional benefit is that this model can be adjusted and used to evaluate other countries social investment tendencies as well.

Research results show, that not all indicators that have been positively connected with ECEC attendance rate in previous studies have been positively influenced by ECEC attendance rate in chosen case study countries. Analyses showed that only in Croatia positive connections can be observed when it comes to all three indicators. Further studies should be carried out on this topic to get a deeper understanding on ECEC attendance rate effect on these and other indicators, that in the previous studies have been positively linked with ECEC attendance rate.

***Keywords*—Child development, Early childhood education, Female employment, Social investment.**

Introduction

Social investment is about investing in people. It means policies designed to strengthen people's skills and capacities and support them to participate fully in employment and social life. One of the main policy areas of social investment is quality childcare [1]. Over the years literature and previous studies have shown that social investments in early age (until age of 6) has the highest return in the human capital (see Fig. 1, [2]) and their effect, mainly through quality childcare, can be observed in many areas of life both in long and short term, including [2], [3], [4], [5], [6], [7], [8], [9], [10]:

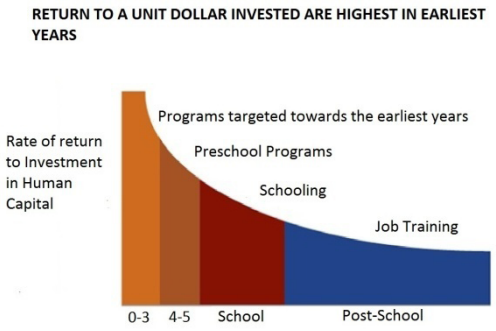


Fig. 1. Return to a dollar invested in different levels of education [2].

- education outcome increase (lower rate of early leavers from education and training, higher school graduation rates, etc.)
- social economic outcome increase (higher fertility rate and increased female workforce participation, etc.);
- exchequer returns increase (higher tax returns, etc.)
- long-term economic outcome increase (increased earnings and labour-force participation, etc.)
- health, social well-being and welfare dependency reduction (reduction in expenditures for the child welfare system, etc.)

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4075
© 2019 Katrīne Kūkoja, Feliciana Rajevska
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

- crime rate reduction (reductions in criminal justice system expenditures for youth and adult crime, etc.)

Previous researches show that indicators mentioned above are positively connected with early childhood education and care service (henceforth – ECEC) attendance rate, that is why, within this article three of these indicators were selected randomly – rate of early leavers from education and training, school graduation rates and female (mothers') employment rate, to see if these tendencies take place in chosen case study countries as well. Literature shows that ECEC services have as much to do with promoting mothers' employment as it does with improving development of young children [2],[7],[8]. Researches show that attending ECE classes helped pupils to better prepare for school and increases possibility for a child to graduate and gain better education that will later on help them to succeed in work life [2],[3],[8].

Importance of social investment in early childhood has been stressed in both Latvian, as well as European Union (henceforth – EU) planning documents. One of the policy priorities defined by both the European Commision and the European Parliament in order to meet the Europe2020 targets is to "ensure universal provision of ECEC" [11], [12].

Since the main form of social investment in this period of life is accessible and qualitative ECEC services, authors have analysed exactly this form of investments in selected country case studies – Denmark, Latvia and Croatia. These countries were selected based on the ECEC attendance rate - one EU country with high, average and low ECEC service attendance rate were chosen. Authors chose this approach to have wider view of social investment effect in different EU countries. Eurostat data show that in 2016 98,1% children in Denmark, 95.5% children in Latvia and 75.1% children in Croatia aged between 4 and the age of starting compulsory education were in early childhood education [13]. When it comes to children before age of four (including), ECEC attendance rate is considerably lower. In 2016 73.44% of children until age of four (including) attended ECEC in Denmark, while in Latvia it was 54.71% and in Croatia 35.64%. As we can see, if compared, ECEC attendance rate in Denmark is the highest, in Croatia the lowest, while in Latvia it has been relatively average [14], [15].

The aim of this research was to see if and how availability of ECEC over the past years has influenced selected indicators in three EU countries, as well as to see how situation may change due to the ECEC service attendance rate changes in each country. To do that simulation model was created.

Materials and methods

To achieve aim of this research analysis of previous researches and policy documents was performed, as well as historical data analysis of ECEC attendance rate and chosen indicators - women employment rate, rate of early leavers from education and training and upper–secondary education graduation rate in selected countries in period from 2013 to 2017, with exception in case of Croatia, since data only about period from 2013 to 2016 were available. Data were selected and analyzed in five (in case of Croatia – four) year period due to the fact that there is a lack of open-access data that could be used for comparison between the countries. That is why it was not possible to analyze longer period of time. In many cases data were not available or they were structured differently between countries.

To get a wider prospect, case studies in three countries with different ECEC attendance rates were carried out. EU countries with high - Denmark, medium – Latvia, and relatively low – Croatia, ECEC service attendance rates were chosen.

Imitation model was developed using ISEE system STELLA MODELING and Imitation Software environment to better illustrate, analyze and compare the situations in selected countries. Based on historical statistic data about ECEC attendance rate and chosen indicators in selected countries, model has the possibility to generate how each of the above mentioned indicators will change if ECEC attendance

TABLE I. ECEC attendance rate and selected indicator average values in Denmark, Latvia and Croatia

	ECEC attendance rate	Graduation rate	Rate of early leavers from education and training	Women employment
DK	73.72	11.3	7.92	77.97*
LV	52.30	8.5*	9.36*	65.78
HR	34.02	9.67	3.23	61.33

*Won't be used in the model

rate will change in the each country. Model is static, determined, based on discrete events and divided into three sectors, where each of the sectors represents one of the countries.

Before date usage in model, data were analyzed and edited using SPSS Software to see if these data are statistically relevant, usable and valid, as well as to see if correlations between ECEC attendance rate and selected indicators in all three countries can be observed in the researched period.

Since the One-Sample Kolmogorov-Smirnov Test at significance level 0.05 confirmed that with 95% probability ECEC attendance rate, as well as women's employment, graduation rate and rate of early leavers from education and training in analyzed period corresponds to normal data distribution in Denmark, Croatia and Latvia, the average values of the indicators where used in this model as status quo values, see table I [14], [15].

Model description

At the interface level of the Social investment effect simulation model, there are two buttons for easy operation of the model and six displays for results that show how indicator values change, if we change ECEC attendance rate. The button "Run" provides the opportunity to start the data processing of components after the input

of variable data. The button "Restore Sliders" is made for user's convenience. At the model-making level the current ECEC attendance rate and indicator parameters are entered, then this button allows resetting original parameters (status quo) after the changes are made. In Fig.2 we can see the model interface view at status quo position.

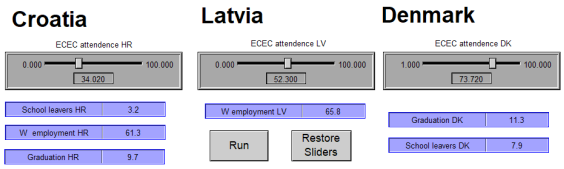


Fig. 2. Model interface view–status quo position.

RESULTS AND DISCUSSION

TABLE II. ECEC ATTENDANCE RATE UNTIL AGE OF 4 (INCLUDING) IN DENMARK, LATVIA AND CROATIA (2013-2017/2016).

	2013	2014	2015	2016	2017
DK	74.68	74.76	73.82	73.44	71.89
HR	32.68	33.45	34.31	35.64	
LV	49.39	50.32	52.64	54.71	54.44

When we analyse social investment in early childhood it is important to look at the ECEC service availability, especially in the early age. As we saw in the previous selection – while the ECEC attendance rate for children aged between 4 and the age of starting compulsory education were relatively high and didn't have significant differences between selected case study countries (especially Latvia and Denmark), when it came to younger children the results were not so positive. Therefore, authors decided that in this article data regarding ECEC attendance before age of 4 (including) will be used.

As we can see in the table II attendance rate of ECEC until age of 4 (including) has been the highest in Denmark. However we can see a tendency of ECEC attendance rate to increase in Latvia and Croatia since 2013, while in Denmark the rate has been steady with little decrease in 2017 [14], [15].

As we can see in the table III [14], [15], despite the previous research results, not all indicators correlated positively with the ECEC attendance rate in the case study countries. When it comes to Latvia, only women employment in the period from 2013 till 2017 had a positive correlation with ECEC attendance rate. In Denmark situation, we can see that there is a positive correlation between ECEC attendance rate and women employment, as well as relatively average negative correlation between ECEC attendance rate and rate of early leavers from education and training. While only in the Croatia ECEC attendance rate has had positive impact on all three indicators, although the correlations were average when it came to women employment, and little bit above average when it came to graduation rates and rate of early leavers from education and training.

TABLE III. CORRELATION BETWEEN RVRV ATTENDANCE RATE AND SELECTED INDICATORS IN DENMARK, LATVIA AND CROATIA IN 2003 - 2017

ECEC attendance rate	Indicators		
	Women employment	Graduation rate	Rate of early leavers from education and training
DK	0.941	-0.638*	-0.546
LV	0.885	-0.893*	0.073*
HR	0.564	0.602	-0.704

* Won't be used in the model

TABLE IV. ECEC ATTENDANCE RATE AND SELECTED FACTOR STATISTIC DATA IN 2015 IN DENMARK, LATVIA AND CROATIA.

	ECEC attendance rate	**Graduation rate**	**Rate of early leavers from education and training**	**Women employment**
DK	73.82	11.24	7.92	76.88
LV	52.64	9	9.36	64.13
HR	34.31	9.277	3.23	61.87

Since analyses of the data didn't show positive correlations between ECEC attendance rate and graduation rate in upper secondary level of education in Latvia and Denmark, as well as negative correlation with rate of early leavers from education and training in Latvia, these indicators were not included in the model in the case of Latvia and Denmark. But authors suggest that further studies regarding these factors and the research results should be carried out. All the other indicators were included in the model.

Despite the fact that not all of the chosen indicators correlated positively with ECEC attendance in all countries, Denmark still showed best performance in almost all indicators, except in the case of rate of early leavers from education and training. As we can see in table I and IV, lowest school leavers rate in researched period was in Croatia. Croatia showed better performance then Latvia in two out of 3 indicators - rate of early leavers from education and training and upper-secondary level of education graduation rate. In table IV situation in year 2015 can be observed [14], [15].

As we can see in the Fig. 2, where model interface view at status quo position is shown, average ECEC attendance rate in Croatia in researched period was 34,02% for children until age of four (including), in Latvia - 52,3% and in Denmark – 73.72%. Average rate of early leavers from education and training was the lowest in Croatia 3.23% and in Denmark it is 7.92%. Data for the second indicator, women employment rate in Croatia was 61.3% and in Latvia – 65.8%. For the last selected indicator – graduation rate in upper-secondary education, that was included in the model only for Denmark and Croatia, the average rate was 9,7% in Croatia and 11,3% in Denmark.

The developed simulation model not only helps to analyze and compare social investments effect on different indicators, but also gives an opportunity to see how changes in ECEC attendance rate in each country may influence selected indicators.

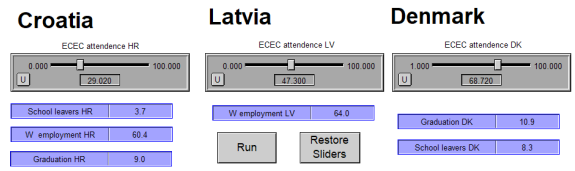


Fig. 3. Model interface view– second case where ECEC attendance decreasd by 5%.

During this study, the simulation model verification was carried out to avoid gaps in the model. Two case scenarios were simulated using this imitation model – the case when ECEC attendance rate decreases for 5% in each country and the case where it increases for 3%.

In the first case simulation model shows that by decreasing ECEC attendance by 5% it negatively influences all of the selected indicators in case study countries. As we can see in Fig. 3 in Croatia case data shows that decrease in ECEC attendance can increase school leaver rate for averagely 0.5%, women employment for 0.9 % and graduation rate could decrease for 0.7%. When it comes to Latvia, simulation shows that with these changes women employment rate could drop for 1.8%, since the correlation between these indicators were relatively strong. As for the Denmark, based on the simulation results, decrease in the ECEC attendance rate could lower graduation rate for 0.4%, and increase school leaver's rate for 0.4%.

In the second case simulation model shows that by increasing ECEC attendance for 3% it positively influences all of the selected indicators in three countries. As we can see in the Fig.4 this little increase in attendance rate can positively influence many areas of life. For example, women employment rate in Croatia could be increased for 0.6% and in Latvia for 1%, graduation rates increased for 0,2% in Denmark to 0.3% in Croatia, and rate of early leavers of education and training decreased for 0.2% in Denmark to 0.2% in Croatia.

However, it is necessary to take into consideration that these simulation results are approximate and are reliable with the condition, that all the other factors that influence these indicators stay the same.

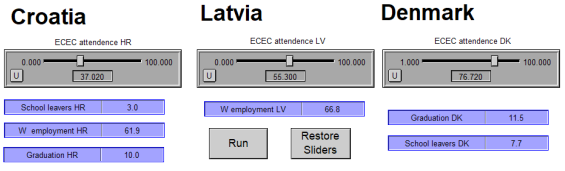


Fig. 4. Model interface view– second case where ECEC attendance increases by 3%.

Conclusion

Research results shows that during the last years there has been increase in social investments in early childhood, especially when we look at ECEC service availability- attendance rate increase in Latvia and Croatia. However, despite the fact that attendance rate is the highest in Denmark, we cannot observe increase in it during the last five years. This may be due to the fact that Denmark already has rather high and stable result in this area as outcome of previous investments, but further studies should be carried out concerning this question.

Statistic data show that Denmark has the highest ECEC attendance rate and has been the leader when it comes to women employment and graduation of upper-secondary education, but level of early school leavers has been relatively high in Denmark, even higher than in Croatia, that showed the lowest rate of early leavers from education and training in the period from 2013-2017.

Despite the fact that Latvia has the second highest ECEC attendance rate of the case study countries, only one of the studied indicators showed to be higher in Latvia, then in Croatia – women employment. This may mean that in case of the graduation rate and rate of early school leavers, ECEC attendance doesn't show such an impact in Latvia, and other factors, for example, ECEC quality, may have a stronger impact on these results. Further studies should be carried out in this field.

Research results show that not all indicators that have been positively connected in previous studies with increased ECEC attendance showed positive correlation in Latvia and Denmark. The only country where ECEC attendance showed high or at least average connection with all of the selected indicators was Croatia. In Denmark case positive correlation could not be observed in the case of graduation rate in upper-secondary education level, but in Latvia – in the case of graduation rate and rate of early leavers of education and training.

Research shows that ECEC attendance rate can positively influence child's development and women employment in most of the selected countries, but in case of education attainment (child's development) situation is more complicated, so authors recommend to carry more studies in longer period of time and with more indicators, to have even wider view on the situation of investment effect.

As a result of this research, a simulation model in the field of ECEC effect on child's development and women employment was developed. Such approach for the determination of the ECEC effect on different indicators, that have previously been positively linked to ECEC attendance rate, has not been used so far before. This model gives a possibility to look at the effect of ECEC attendance on different indicators in different EU countries.

Acknowledgments

Authors would like to express their gratitude to the Institute of Social, Economic and Humanities Research of Vidzeme University of Applied Sciences for provided support in doing this research.

References

[1] European Commision, "Home page - European Commision," February 2019. [Online]. Available: https://ec.europa.eu/social/main.jsp?catId=1044&langId=en [Accessed Jan. 28, 2019]

[2] J. J. Heckman, „Schools, skills, and synapses," Economic Inquiry, vol. 46, no. 3, p. 289□324, 2008 [Online] Available: NBER, https://www.nber.org/papers/w14064 [Accessed June 1, 2018], DOI: 10.3386/w14064

[3] R.M. Jacinta and K.S. Jacinta, "Impact of Early Childhood Education on Pupils' Learning in Primary Schools in Kenya," Global Journal of Educational Studies, vol. 1, no. 1, p. 52-61, May, 2015 [Online] Available: Macrothink Institute, http://www.macrothink.org/journal/index.php/gjes/article/view/7686 [Accessed Nov. 20, 2018], http://dx.doi.org/10.5296/gjes.v1i1.7686

[4] A. J. Reynolds, J. A. Temple, B. A. White, S. R. Ou and D. L. Robertson, "Age-26 Cost-Benefit Analysis of the Child-Parent Center Early Education Program," Child Development, vol 82, no. 1, p. 379-404, January-February 2011 [Online]. Available: NCBI, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3817956/ [Accessed May 9, 2018], https://doi.org/10.1111/j.1467-8624.2010.01563.x

[5] Education, Audiovisual and Culture Executive Agency and Eurydice, "Early Childhood Education and Care in Europe: Tackling Social and Cultural Inequalities," Education, Audiovisual and Culture Executive Agency and Eurydice, p. 188, 2009 [Online]. Available: https://publications.europa.eu/en/publication-detail/-/publication/83f56fb3-62f0-440e-b646-c782f2abef18/language-lv [Accessed: Dec. 20, 2018].

[6] T. Havnes and M. Mogstad, "No Child Left Behind Universal Child Care and Children's Long-Run Outcomes," American Economic Journal: Economic Policy, vol. 3, no. 2, p. 97-129, May 2011 [Online]. Available: SSB, https://www.ssb.no/a/publikasjoner/pdf/DP/dp582.pdf [Accessed Dec. 20, 2018], DOI: 10.1257/pol.3.2.97

[7] Janna van Belle, "Early Childhood Education and Care (ECEC) and its long-term effects on educational and labour market outcomes," Janna van Belle, p.33, 2016. [Online] Available: https://www.rand.org/pubs/research_reports/RR1667.html [Accessed: Jan. 12, 2019], https://doi.org/10.7249/RR1667

[8] N. Morel, B. Palier and J. Palme, Towards a Social Investment Welfare State?: Ideas, Policies and Challenges The Policy press, Great Britain, 2012

[9] Editor-in chief F. Rajevska, Turns of Social Policy Transformation in post-crisis years, National research program EKOSOC_LV, project 5.2.6. LU, RSU, ViA, Riga, 2018.

[10] F. Campbell, G. Conti, J.J. Heckman, S.H. Moon, R. Pinto, E. Pungello, Y. Pan , " Early Childhood Investments Substantially Boost Adult Health," Science, vol. 343, issue 6178, p. 1478-1485, March 2014 [Online] Available: Science, http://science.sciencemag.org/content/343/6178/1478.long [Accessed Nov 2, 2018], DOI: 10.1126/science.1248429

[11] European Commission," Early Childhood Education and Care: Providing All Our Children the World of Tomorrow," European Commission, p.9, 2011.[Online]. Available: file:///C:/Users/Lietotajs/Desktop/Downloads/RAND_RR1667.pdf [Accessed: October. 17, 2018].

[12] European Parliament, "Resolution of 12 May 2011 on Early Years Learning in the European Union, 2010/2159/INI," European Parliament, p. 7, 2011.[Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011IP0231&-from=EN [Accessed: October. 17, 2018].

[13] European Commision, "Home page - European Commision," September 2018. [Online]. Available: https: //ec.europa.eu/ eurostat/statistics-explained/index.php/Early_childhood _and_primary_education_statistics [Accessed: Jan, 2019]

[14] Central Statistical Bureau of Latvia , "Home page - Central Statistical Bureau of Latvia," December 2018 [online]Available: http://www.csb.gov.lv/en [Accessed: Jan. 15, 2019].

[15] European Commision, "Home page - European Commision," September 2018. [Online]. Available: https://ec.europa.eu/eurostat/data/database [Accessed: Jan. 15, 2019].

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 205-207

Fuzzy Modelling of the Academic Staff Attestation Process

Milena Lambeva
Artillery, Air Defense and CIS Faculty
Vasil Levski National Military University
Shumen, Bulgaria
uzunova_milen@abv.bg

Valentin Atanasov
Artillery, Air Defense and CIS Faculty
Vasil Levski National Military University
Shumen,Bulgaria
valopaint@yahoo.com

***Abstract*—A large class of real-world problems, arising in the administrative - organizational systems is connected with the attestation of the academic staff in accordance with predefined and linguistically described evaluation criteria. Such problems are solved in the circumstances of uncertainties of different types. The process of interpretation goes together with some ambiguity and inaccuracy due to the subjective interpretation of the criteria by experts. The Fuzzy Set theory provides a convenient apparatus for formalization and creation of new intelligent decision-making methods, based on analysis and processing of the expert knowledge. This paper offers a method of fuzzy modelling of the problem of scientific and pedagogical staff attestation which gives the possibility of an adequate interpretation of the attestation results. The method can be implemented programmatically and can be incorporated in the systems of education quality management.**

Keywords— Artificial intelligence methods, attestation of the academic staff, fuzzy models.

I. Introduction

The attestation of the academic staff is a "procedure of systematic formalized assessment according to the specified criteria for the correspondence of the employee's activity to the standards of work performance at a given workplace, for a given position, for a certain period of time" [1].

Attestation of scientific and pedagogical staff is an important condition for establishment of an academic work climate in scientific communities and it contributes to the achievement of scientific results. The implementation of quality management systems of training requires automation of this process through the development of mathematical methods for an adequate assessment of cadre staff. The development of automated intelligent systems for attestation of the academic staff is an actual problem on which scientists from many universities work [2] – [6].

The use of such systems in practice of educational institutions can significantly facilitate the work of specialists in the field of pedagogical diagnostics and contribute to the objectification of the evaluation process of academic staff and elimination of subjectivity and randomness.

II. Analysis Of The Decision-Making Process For The Estimation Of The Scientific And Pedagogical Staff

The attestation of the personnel in departments is carried out by the head of the department and is sometimes discussed by the scientific council of the department.

Analysing the process of thinking of the expert (the head) in the decision-making process for evaluation of the scientific and pedagogical staff, we can draw some conclusions related to the essence of this process.

The analysis of the decision-making process for evaluation of scientific and pedagogical staff shows that:

1. Evaluation criteria are fuzzy and are given by the linguistic terms of the type:
   - employee always (sometimes, never, often, rarely) takes part in the activities of the department;
   - the employee has good (bad, neutral, excellent) relations with the colleagues (students);
   - has a lot of (a few, several, no) scientific publications;
   - the scientific results are excellent (good, bad, insignificant);
   - conducts classes at a high (bad, low, satisfactory) methodological level;
   - has excellent (bad, insufficient, good) scientific preparation;
   - helps colleagues (students) always (very rarely, it's hard to make them help someone, never helps anyone), etc.
2. The choice of criteria is carried out by experts and strongly depends on their experience, professionalism and competence.
3. In real problems, the criteria are formulated linguistically; the process of interpretation is accompanied by inaccuracy and ambiguity due to the subjective interpretation by experts.

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4126
© 2019 Milena Lambeva, Valentin Atanasov.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

4. Expertise and ranking of scientific and pedagogical staff is a result of establishing fuzzy relations between evaluations according to the fuzzy criteria and are formulated on linguistic scales.
5. In the process of resources ranking, we want to obtain a scalar quantitative or qualitative assessment on the basis of particular quantitative or qualitative estimates in accordance with pre-determined criteria.

Therefore, the task of attestation of scientific and pedagogical staff belongs to the class of multicriteria, non- formalized problems that are characterized by fuzzy criteria and linguistic scales of assessments.

The decision-making process is subjective, which does not allow the application of well-known classical methods and models of decision-making theory. Therefore, for the formalization and decision-making it is appropriate to use the apparatus of fuzzy sets, which makes it possible to model the uncertainties and to formalize the fuzzy decision-making process by the expert.

III. Formalization Of The Problem Of The Academic Personel Attestation

We assume that a linguistic variable (LV) [7] <*teacher's work*> with a given term-set, for example {bad, satisfactory, good, excellent}, is subject to assessment. For estimation we use the criteria representing linguistic variables of lower rank, which can be more easily assessed.

Let:
D - discrete or continuous set of diagnoses;
P - $\{p_1, p_2, \ldots p_m\}$ - a discrete set of teachers which are subject to attestation;
C - $\{C_1, C_2, \ldots C_n\}$- a discrete set of evaluation criteria defined by experts;
$A = || a_{ij} ||, i=1,2,\ldots m, j=1,2,\ldots n$ - the matrix containing the results of evaluating teachers;
$a_{ij} \in L_j$ - an estimation of the i^{-th} teacher in accordance with the C_j-criterion;
L_j - a discrete or continuous scale of estimates in accordance with the C_j-criterion;
In practice for an estimator, it is easier to use the same scales for the assessment according to different criteria, i.e

$$L_1 = L_2 = \ldots = L_n = L, \quad (1)$$

for example

$$L = \{bad, satisfactory, good, excellent\} \quad (2)$$

The greatest meaningful correspondence of the linguistic model of decision-making process for attestation of scientific and pedagogical staff can be achieved if we consider it as a diagnostic problem of the kind:

$$< \mathrm{P}, \mathrm{C}, \mathrm{L}, \mathrm{A}, \mathrm{D} >, \quad (3)$$

with the following formulation:

For each teacher $p_i \in P$ the diagnosis $d \in D$ must be determined on the basis of the results in the matrix A containing the estimations in accordance with the criteria *C* given in the *L* scale.

Formally, this means to find an injective image:

$$\Omega: P \rightarrow D \quad (4)$$

of the set of teachers subject to attestation to the set of diagnosis D, in which each element $p_i \in$ P corresponds to exactly one element d $\in$ D.

IV. Fuzzy Model Of The Problem Of The Academic Staff Attestation

The identification of the image Ω can be obtained on the basis of the analysis of the evaluation criteria and the decision-making process.

Most often in practice, fuzzy criteria of the kind <β isβ^*> are used [8], where β is the name of the linguistic variable:

$$(LV) < \beta, \mathrm{T}_\beta, \mathrm{X}_\beta, G_\beta, M_\beta > \quad (5)$$

on the universum X_β, T_β is the term set, $\beta \in \mathrm{T}_\beta$, G_β is the set of rules for obtaining new *LP* values, M_β is the set of rules for transforming new values into a fuzzy variable [3]. For example, an estimation $\{good\ skills\ for\ teamwork\}$, which is a fuzzy set on an arbitrary set of teachers, can be modeled by a linguistic variable of the form < $\beta, \mathrm{T}_\beta, \mathrm{X}_\beta, G_\beta, M_\beta >$, where:

$$\beta = < teamwork\ skills > \quad (6)$$

$$\mathrm{T}_\beta = \{bad, satisfactory, good, excellent\} \quad (7)$$

X_β- a non-numeric base set.

The decision-making process is a fuzzy implication [9] of the form:

```
If for :
<C1 is C1*> AND <C2 is C2*> AND ....
AND <Cn is Cn*>
THEN d = d * ∈ D
```

In the process of attestation of teachers, the head of the scientific section actually defines a vector-valued degree of membership to the Cartesian product $C = C_1 \times C_2 \times \ldots \times C_n$ of fuzzy criteria with the membership function [9]:

$$\mu_{c1xc2x\ldots xcn}(p_i) = (\mu_{c1}(p_i), \mu_{c2}(p_i), \ldots \mu_{cj}(p_i), \ldots, \mu_{cn}(p_i)) = (l_{i1}, l_{i2}, l_{i3}, \ldots l_{in}), \quad (8)$$

is a multidimensional fuzzy set over the set of attested ones, heterogeneous in the general case when : $L_1 \neq L_2 \neq \ldots \neq L_n$:

$$\mu_{c1xc2\ldots xcn}: P \rightarrow L_1 \times L_2 \times \ldots \times L_n \quad (9)$$

and homogeneous, provided that $L_1 = L_2 = \ldots = L_n$:

$$\mu_{c1xc2\ldots xcn}: P \rightarrow = L_n \quad (10)$$

To obtain the diagnosisd $\in D$ for $p_i \in P$, we need to transform the evaluation vector into a scalar. This can easily be done by assuming that estimation scales coincide, i.e.

$$L = L_1 = L_2 = \dots = L_n = \{l_1, l_2, \dots l_k\} \text{ and } D = \{d_1, d_2, \dots d_k\} \quad (11)$$

For each element $(l_{i1}, l_{i2}, l_{i3}, \dots l_{in})$ of the set C, we introduce the following norm:

$$: C \to [0,1], \ (p_i) = (l_{i1}, l_{i2}, l_{i3}, \dots l_{in}) \Big| = \frac{i_1 + i_2 + \dots + i_n}{\sum_{i=1}^{i=n} i} \quad (12)$$

and multivalued function:

$$\varphi: [0,1] \to D, \varphi(x) = d_j \text{ for } x \quad [(j \quad 1)/k, j/k]. \quad (13)$$

The required image Ω, can be represented as a composition μ, h, φ in the following way:

$$\Omega : P \overset{\mu}{\to} L^n \overset{h}{\to} [0,1] \overset{\varphi}{\to} D \quad (14)$$

V. Fuzzy Model Analysis

In general, the manager can choose an assessment from:

- a linearly ordered set $L = \{l_1, l_2, \dots l_k\}$, that meets the requirements of the finiteness and completeness;
- Interval [0,1].

The scales used in practice can be reduced to these two simple transformations. Scales of the form $L = \{l_1, l_2, \dots l_k\}$ correspond to a greater extent to the characteristics of human thinking.

It is easier for the manager to assess the work of an employee using a qualitative scale, for example:
<X is well prepared, has many publications, always helps colleagues, etc.>.

In the same way, diagnoses can be chosen from:

- a linearly ordered set $D = \{d_1, d_2, \dots d_k\}$;
- the Interval [0,1], showing the intensity of the investigated LV.

In the first case, the staff will be divided into disjoint classes of excellent workers, poor workers, etc. Bearing in mind that there not many people in scientific task groups (7-15), this is quite suitable for the attestation process. The interval [0,1] will provide the possibility to get a quantitative assessment and rank the employees.

VI. Conclusions And Future Work

The proposed method of fuzzy modelling of academic staff estimation problem makes it possible to adequately interpret the results of attestation.

Practice shows that we think in fuzzy, qualitative categories, so the fuzzy model is more adequate to human thinking than the models of conventional mathematics.

A fuzzy model can be used in practice as, because it allows a software application Fuzzy system for the academic staff attestation is being developed in the National Military University of Bulgaria. This intelligent system for attestation uses the above described method.

References

[1] Борисова Е., "Оценка и аттестация персонала.", Питер, 2002, ISBN 5-94723-388-6

[2] Saint Petersburg State University, "Two-factor multi-criteria methodology of certification of scientific and pedagogical workers of St. Petersburg state University on the basis of indicators of efficiency of their research and educational activities based on their personal achievements. Report on the research work". Available: https://docviewer.yandex.ru/view/0/?*=un9UldlP55D4gd6Mye1Yi5KMeqV7InVybCI-6Imh0dHA6Ly9jc3Iuc3BidS5ydS93cC1jb250ZW50L3Vw-bG9hZHMvMjAxMi8wMy9yZXBvcnRfYXR0ZXN0YX-Rpb24uZG9jIiwidGl0bGUiOiJyZXBvcnRfYXR0ZXN0YXRp-b24uZG9jIiwidWlkIjoiMCIsInl1IjoiOTM4NTg3OTAxMTQ4N-jgzMTk0NiIsIm5vaWZyYW1lIjp0cnVlLCJ0cyI6MTQ5M-jE1MDUwMDUxNH0%3D&lang=ru [Accessed: April 14, 2018]

[3] Vasil Levski National Military University, "Rules for recruitment, development, assessment and certification of academic staff in the National Military University "Vasil Levski"". Available: http://nvu.bg/sites/default/files/nmu_files/norm_baza/pravilnici_nvu/4-Prav_podb-ocen-razv-atestir-akadem-sastav.pdf [Accessed: April 14, 2018]

[4] Dubois D., Grabisch M., Prade H., Smets P. Assessing the value of a candidate. Comparing belief functions and possibility theories. XVth Conf. Uncertainty in Artificial Intelligence, San Francisco, 1999

[5] Dubois D., Grabisch M., Prade H., Smets P. Using the transferable belief model and a qualitative possibility theory approach on an illustrative example:the assessment of the value of a candidate. Available: http://iridia.ulb.ac.be/~psmets/SelCandiFusion.pdf [Accessed April 14, 2018]

[6] Slavyanov K.O., An algorithm of fuzzy inference system for human resources selection tools, "Society. Integration. Education", Proceedings of the International Scientific Conference. Volume V, Rezekne, May 25th-26th, 2018. 445-454, http://dx.doi.org/10.17770/sie2018vol1.3311

[7] Заде Л. Понятие лингвистической переменной и ее применение к принятию приближенных решений. - М.: Мир, 1976

[8] Нечеткие множества в системах управления. /Под ред. Ю.Н.Золотухина. Курс лекций в Новосибирском Государственном Университете, 2000

[9] Василева М., Размити множества. Теория и практика. ЦВКБ-Троян, 2008., ISBN 978-954-9681-33-8

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 208-213

Improving Computer Engineering and Information Technologies Undergraduate Students' Training Through Combination of Formal, Non-Formal and Informal Learning

Oleksandr Malykhin
The Department of Romance and Germanic Languages and Translation
National University of Life and Environmental Sciences of Ukraine
Kyiv, Ukraine
malykhinalex1972@gmail.com

Nataliia Aristova
The Department of English Philology
National University of Life and Environmental Sciences of Ukraine
Kyiv, Ukraine
n.aristova.na@gmail.com

***Abstract*—The paper analyses benefits of combining formal, non-formal and informal learning in Computer Engineering and Information Technologies undergraduate students' training. The results of research conducted in the 2017/2018 academic year are shown. A total of 106 students earning BA in Computer Engineering and Information Technologies at National University of Life and Environmental Sciences of Ukraine volunteered to participate in the research. The aim of the research was to investigate the influence of combining formal, non-formal and informal learning on improving Computer Engineering and Information Technologies undergraduate students' training through the increase in their general self-efficacy. To collect data from 106 computer engineering and information technologies undergraduate students we applied a mixed methods approach implying the combination of qualitative and quantitative methods. The quantitative data were collected using a pre-test and a post-test based on a questionnaire on general self-efficacy. Qualitative methods included classroom observations and interviews with computer engineering and information technologies undergraduate students. The authors compared the results obtained in the experimental and control groups and drew conclusions concerning the positive effects of combining formal, non-formal and informal learning on improving Computer Engineering and Information Technologies undergraduate students' training through the increase in their general self-efficacy.**

***Keywords*—Computer Engineering and Information Technologies undergraduate students' training, formal learning, general self-efficacy, non-formal learning, informal learning.**

I. Introduction

The intensive development of computer science and information technology due to the invention of the World Wide Web influences all aspects of human life in the digital era. As a result, specialists with jobs in the career fields of computer engineering and information technology are becoming more and more attractive in the eyes of prospective employers. In Ukraine the computer engineering and information technology training is regulated by the Ministry of Education and Science according to which earning a degree in this field requires students to complete at least undergraduate programmes at a higher educational institution. In this way students who are going to be computer engineers and IT professionals acquire hard and soft skills relevant to the range and complexity of tasks to be done at the working place. But are these skills enough to become successful after employment? Is formal learning provided by higher educational institutions enough to achieve the desired outcomes? What influences the strengthening of students' self-efficacy to fulfill various tasks during their training at university? Are students with a strong sense of self-efficacy more successful in acquiring the necessary skills and performing various tasks? These are some of the questions which are most frequently asked and for which educators and scientists all over the world are seeking answers to.

A considerable amount of scientific literature is documenting the fact that traditional learning regarded as formal one is no longer the only sufficient means of improving students' self-efficacy and as a consequence their skills necessary to meet ever changing requirements of the labour market [1], [2], [3], [4]. In almost all countries the standard education systems which regulate the professional training of specialists who will work in various industries have a static nature. It means that these systems do not always keep pace with the rapid development of new human knowledge and as a consequence respond to changing requirements by their up-to-date training. In this regard, the influence of the combination of formal, non-formal and informal learning

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4113
© 2019 Oleksandr Malykhin, Nataliia Aristova.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

on improving computer engineering and information technologies undergraduate students' training through the increase in their general self-efficacy is of particular interest in our research.

The research has the following purposes:

1. To share the experience of combining formal, non-formal and informal learning in computer engineering and information technologies undergraduate students' training at a higher educational institution.

2. To investigate the influence of combining formal, non-formal and informal learning on improving computer engineering and information technologies undergraduate students' training through the increase in their general self-efficacy.

II. The Role of Formal, Non-formal and Informal Learning in Improving Computer Engineering and Information Technologies Undergraduate Students' Training through the Increase in their Self-efficacy

Many scientists are deeply convinced that self-efficacy plays a very important role in our lives. In fact, it was Bandura [5] who was the first one to state that self-efficacy influences people's beliefs on feelings, viewpoint, motivation and behaviour. Likewise, Tsang, Hui & Law [6] specify that functioning as a multilevel and multifaceted set of views self-efficacy denotes people's confidence about their ability to fulfil various tasks. Cherry [7], in her study of self-efficacy, made it very clear that people with a strong sense of self-efficacy consider challenging issues as tasks to be mastered, take a greater interest in the activities they participate in, build a stronger sense of commitment to their interests and activities etc. At the same time Cherry [8] thinks that people whose sense of self-efficacy is weak believe that difficult tasks are beyond their capabilities, focus more on personal failings and negative outcomes, quickly lose confidence in personal abilities.

During professional training students usually face different challenges they have to overcome. Researchers' practical experience shows that students' abilities to overcome these challenges depend directly on the level of their self-efficacy [9], [10]. And although higher educational institutions provide students with various opportunities to improve their self-efficacy, in most cases formal learning cannot fully exert its influence on this process. We strongly believe that in order to maximize the impact of formal learning on improving students' self-efficacy universities have to provide non-formal and informal learning as well. Formal, non-formal and informal learning to different degrees provides four sources for developing self-efficacy beliefs, namely mastery experiences, vicarious experiences, verbal persuasion, emotional and psychological states [11]. Thus, formal learning is more connected with verbal persuasion, emotional and psychological states. It is probable due to the fact that during professional training students are surrounded by the faculty who try to make them think they can cope with any difficult task and they do not have to be afraid of any challenges. Non-formal and informal types of learning are linked with vicarious experiences as watching the success of their peers can play a crucial role in believing in your own abilities to cope with any task no matter how difficult it is.

Bearing in mind information mentioned above we can theorize that combination of formal, non-formal and informal learning influences the improvement of computer engineering and information technologies undergraduate students' training through the increase in their general self-efficacy.

Formal learning, for the purposes of our research, is defined as learning that occurs in an organised and structured environment [12]. It means that provided by different educational institutions it leads to validation and obligatory certification. In its turn, non-formal learning which occurs both at the learners' initiative or as a by-product of educational establishment activities is rather organized and can have learning objectives [13]. It is provided by university partners who in most cases are prospective employers interested in young, active and highly motivated specialists. Unlike the formal and non-formal learning, informal one is unintentional, does not lead to certification and is not organised or structured in terms of objectives, time or learning support [14]. It is often defined as learning by experience as it does not have any objectives in terms of learning outcomes [15]. At higher educational institutions it is provided by scientific circles, clubs, themed contests etc.

At the Faculty of Information Technologies at National University of Life and Environmental Sciences of Ukraine formal learning is provided by five Departments, namely the Department of Computer Science, the Department of Computer systems and networks, the Department of Economic Cybernetics, the Department of Information Systems and the Department of Informational and Distant Technologies.

Non-formal learning is provided by the functioning of IT-Academies. Microsoft Imagine Academy and Cisco Academy enable students to get additional education and be certified as Microsoft Office Specialists, Microsoft Technology Associates, Microsoft Certified Solutions Developers and Microsoft Certified Solutions Experts and Cisco Certified Network Associates.

Four scientific circles which function at the Faculty of Information Technologies provide informal learning for students who earn their degrees in Computer Engineering and Information Technologies. These scientific circles are: the scientific circle in programming, the scientific circle "iTeam", the scientific circle "The Internet of things" and the scientific circle "Cybertonus". The operation of these circles aims at development of students' scientific and creative potential and what is more important at the improvement of their general self-efficacy. The university and faculty administration also organize various themed contests and club activities for students which assist in developing communication, team-working and time-management skills, ability to work under pressure etc.

III. Methodology of Research

Computer Engineering and Information Technologies undergraduate students were selected as a research sample using a convenience sampling technique. The research was conducted at the Faculty of Information Technologies, National University of Life and Environmental Sciences of Ukraine (Kyiv). Experimental data were collected from 106 students earning BA in Computer Engineering and Information Technologies, 54 undergraduate students in the experimental group and 52 undergraduate students in the control group. Thus, 106 selected undergraduate students were informed about all types of activities organized by the faculty and university administration. But 54 computer engineering and information technologies undergraduate students who represented the experimental group were offered to become either members of scientific circles which function at the faculty or members of students' organization. The representatives of the experimental group were actively involved in all activities provided by non-formal and informal learning. The control group included 52 undergraduate students who were provided with formal learning only.

A mixed methods approach implying the combination of qualitative and quantitative methods was used to collect data from 106 Computer Engineering and Information Technologies undergraduate students.

The quantitative research used a pre-test and a post-test based on a questionnaire on general self-efficacy created by R. Schwarzer & M. Jerusalem [16]. All the respondents were asked to rate the statements according to a 4-point scale (not at all true – 1 point, hardly true – 2 points, moderately true – 3 points and exactly true – 4 points) at the beginning and the end of experiment. In order to get a sum score we added up all the responses supplied by the respondents. The obtained results ranging from 10 to 26 points showed the low level of general self-efficacy, from 27 to 35 points – medium level of general self-efficacy and from 36 and higher – high level of general self-efficacy.

Statements for assessment

1. I can always manage to solve difficult problems if I try hard enough.
2. If someone opposes me, I can find means and ways to get what I want.
3. It is easy for me to stick to my aims and to accomplish my goals.
4. I am confident that I could deal efficiently with unexpected events.
5. Thanks to my resourcefulness, I know how to handle unforeseen situations.
6. I can solve most problems if I invest the necessary effort.
7. I can remain calm when facing difficulties because I can rely on my coping abilities.
8. When I am confronted with a problem, I can usually find several solutions.
9. If I am in a bind, I can usually think of something to do.
10. No matter what comes my way, I am usually able to handle it.

The qualitative methods used in the research comprised observation and interviews with Computer Engineering and Information Technologies undergraduate students. Like the quantitative methods, the qualitative ones made a great contribution to understanding the role of formal, non-formal and informal learning on increasing respondents' general self-efficacy and respondents' attitude towards participating in activities provided by three types of learning during Computer Engineering and Information Technologies professional training.

Two weeks prior to the beginning of the experimental work we at first interviewed 106 Computer Engineering and Information Technologies undergraduate students to understand their attitude towards opportunities provided by combining formal, non-formal and informal learning during their professional training. At second, all the respondents were asked to respond to the questionnaire developed by R. Schwarzer & M. Jerusalem [17].

After interviewing and initial assessment of general self-efficacy the experimental and control groups were formed and the experimental work started. During the experimental work 54 Computer Engineering and Information Technologies undergraduate students were trained using the combination of formal, non-formal and informal learning whereas 52 students by means of formal learning only. To achieve the main aim of the research two hypotheses were formulated:

H 1: The combination of formal, non-formal and informal learning influences the improvement of Computer Engineering and Information Technologies undergraduate students' training through the increase in their general self-efficacy.

H 2: Formal learning which constitutes the standard learning required by the Ministry of Education and Science of Ukraine for earning a degree in the spheres of computer engineering and information technology provides the substantial increase in undergraduate students' general self-efficacy during their training.

IV. Results and Discussion

The results of the interviews carried out at the beginning of the experiment demonstrate that the students' opinions on combining formal, non-formal and informal learning during their training are different. Thus, 37,74% (n=40) of respondents stated that formal learning is not enough to become a real professional. These respondents shared the common point of view:

Personally I think that formal learning is not enough for students who want to be good specialists in such spheres as computer engineering and information technology. These spheres are being developed rather quickly which means that knowledge and skills you obtain at university can become obsolete in a very short period of time. That is why during our training we have to pay more attention to all opportunities which are provided by university and which influence our self-development. There are different activities we can take part in beyond our classes. Some activities enable us to acquire additional knowledge and

skills which are essential for becoming real professionals and which help us improve our confidence. But despite the fact that faculty administration always provides us with comprehensive information on seminars, webinars and various contests, it is your own choice to take part in them or not. It's a pity that some of us do not seize these opportunities.

28,30% (n=30) of all the respondents shared the opinion that

Although during our training we are given opportunities to obtain additional training which will enhance our employability. For example, participating in scientific circles can improve not only our practical skills and intellectual abilities but also our abilities to work in a team, to cooperate and collaborate with other team members. Moreover, we can get some certificates completing different courses but you have to pay extra money to get them.

33.96% (n=36) of all the respondents argued that

To my mind, studying at university gives you basic background for being a good specialist in any sphere of human life. Skills and knowledge one can obtain at university is enough to start a career at any private or public organization. While you are a student it is difficult to understand what skills will be necessary at the working place. Another thing is that even if you get additional qualifications while you are training at university it doesn't mean you might need them in the future. You will understand what skills you lack when you start working. Only in this case you have to be involved in non-formal learning. As for informal learning, although university provides us with various opportunities, to tell the truth, it takes much time and effort to participate in activities after classes.

The analysis of respondents' ideas on incorporating the combination of formal, non-formal and informal learning into professional training of Computer Engineering and Information Technologies undergraduate students expressed during their interviewing enabled researchers to form maximally homogenous experimental and control groups. As it was mentioned above there were 54 Computer Engineering and Information Technologies undergraduate students in the experimental group and 52 Computer Engineering and Information Technologies undergraduate students in the control one. A fairly equal distribution of respondents who expressed different points of view on combining formal, non-formal and informal learning during their professional training enabled researchers, on the one hand, to do the experiment in the natural conditions and, on the other hand, to consider their attitudes towards additional opportunities provided by non-formal and informal learning.

The results of data analysis related to finding out the levels of general self-efficacy by the experimental and control groups at the beginning of the experiment are presented in Table 1.

Table 1 Respondents' general self-efficacy levels at the beginning of the experiment

Distribution of Respondents	Respondents' general self-efficacy levels (N)		
	Low N %	Medium N %	High N %
Experimental group (54)	17 31.48%	33 61.11%	4 7.41%
Control group (52)	15 28.85%	32 61.53%	5 9.62%

As we can see in Table 1, there was no clear difference in computer engineering and information technologies undergraduate students' general self-efficacy levels at the beginning of the experiment. Thus, 31.48% (n=17) of respondents in the experimental group and 28.85% (n=15) of respondents in the control group had a low level of general self-efficacy. Medium level of general self-efficacy was shown by 61.11% (n=33) of the experimental group respondents and 61.53% (n=32) of the control group respondents. High level of general self-efficacy was shown by 7.41% (n=4) of respondents in the experimental group and 9.62% (n=5) of respondents in the control group.

The results of data analysis related to finding out the levels of general self-efficacy by the experimental and control groups at the end of the experiment are presented in Table 2.

Table 2 Respondents' general self-efficacy levels at the end of the experiment

Distribution of Respondents	Respondents' general self-efficacy levels (N)		
	Low N %	Medium N %	High N %
Experimental group (54)	0 0%	44 81.48%	10 18.52%
Control group (52)	10 19.23%	37 71.15%	5 9.62%

As it is shown in Table 2 at the end of the experiment the experimental group whose formal learning at the Faculty of Information Technologies was combined with non-formal and informal learning showed higher levels of general self-efficacy that the control group. Thus, low level of general self-efficacy was demonstrated by 0% (n=0) of respondents in the experimental group, medium level of general self-efficacy by 81.48% (n=44) of respondents and high level by 18.52% (n=10) of respondents. Unlike the experimental group, the results shown by the control group were considerably lower. 19.23% (n=10) of respondents in the control group had a low level of general self-efficacy, 71.15% (n=37) – medium level and only 9.62% (n=5) – high level.

The results of data analysis regarding comparison of the general self-efficacy levels by the experimental and control groups at the beginning and the end of the experiment are given in Table 3.

TABLE 3 THE COMPARISON OF RESPONDENTS' GENERAL SELF-EFFICACY LEVELS AT THE BEGINNING AND THE END OF THE EXPERIMENT

Distribution of Respondents		Respondents' general self-efficacy levels					
		The beginning of the experiment			The end of the experiment		
		Low	Medium	High	Low	Medium	High
Experimental group (54)	N	17	33	4	0	44	10
	%	31.41	61.11	7.41	0	81.48	18.52
Control group (52)	N	15	32	5	10	37	5
	%	28.85	61.53	9.62.	19.23	71.15	9.62

By comparing the results obtained at the beginning and the end of the experiment, we can state that if before the experimental work a low level of general-self efficacy was shown by 31.48% (n=17) of respondents in the experimental group and 28.85% (n=15) of respondents in the control group after the experimental work it was demonstrated by 0% (n=0) of respondents in the experimental group and 19.23% (n=10) of respondents in the control group. At the beginning of the experiment a medium level of general self-efficacy was demonstrated by 61.11% (n=33) of the experimental group respondents and 61.53% (n=32) of the control group respondents while at the end of the experiment it was shown by 81.48% (n=44) in the experimental group and 71.15% (n=37) in the control group. At the beginning of the experiment 7.41% (n=4) of respondents in the experimental group and 9.62% (n=5) of respondents in the control group showed a high level of general self-efficacy whereas at the end of the experiment a high level of general self-efficacy was shown by18.52% (n=10) of respondents in the experimental group and 9.62% (n=5) of respondents in the control group. Thus, taking into account the obtained results we can assert that there has been a significant increase in general self-efficacy levels by the experimental group. Moreover, the results of the research confirm our idea about necessity of incorporating non-formal and informal learning into computer engineering and information technologies undergraduate students' training. Such a combination of formal, non-formal and informal learning increased computer engineering and information technologies undergraduate students' general self-efficacy and as a consequence improved their training.

Thus, the first hypothesis concerning the influence of combining formal, non-formal and informal learning on the improvement of computer engineering and information technologies undergraduate students' training through the increase in their general self-efficacy was confirmed.

The second hypothesis that formal learning provides the substantial increase in computer engineering and information technology undergraduate students' general self-efficacy during their training was rejected.

V. CONCLUSION

The main purposes of the research were to share the experience of combining formal, non-formal and informal learning in computer engineering and information technologies undergraduate students' training at a higher educational institution and to investigate the influence of combining formal, non-formal and informal learning on improving computer engineering and information technologies undergraduate students' training through the increase in their general self-efficacy. The obtained findings showed that undergraduate students involved in additional activities provided the combination of formal, non-formal and informal learning at the Faculty of Information Technologies benefited greatly in increasing their general self-efficacy.

The research concluded that providing formal, non-formal and informal learning in isolation denied computer engineering and information technologies undergraduate students the possibility of substantial increase in their general self-efficacy. Formal learning provides students with skills and competences necessary for their employability but it is not enough to teach them to cope with all the difficulties they face. Understanding what way the incorporation of non-formal and informal learning into the formal one relates to the increase of general self-efficacy allows the faculty to benefit fully in improving computer engineering and computer technologies undergraduate students' training.

Although educators and researchers dealing with improvement of undergraduate students' training through the increase in their self-efficacy can take into consideration the implications of the present study, further research is recommended in this area.

REFERENCES

[1] M. E. Gist, " The Effects of Self-Efficacy Training on Training Task Performance," Academy of Management Proceedings, vol. 1986, no. 1, pp. 250-254, 1986.

[2] O. Malykhin and N. Aristova, "Investigation into participation activity rate of foreign languages teachers in specially focused network pedagogical communities," The New Pedagogical Review, vol. 53, no. 3, pp. 227-238, 2018, DOI: 10.15804/tner.2018.53.3.19

[3] Recognition of Non-formal and Informal: Learning Pointers for policy development, March 2010. [Online]. Available: file:///C:/Users/Asus/Desktop/non-formal%20and%20formal%20learning.pdf [Accessed: Dec. 22, 2018].

[4] M. Radaković and S. Antonijević, " Balance between formal and informal learning − experience and challenges of civil servants training in Serbia", 2013. [Online]. Available: http://www.nispa.org/files/conferences/2013/papers/201304151202190.Paper_Radakovic.pdf. [Accessed: Dec. 12, 2018].

[5] A. Bandura, "Perceived Self-Efficacy in Cognitive Development and Functioning,". Educational Psychologist, vol. 28, no. 2, pp. 117-148, 1993.

[6] S. K. M. Tsang, E. K. P. Hui and B. C. M. Law, "Self-Efficacy

as a Positive Youth Development Construct: A Conceptual Review", The Scientific World Journal, April 2012. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3351108/ [Accessed: Nov. 7, 2018].

[7] K. Cherry, "Self-Efficacy: Why believing in yourself matters," August 2018. [Online]. Available: https://www.verywellmind.com/what-is-self-efficacy-2795954. [Accessed: Dec.10, 2018].

[8] K. Cherry, "Self-Efficacy: Why believing in yourself matters," August 2018. [Online]. Available: https://www.verywellmind.com/what-is-self-efficacy-2795954. [Accessed: Dec.10, 2018].

[9] M. E. Gist, "The Effects of Self-Efficacy Training on Training Task Performance," Academy of Management Proceedings, vol. 1986, no. 1, pp. 250-254, 1986.

[10] K. Cherry, "Self-Efficacy: Why believing in yourself matters," August 2018. [Online]. Available: https://www.verywellmind.com/what-is-self-efficacy-2795954. [Accessed: Dec.10, 2018].

[11] A. Bandura, (1977). "Self-efficacy: Toward a Unifying Theory of Behavioral Change," Psych. Rev., vol. 84, no. 2, pp. 191-215, 1977.

[12] Terminology of European education and training policy. (2008). [Online]. Available: http://www.cedefop.europa.eu/files/4064_en.pdf

[13] Recognition of Non-formal and Informal: Learning Pointers for policy development, March 2010. [Online]. Available: file:///C:/Users/Asus/Desktop/non-formal%20and%20formal%20learning.pdf [Accessed: Dec. 22, 2018].

[14] Terminology of European education and training policy. (2008). [Online]. Available: http://www.cedefop.europa.eu/files/4064_en.pdf

[15] Recognition of Non-formal and Informal: Learning Pointers for policy development, March 2010. [Online]. Available: file:///C:/Users/Asus/Desktop/non-formal%20and%20formal%20learning.pdf [Accessed: Dec. 22, 2018].

[16] R. Schwarzer and M. Jerusalem, "Generalized Self-Efficacy scale," in J. Weinman, S. Wright, & M. Johnston, Measures in health psychology: A user's portfolio. Causal and control beliefs, Windsor, England: NFER-Nelson, 1995, pp. 35-37.

[17] R. Schwarzer and M. Jerusalem, "Generalized Self-Efficacy scale," in J. Weinman, S. Wright, & M. Johnston, Measures in health psychology: A user's portfolio. Causal and control beliefs, Windsor, England: NFER-Nelson, 1995, pp. 35-37.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 214-217

Mechatronics Education: Needs and Challenges

Gilberto Marzano
Rezekne Academy of Technologies
Rezekne, Latvia
gilberto.marzano@rta.lv

Andris Martinovs
Rezekne Academy of Technologies
Rezekne, Latvia
andris.martinovs@rta.lv

Svetlana Usca
Rezekne Academy of Technologies)
Rezekne, Latvia
svetlana.usca@rta.lv

***Abstract*— Mechatronics is a multidisciplinary branch that combines electronics, mechanics, and computer science. Due to the digital revolution, mechatronics is growing fast, rapidly spreading from manufacturing to many new sectors such as agriculture, healthcare, security, and transportation. As a consequence, mechatronics training courses are proliferating at the undergraduate and graduate levels. Nowadays, a considerable number of educational institutions, both academic and vocational, are offering mechatronics education, however their programs are not homogeneous and depend on teaching staff competence and available resources.**

This paper presents NewMetro, an EU funded project that aims to develop an innovative European framework of competences for mechatronics education and an educational model able to address the needs of young people as well as adult workers undergoing requalification programs.

The philosophy of the project and the first activities carried out are illustrated and discussed.

Keywords—Education 4.0, Mechatronics, Mechatronics Education, Smart Learning Environments.

I. Introduction

The concept of mechatronics was introduced nearly 50 years ago as a consequence of the increasing use of computers for the control of mechanical processes and systems [1]. The term derives from the contraction of the words mechanical and electronic. It was coined in the early 1970s by Tetsuro Mori, an engineer from the Yaskawa Electric Corporation of Japan [2], to pinpoint synergetic systems composed of mechanical and electrical elements.

From the1980s, mechatronics grew as a multidisciplinary branch of engineering encompassing a mix of electronics, computing, telecommunications, and systems engineering. The basic components of mechatronics are sensors, microcontrollers, and actuators, as well as real-time software.

Over the last few years, interest in mechatronics has been reinvigorated due to the advances in sensing, communication, and computing. The extension of internet connectivity to physical devices has brought to the spread of the Internet of Things (IoT), a notion that encompasses everything is connected to the internet. Devices equipped with smart sensors can communicate with each other as well as with people who use smartphones or wearables [3], [4].

Many of the smart components associated with the IoT are mechatronic since the increasing complexity of the IoT applications entails a notable integration of the mechanical domain with the information technology and electronics domains.

Given that mechatronics is a multidiscipline and represents the combination of various systems, its scope is very vast and relates to multiple fields and domains [5], [6]:

- The Medical field: in areas such as surgery, radiology, and emergency medicine;
- Robotics Industry: for industrial robots and robotics systems;
- Automotive/Automobile engineering: in the design and manufacture of motorcycles, automobiles, and trucks that integrate mechanical, electrical, electronic, and software technologies as well as safety engineering;
- Research Organizations: with uses in instrumentation and sensors, microfluidic systems and MEMS (Micro-Electro-Mechanical Systems), and energy conversion;
- Mechanical Industry: with uses in designing, analyzing, manufacturing, and maintaining mechanical systems;
- Computer-Aided Design (CAD): with the use of computer systems to support and optimize mechanical design.
- Manufacturing Industry: in the production of industrial goods, machines, or tools;
- Mining: with uses in the extraction of valuable minerals or other geological materials from the earth;
- Inspection: for uses such as the inspection oil and gas pipelines via drones.

The increasingly numerous scope of applications of mechatronics brings us to the problem of what competencies are required to be able to work in so many different fields.

What are the competencies that should be included in an effective mechatronics curriculum?

Definitely, these competencies must relate to the six Key Enable Technologies (KETs) indicated, in 2014, by the European High-level Expert Group (HLG) in the Sta-

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4199
© 2019 Gilberto Marzano, Andris Martinovs, Svetlana Usca.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

tus Implementation Report of the second Key Enabling Technologies [7]. In this regard, the HLG observes that new technologies cannot be developed and brought to the market if the European work-force does not possess the necessary skills. In addition, the rapidly growing market in KETs related sectors requires an increasing number of professionals at all technical levels and from multiple disciplines. In this respect, the first HLG KET report noted that Europe is facing a damaging shortage of skilled labor qualified to master the multi-disciplinary nature of KETs.

Accordingly, a mechatronics curriculum should provide the basis for innovation in a wide range of sectors such as the automotive industry, food, chemicals, electronics, energy, pharmaceuticals, construction, and telecommunications.

Furthermore, this curriculum should be flexible and it should be continuously updated to keep abreast of continuing technological advances.

This paper presents NewMetro, an EU funded project that aims to develop an innovative European framework of competences for mechatronics education and an educational model able to address the needs of young people as well as adult workers undergoing requalification programs.

II. The NewMetro project

The structure of manuscripts:

The NEWMETRO project is a three-year Erasmus+ Sector Skills Alliance project that started in December 2018. It involves 11 partners from 7 European countries (Italy, Latvia, Austria, Germany, Poland Greece, and Spain).

This project is articulated around three different layers:

1. Educational and training layer, where partners design a model for mechatronics education, develop a curriculum, implement training activities, and run pilots. They involve enterprises with a view to individuating the mechatronics competency model, and evaluate the results of the pilot projects. They adopt a learning strategy that engages learners in work-based learning activities including ubiquitous learning, online collaborative learning, and digital social learning.
2. Inter-organizational layer, where partners, associated partners, and selected stakeholders cooperate in disseminating the project results. They bring about actions and organize events to show how remodeled mechatronics education can address the needs of Industry 4.0.
3. Policy-making layer, where the project results and recommendations are illustrated and discussed with policymakers and governmental authorities to frame new policies directed at sustaining continuous learning on KETs and offering workers flexible and advanced learning environments.

Currently, partners are engaged in the phase of Work package 1 "Participatory review and validation of already available training need analysis", led by The Rezekne Academy of Technologies. Their activities aim at:

- Understanding, sharing, and updating the current European surveys in the mechatronic field;
- Analyzing competence and related professional profiles;
- Providing a methodological framework to support and facilitate project implementation.

Partners should collect data from statistical sources, literature analysis, and structured interview. At this time, the literature analysis has already been carried out while interviews are still ongoing. Each partner is required to interview 50 key-persons involved in industrial activities (30 workers, 10 entrepreneurs or managers of enterprises, and10 policymakers) in order to create a comparative multidimensional matrix crossing the data.

The literature analysis has been performed on articles available in various academic databases (Scopus, Web of Science, Science Direct, and IEEE Xplore) and selected scientific journals. We adopted the PRISMA methodology [8].

Statistical data has only been collected from official sources and scientific surveys.

The following paragraphs report the preliminary results concerning the competence for a new mechatronics curriculum and the indications emerging for a teaching-learning environment.

.

III. Competencies for mrchatronics education

To design the competency model for mechatronics, we carefully analyzed the European multilingual classification, ESCO (European Skills, Competences, Qualifications, and Occupations) [9]. ESCO provides descriptions of 2,942 occupations and 13,485 skills linked to those occupations, translated into 27 languages (all the official languages of the EU plus Icelandic, Norwegian, and Arabic).

For mechatronics, ESCO indicates robotics as a key knowledge. Robotics is defined as a part of mechanical engineering, electrical engineering, and computer science that overlaps with mechatronics and automation engineering.

The skills and competence for a mechatronics engineer given by ESCO [10] are:

- maintain mechatronic equipment;
- test mechatronic units;
- install mechatronic equipment;
- develop mechatronic test procedures;
- simulate mechatronic design concepts;
- calibrate mechatronic instruments;
- assemble mechatronic units;

- micro-mechatronic engineering.

From our analysis, the mechatronics competences should primarily encompass the following areas, since they are the most involved in Industry 4.0 and the ongoing technological revolution:

- Industrial design using advanced materials (e.g. biomaterials, metals, ceramics, polymers, powders);
- Microelectronics applied in the mechanical systems sector;
- Assembly line management (automation, supervision, measurement, data transmission and storage, etc.);
- Capabilities in the mechanical systems sector;
- LCA (Life Cycle Approaches) prognostics and environmental footprint evaluation;
- Integrating advanced systems design and manufacturing (modeling, simulation, virtual testing, data management, etc.);
- Advanced manufacturing systems (logistics management, cloud manufacturing, etc.);
- New equipment for telecare, telemedicine, and telerehabilitation based on intelligent mechanical objects;
- Domestic robots based on the Internet of Things paradigm;
- Applications in the scope of Smart and Connected Communities.

In some countries, standards for the profession of mechatronics engineer and mechatronics technician have been approved. They define knowledge, skills and competence required by these professionals. In Latvia, the first standard for the profession of mechatronics engineer was approved in 2002. [11]. In 2006, a new version that is currently being used was approved [12].

Focusing on the above areas, we create a preliminary list of skills and competencies based on the 2010 standard:

- designing algorithms for automation processes and developing tasks for machine design;
- using software for machine control;
- monitoring mechatronic machines;
- detecting failures, repairing, and maintaining automated machines;
- defining production technology plans;
- assessing production automation levels;
- choosing suitable materials for machine construction;
- developing software for automated controls;
- determining the lifetime of a mechatronic system;
- designing processes to ensure long-term quality operability of machines;
- assessing economically advantageous technical solutions;
- planning timely completion of tasks;
- managing staff;
- ensuring environmental, health, and safety requirements and standards.

We utilized the above lists in the structured interview to solicit feedback from stakeholders.

IV. A SMART TEACHING LEARNING ENVIRONMENT

Our preliminary analysis confirmed that, in order to create a robust educational framework for mechatronics, a smart teaching-learning environment should be implemented.

The adjective 'smart' is very popular nowadays. In everyday usage, it refers to an action or decision that involved careful planning, cleverness, innovation, and/or a desirable outcome [13]. In the field of technology, instead, it is a label used in a wide variety of ways to underlie the combined use of intelligent devices, advanced analytics, and artificial intelligence algorithms (Fig. 1).

Lately, smart has also been associated with educational practice: smart schools, smart classrooms, and smart learning technologies are spawning seemingly everywhere. In the field of education, SMART is also an acronym for Sensitive, Manageable, Adaptable, Responsive, and Timely, and refers to interactive technology that has a flexible and tailored approach to meet diverse individual learning needs.

A smart learning environment for mechatronics (Fig. 1) should encompass intelligent tutoring, adaptive learning, technology-enhanced learning, mobile learning, and context-aware ubiquitous learning [14], [15]. Sensing technologies, as well as virtual reality, should enable learners to access learning resources that simulate real situations. Moreover, supportive tools should be implemented to provide them with contextual guidance and hints.

In essence, then, the smart learning environment for mechatronics education should be a technology-enhanced teaching-learning system that simulates the real-world, allows access to different types of resources, provides collaborative functions, and can be easily adapted for work-based learning.

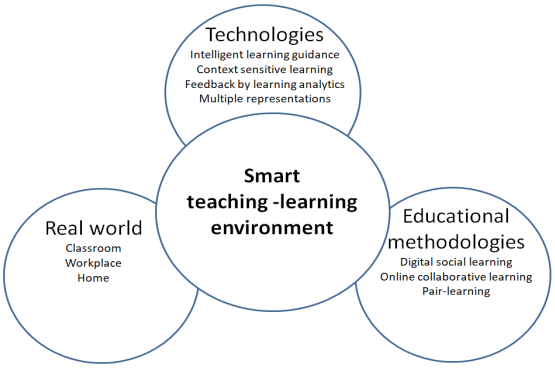


Fig. 1. Components of a smart learning environments

The primary criteria identified in the literature for smart learning environments are:

1. Context-sensitiveness; that is, learning activities are lead according to the real-world environment in which the learner is (e.g., classroom, workplace, home);
2. Adaptive support; this is, learners are led by immediate analyses of their needs from various different perspectives (e.g., learning performance, learning behaviors, profiles, personal factors) as well as the online and real-world contexts in which they are situated;
3. Flexible user interface; this is, information is presented that meets the learner's personal factors (e.g. learning styles and preferences) and their learning status (e.g. learning performance).

V. Conclusion

It has been observed that: The collaboration between science and technology, and the mechatronic integration of technology make possible the biggest and fastest evolution of humanity in the history of technology [16].

Mechatronics creates new complex products using and integrating the discoveries of other disciplines. As a consequence, mechatronics can be found embedded in most of the new advanced products today. Appropriately, integration should also be a fundamental characteristic of mechatronics education.

It should not only encompass the skills and competencies of a mechanical engineer, a computer engineer, and an electrical engineer, but also adopt a novel educational model to prepare the students to meet the new industrial and societal needs.

Traditional courses in engineering as well as the mechatronic courses currently on offer are compartmentalized and taught by an individual instructor. In the near future, this approach will be inadequate since the solution of real problems invariably requires integrating different subjects and disciplines, both technical and non-technical [17].

This is very challenging. Indeed, it immediately appears evident that the current distinct engineering modules should be revised and improved in order to bring about a seamless .transition from university to industry and society.

The NEW METRO project tackles this challenge. It is still in its early stages, but the crucial role of the teaching-learning environment has appeared evident.

In this article, we have highlighted the smart teaching-learning environment that should support the new mechatronics curriculum, ensuring two primary needs. Firstly, allowing the continuous adaptation of the technological components and the learning content. Secondly, fostering the minds-on and hands-on capacities of learners through teaching-learning methodologies such as Digital Social Learning, Pair-learning, and Online Collaborative Learning.

References

[1] D. Bradley, D. W. Russell, I. Ferguson, J. Isaacs, A. MacLeod and R. White, "The Internet of Things–The future or the end of mechatronics". Mechatronics, *27*, pp. 57-74, 2015.
[2] D. Bradley, D. W. Russell, Mechatronics in action, Springer, 2010.
[3] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications", IEEE communications surveys & tutorials, 17(4), pp. 2347-2376, 2015.
[4] P. P. Ray, P. P. "A survey on Internet of Things architectures", Journal of King Saud University-Computer and Information Sciences, 30(3), pp. 291-319, 2018.
[5] S. E. Lyshevski, Electromechanical systems, electric machines, and applied mechatronics, CRC press, 2018.
[6] O. O. V. Villegas, M. Nandayapa and I. Soto (Eds.). Advanced Topics on Computer Vision, Control and Robotics in Mechatronics, Springer, 2018.
[7] Expert Group on Key Enabling Technologies (KETs), Status Implementation Report July 2013, European Commission, 26/11/2014. Available at: https://ec.europa.eu/growth/content/status-implementation-report-%E2%80%93-second-high-level-expert-group-key-enabling-technologies-kets_en; last accessed 1.04.2019.
[8] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, "The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement", PLoS Med 6(7): e1000097, 2009. Available at: https://doi.org/10.1371/journal.pmed.1000097; last accessed 1.04.2019.
[9] ESCO (European Skills/Competence, qualification and Occupations), https://ec.europa.eu/esco/portal/; last accessed 01.04.2019.
[10] ESCO Portal, Mechatronics. Available at: https://ec.europa.eu/esco/portal/search?term=mechatronics; last accessed 01.04.2019.
[11] Ministry of Education and Science, "PROFESIJAS STANDARTS.", n. 405, July 2002. Available at: https://visc.gov.lv/profizglitiba/dokumenti/standarti/ps0097.pdf; last accessed 01.04.2019.
[12] Government cabinet, "2.pielikums Ministru kabineta", n. 461, May 2000. Available at: http://www.lm.gov.lv/upload/darba_devejiem/profesiju_standarti_0811.pdf; last accessed 01.04.2019.
[13] J. M. Spector (Ed.), The SAGE encyclopedia of educational technology, Sage Publications, 2015.
[14] R. Huang, J.Yang and Y. Hu, "From Digital to Smart: The Evolution and Trends of Learning Environment", Open Education Research, 1,pp. 75-84, 2012.
[15] G. J. Hwang, "Definition, framework and research issues of smart learning environments-a context-aware ubiquitous learning perspective", Smart Learning Environments, *1*(1), 4,2014.
[16] I. Vlașin, S. Greta, L. Dache and V. Mătieș, "Integrative platform based on the mechatronics model for educational technologies focused on competence" In IOP Conference Series: Materials Science and Engineering (147)1, IOP Publishing, August 2016. Available at: https://iopscience.iop.org/article/10.1088/1757-899X/147/1/012054/pdf; last accessed 01.04.2019.
[17] Y. Wang, Y. Zhu, Y. Yu, X. Zhang, and C. Xie, C, "Simulating Industry: A Holistic Approach for Bridging the Gap between Engineering Education and Industry. Part II: Practice in Mechatronics Engineering", International Journal of Engineering Education, 31(1), pp. 174-180, 2015.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 218-221

The Personification of the User's Interface: Classification vs. Clusterization of Users of Online Courses

Nataliia D. Matrosova
Faculty of Software Engineering and Computer Systems
Saint-Petersburg National Research University of Information Technologies, Mechanics and Optics
Saint-Petersburg, Russian Federation
ndmatrosova@corp.ifmo.ru

Dmitry G. Shtennikov
Faculty of Software Engineering and Computer Systems
Saint-Petersburg National Research University of Information Technologies, Mechanics and Optics
Saint-Petersburg, Russian Federation
dshtennikov@corp.ifmo.ru

***Abstract*—Researchers compared the classification and the clusterization of users of online course for the personification of the users' information system interface. When interacting with control and information systems, users may manifest individual features, including implicit characteristics that may affect one's results within the system. At the same time due to information system building peculiarities one of the most comprehensive statistics can be collected via e-learning systems. When using a course, the user leaves a wide trail of activity that may contain different information depending on the learning environment structure. Online blended learning courses draw the researcher's attention to the impact of digital teaching models on students as well as its ability to adjust distant learning courses to individual students' needs and differences.**

Information personalization is a highly relevant content presentation at the most individual level. Therefore, the task of personalization is to show users information that meets their needs and interests. Personalization gives the opportunity to focus on points that have real value for users.

***Keywords*—machine learning, dataset, classification, clusterization, personification.**

I. Introduction

Firstly, it is important to clarify the term *user*. A user of information system is a specialist in the system's subject area, for whom the system was created to satisfy his informational needs.

The information system user interface consists of various elements. With a large amount of information, the user's in-system work efficiency decreases, due to attention diffusion. Through changes in the graphical part of the interface it is possible to implement recommendation modules. This change in interface is called user interface personification.

It is more efficient to carry out interface upgrades for each individual user than to personalize the interface for all users at once. However, with a large number of users there may be too many adapted options to store. The solution to this challenge is grouping users and conducting personification procedures for each group (cluster) separately.

The task of personification is to display information that meets the needs and interests of users. Personification allows you to focus users on important details. The simplest type of interface personification is a ranked list of items [1].

To achieve the best results in personification it is wise to use not only user explicit characteristics but also non-explicit ones, such as circadian rhythms for example. There are detailed studies of the effects of circadian rhythm on learning [2, 3].

Interface personification is based on algorithm adaptation, which describes rules by which interface changes depending on the user's actions. Among the characteristics of the adaptive algorithm, the following are emphasized: prediction of accuracy, predictability of adaptive behavior, and frequency of interface changes [4].

In this investigation the authors considered only the question of comparing classification and clustering for user interface personification on how to exert the influence of the circadian rhythm on learning. It did not address the question of whether all users from each identified class really require different user interface compared to users from another class.

A. Literature review

The possibilities of interface characterization for information systems, especially for training information systems, are discussed in detail in various studies.

The question regarding the characteristics of the

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4080
© 2019 Nataliia D. Matrosova, Dmitry G. Shtennikov.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

learner is discussed in detail in Hendrik Dranchsler and Paul A. Kirschner' article [5]. The authors suggest that learner characteristics can be personal (age, gender, maturation, language and etc.), academic (learning goals, prior knowledge, education type and level), social/ emotional (sociability, self-image, mood), cognitive (memory, mental procedures, intellectual skills) [5]. These characteristics are highly individual and vary for each learner.

Christoph Fröschl, Loc Nguyen, and Phing Do in the study [6] described an adaptive system based on the "description of learner's properties" (user model or learner model). In their work "the user model can contain information from two categories: domain specific information (reflects status and degree of knowledge and skills) and domain independent information (may include goals, interests, background and experience, personal traits and demographic information)" [6]. The authors offered to classify the user model into three kinds: stereotype, overlay, and plan models.

In the study of the influence of student characteristics on learning paths and strategies [7] the authors considered the following characteristics of students: prior knowledge, study level, gender, and intrinsic motivation. The results showed that students do indeed follow individual learning paths and some student characteristics are related to their learning paths (gender and prior knowledge did not have an effect, but intrinsic motivation had a stronger influence than prior knowledge) [7].

Using the classification and clustering of users to analyse the characteristics of users are considered Ronald G. Leppan, Johan F. van Niekerk, Reinhardt A. Botha in their study [8]. The authors suggested that online learning design should be informed by behavioural patterns. And learner characteristics are inferred using data analysis. The classification used as "predictive modelling to model something that cannot be directly observed by using readily available features as input" [8], and the clusterization - "structure discovery to find patterns in data that are not obvious" [8].

The authors suggest in further research to experiment with the personification of the user's interface to confirm this theoretical research.

II. Materials and methods

This research uses the dataset from one massive open online course (MOOC) from the national open education platform of the Russian Federation as source data.

The set contains data from one batch of students (spring 2018). The students were offered to study learning materials (lecture in video format), complete an after video mini-assessment, working in a virtual laboratory.

The data set contains over 900 000 logs from students and their activities with approximately 90 features. User classification and clusterization were compared using students' latent features and course success rates.

The equation for course user success can be written as Eq.1:

$$s = \sum_1^i \frac{g_i}{k} + \frac{\sum_1^l g_{ml}}{l} + \frac{\sum_1^p g_{fp}}{p} \quad (1)$$

Where s – course success, – after video mini-assessments grade i, –midterm grade, - final test grade, k, l, p – the number of examination passing tries (mini, midterm, and final accordingly).

A. Research data

After data preparation (null rows and Nan values deletion, choosing features), about 1450 users and 5 features were chosen (see Table 1).

It can be noted in Figure 1 that there is a weak correlation between features, but a strong one between weekday_video and weekday_lms, and also hour_video and hour_lms at this stage of the study.

TABLE I. choosing features from the dataset

Name features	Description
w e e k d a y _ video	the highest day of the week lection materials activity (lecture-activity weekday)
hour_video	the highest hour of a day lection materials activity (lecture-activity hour-day)
weekday_lms	the highest day of the week virtual laboratories activity (lms-activity weekday)
hour_lms	the highest hour of a day virtual laboratories activity (lms-activity hour-day)
grade	course student success rate
class	course user group (class) success (target variable for classification)

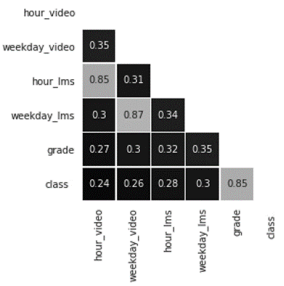


Fig. 1. Correlation between features

Figure 2 presents *the* distribution of features. It can be seen that most features do not have normal distribution.

B. Classification

Classification – the process of streamlining or distributing objects (observations) into classes in order to reflect relations between them [9].

Calculations were carried out for some classification models in this part of the investigation:

- classification and regression trees (CART) - solves classification and regression problems by building a decision tree;
- k-nearest neighbors algorithm (k-NN) – assigns the object to the class that is most common among its k-neighbors whose classes are already known;

- linear regression - estimates coefficients of the linear equation containing one or more independent variables, allowing better value prediction of the dependent variable;
- support vector machines (SVM) – has a special feature which is a continuous decrease in empirical classification error and an increase in the gap. The main idea of the method is a translation of initial vectors into space of higher dimension and search for separating hyperplane with the maximum a gap in that space. Fit time complexity is more than quadratic with the number of samples which makes it hard to scale to the dataset with more than a couple of 10000 samples. Therefore, the authors will not use this model further in this study;
- logistic regression - linear classifier construction method, which allows posterior probabilities evaluation of objects belonging to classes [10];
- Bayesian classification approach - is based on a theorem stating that if the distribution densities of each class are known, then the desired algorithm can be written in explicit analytical form. Moreover, this algorithm is optimal, as it has minimal error probability [10].

Table 2 was compiled as a result of the construction of all the above models. This table contains information about mean accuracy on given test data and labels. Note that linear regression contains a prediction coefficient of determination .

Also note that models were separately checked several times on test samples to avoid possible overfitting problems.

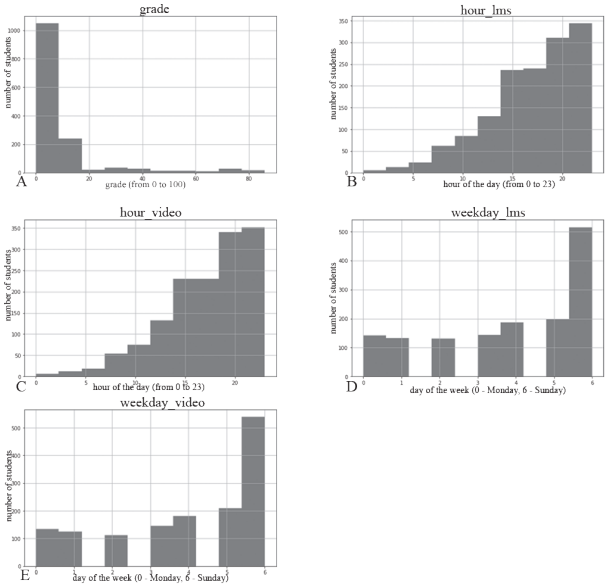


Fig. 2. Distribution of features: A – grade, B – hour_lms, C – hour_ video, D - weekday_lms, E - weekday_video

C. Clusterization

Clusterization (or cluster analysis) – the task of breaking up a set of objects into groups called clusters [11]. Clusterization involves the selection of compact, separate groups of objects characterized by internal homogeneity and external isolation.

In this investigation the authors used the best known method of clusterization - kMeans. To select an appropriate number of clusters, usually the number of clusters chosen from which (the sum of squares of distances from points to centroids of clusters to which they belong, see Eq.2) ceases to decrease sharply (Figure 3). In this example, the number of clusters is 4.

$$J(C) = \sum_{k=1}^{K} \sum_{i \in C_k} \|x_i \quad \mu_k\|^2 \rightarrow \min_C \tag{2}$$

TABLE II. MEAN ACCURACY ON THE GIVEN TEST DATA AND LABELS

Approach	Mean accuracy
classification and regression trees (CART)	0.91
k-nearest neighbors algorithm	0.86
linearregression	0.73
support vector machines (SVM)	0.83
logistic regression	0.66
Bayesian classification approach	0.84

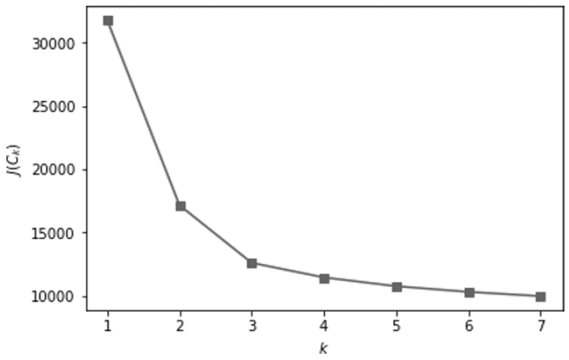


Fig. 3. The graph to define the number of clusters - the sum of the squares of distances from points to centroids of clusters to which they belong

```
[[0.46 -0.14 0.38 -0.88 -0.54]
 [-0.98 -0.08 -0.82 -0.18 -0.30]
 [0.04 0.00 0.01 0.05 1.51]
 [0.34 0.20 0.31 0.98 -0.50]]
```

Fig. 4. Coordinates of centroids

Coordinates of 4 cluster centers were received (Figure 4) as well as an additional column to data, containing information about cluster number to each user_id (Figure 5).

III. RESULTS AND DISCUSSION

Approaches to personification interfaces can be divided into two types: stereotypical and individual [12].

Stereotypical approach states that interfaces are

collected for several users' classes; the system classifies a user and provides one of these interfaces. Usually, several user models are created. For this situation and for predictive purposes (such as defining a user class) classification can be used. For example, the allocation of a user to a certain class based on one's activity and success.

As can be seen from Table 2, classification models give good predictive results. Especially, classification approaches and regression trees.

An individual approach personalizes the interface to improve the layout for a specific user (or a group of users with similar characteristics) based on behavioral data. Therefore, there are individual statistics on in-system actions for each group.

As authors stipulated above, classification and clusterization can address various tasks, as summarized in Table 3.

user_id	hour_video	weekday_video	hour_lms	weekday_lms	grade	cluster
300.0	13	3	13	3	0	1
429.0	15	6	16	6	13	2
525.0	20	6	20	6	2	2
715.0	23	5	23	5	6	2
853.0	14	6	14	6	4	1
1206.0	20	6	20	6	23	2
1874.0	21	6	21	6	82	3
3172.0	20	6	21	6	11	2

Fig. 5. New column in the dataset – cluster

TABLE III. CLASSIFICATION AND CLUSTERIZATION COMPARISON FOR PERSONIFICATION INTERFACE

Classification	Clusterization
Used for stereotypical personification.	Used for individual personification.
Needs prepared data.	Does not need prepared data.
The number of necessary changes for interface personification is reduced due to user grouping, which is defined as closest to the selected one.	Allows cluster centres to replace all users in the cluster, due to similar users clustering principles. Thus, the user database is formed.
Reduces the number of resources used and improves system performance.	Allows a reduction in search time for solutions and memory.
Possible decline in interface personification quality.	Loss of accuracy at cluster boundaries.

Conclusions

The authors studied classification and clusterization on a real data set and their effect on the personification a user's interface. It can be concluded that classification is better used in situations with prepared data, when users are already using a system and administrators can divide people into groups. On the other hand, clusterization is good at "cold start" situations.

For personification of user's interface classification may assist with stereotypical personification but clusterization – with individual personification.

Support for user personification (differentiation rules provision, interface adaptation, required information obtention) can distinguish an information system from a variety of similar competitive systems.

The authors additionally note that an experiment will be conducted to identify that the particular machine learning methods can be used for user interface personification.

References

[1] A. Schade "Customization vs. Personalization in the User Experience" // Nielsen Norman Group World Leaders in Research-Based User Experience.Jul. 10, 2016. [Online]. Available: https://www.nngroup.com/articles/customization-personalization/ [Accesses: Jan. 11, 2019].

[2] Mou X "What Role do Circadian Rhythms Play in Learning and Memory?" J Neurol Neurophysiol, 2016. 7:367. doi:10.4172/2155-9562.1000367. [Accesses: Jan. 15, 2019]

[3] P. Valdez, C. Ramírez and A. García "Circadian Rhythms in Cognitive Processes: Implications for School Learning". Mind, Brain, and Education, vol. 8, pp. 161-168, 2014 doi:10.1111/mbe.12056. [Accesses: Jan. 11, 2019]

[4] L. Findlater, Krzysztof Z. Gajos. "Design Space and Evaluation Challenges of Adaptive Graphical User Interfaces". AI Magazine. Vol. 30, No. 4. p. 68 – 73. 2009. [Accesses: Mar. 3, 2019].

[5] H. Drachsler, P.A. Kirschner "Learner Characteristics". // Encyclopedia of the Sciences of Learning. Chapter: Learning Characteristics. SpringerEditors: N. M. Seel. 2011. doi: 10.1007/978-1-4419-1428-6_347 [Accesses: Apr. 12, 2019].

[6] C. Fröschl, L. Nguyen, P. Do Learner Model in Adaptive Learning. The 2008 World Congress on Science, Engineering and Technology (WCSET2008), Volume 35, November 2008, Paris, France. 2008 [Accesses: Apr. 13, 2019].

[7] J.R. van Seters, M.A. Ossevoort, J. Tramper, M.J. Goedhart, The influence of student characteristics on the use of adaptive e-learning material, Computers & Education, Volume 58, Issue 3, 2012, Pages 942-952, ISSN 0360-1315, https://doi.org/10.1016/j.compedu.2011.11.002. [Accesses: Apr. 13, 2019]

[8] R.G. Leppan, J.F. van Niekerk, R.A. Botha "Process model for differentiated instruction using learning analytics". South African Computer Journal, 30 (2), pp. 17-43, 2018. doi: 10.18489/sacj.v30i2.481 [Accesses: Apr. 14, 2019]

[9] M.W.Sholom "Text mining. Predictive methods of analyzing unstructured information". M.W.Sholom, N.Indurkhya, T.Zhang, F.J.Damarau. — 2004. — 236 p. [Accesses: Jan. 22, 2019]

[10] MachineLearning.ru – [Online]. Available: http://www.machinelearning.ru/ [Accesses: Jan. 21, 2019].

[11] A.Chasovskih [Overview of data clustering algorithms] Obzor algoritmov klasterizacii dannyh"– Habr. Aug 11, 2010. [Online]. Available: https://habr.com/ru/post/101338/ [Accesses: Mar. 1, 2019].

[12] I.S. Morozov [Clustering methods of users for adaptationinterfaces] "Metody klasterizacii pol"zovatelej dlja adaptacii interfejsov", Bachelor thesis, Novosibirsk State University, Novosibirsk, Russian Federation, 2013. [Accesses: Feb. 27, 2019].

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 222-227

Using Activity Theory for Modelling Transformative Digital Learning

Olena Mykhailenko
University of Ontario Institute of Technology
Canada
olena.mykhailenko@collaboritsi.com

Irēna Žogla
Rezekne Academy of Technologies
Latvia
irena.zogla@lu.lv

Velta Lubkina
Rezekne Academy of Technologies
Latvia
velta.lubkina@rta.lv

Todd Blayone
University of Ontario Institute of Technology
Canada
todd.blayone@collaboritsi.com

***Abstract*— In support of ongoing educational transformation in post-Soviet nations, this article positions activity theory (in the tradition of Engeström) as a framework for modelling changes towards innovative forms of collaborative, fully online digital learning. A strength of activity theory is that it adopts a holistic socio-technical perspective in which teachers, learners, technologies, pedagogical values, roles/identities and rules/cultures are considered together as interdependent elements of collective activity. An illustrative example is offered to model a current and envisioned (target) activity system. In addition, a few considerations to guide research are offered. These include an emphasis on measuring the general readiness of students and teachers, and the need to explore gender divides. The goal is to help envision program transformations towards online learning at two partner universities as part of Ukrainian and Latvian, government-funded projects.**

Keywords—activity theory, educational reforms, online learning, transformative digital learning.

I. Introduction

Fundamental changes in society and technology have disrupted traditional socio-economic activities and triggered educational reinvention. Millions of jobs may be lost or reconfigured in the near future owing to advances in machine automation and human-machine symbiosis as envisioned, for example, by Industry 4.0 [1]. At the same time, many new (and currently unknown) professional roles will emerge requiring a more diverse workforce to develop advanced technological knowledge and skills, a positive disposition towards environmental and organizational change, strong competences for collaborating with both humans and intelligent machine agents, and creative problem-solving abilities [2]. The old concept of mass education, designed for the assembly line, focused on transferring standardized content, and enforcing patterns of top-down control cannot satisfy the demands of a digitalized, globalized, democratized, sustainability-focused and increasingly precarious world [3]. But how must education change to align itself with the needs, rights and expectations of citizens?

Transformative digital learning (TDL) shifts the educational focus from predefined outcomes to emancipatory, digitally-mediated and augmented, inquiry and knowledge construction— "the expansion of consciousness through the transformation of worldviews and specific capacities of the self" [4]. Within this perspective, learners move through ongoing processes of cognitive and social change, establishing intermediate contextual learning goals and rethinking them once they are achieved. Are educators, who have trained for an industrial model of top-down instruction and predefined answers ready for this approach? How might teachers transform their roles, competencies and values to facilitate TDL? What transformation model can be used to guide the transition to TDL? These questions guided our previous research on student and teacher readiness for TDL within the post-Soviet space [5], [6], [7], and many questions remain. In this article, we address the needs of two international partnership projects (Latvia-Canada and Ukraine-Latvia) dedicated to educational transformation in higher education.

II. The context: Starting positions for educational transformations in Latvia and Ukraine

As a global "megatrend," digitalization represents a broad program towards the intensive application of digital technologies to transform manufacturing, business models, urban infrastructures and institutions so as to pursue social innovation and address problems of sustainability [8-10]. However, as important as technological innovation may be, transforming human activity is much more than a technological issue. For example, without adequate digital competences, a deep appreciation of the affordances that digital technologies offer, and most importantly, a new vision of teaching and learning aligned with today's world, it is too easy to "pour

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4044
© 2019 Olena Mykhailenko, Todd Blayone, Irēna Žogla, Velta Lubkina.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

old wine into new wine skins."

In post-Soviet nations, the Soviet system of higher education continues to exert cultural force over educational practices. This system was built on principles of practical training, alignment with collective goals, obedience to hierarchy and centrally administered planning and control [11]. The dissolution of Soviet system triggered significant dramatic changes to both Ukrainian and Latvian social institutions. In general terms, education was realigned to with western models, which emphasized legitimacy, transparency, local autonomy and pluralism. Despite economic limitations, each country made progress on several key elements of educational reform— developing new curricula and teaching materials, and establishing new links with international partners [12-14].

In Ukraine, after the student-initiated, "Euromaidan" revolution in 2014, declarations of educational reform emphasized more learning-centric, digitalized, democratized, and globalized education. However, Ukrainian educators are still struggling to achieve such reforms, often debilitated by lack of (personal, financial and technical) resources and a strategic roadmap well aligned with their context and culture [15]. At the micro-level, individual Ukrainian academics often innovate on their own, pursuing opportunities transformation through international partnerships, application of new learning models, greater academic autonomy and richer integration of digital technologies [7], [16]. It is perhaps noteworthy, however, that our measurement of digital competencies of students and professors at two Ukrainian universities [6],[5] indicate that many remain ill-prepared for fully online learning. In some cases, professors' skills even fell behind those of students: e.g., publishing media and ideas online, operating with audio/video files, program or automate procedures. These are creative, independent-thinking skills, which are vital for 21st-century professionals.

In Latvia, the implementation of more democratic educational principles in universities was often more successful than in Ukraine [17], [18]. University autonomy was established, along with a new research infrastructures, a framework for quality assurance and a differentiated higher education system. Importantly, the Soviet restrictions in content and pedagogy were eliminated, which were especially significant in social sciences and humanities. However, some key issues are waiting to be addressed. For example, institutional transformation should include more democratic learning models that fully leverage the affordances of information and communication technologies (ICTs) and global networks. As McGuinness [19] argues, these changes require the professional development of faculty and researchers. To this end, several strategies, including strengthening doctoral programs, should be implemented. Given a shrinking student age cohort as well as migration and fluctuating economic conditions, it is likely that serious transformations of the institutional landscape in Latvia are just beginning [18].

As some researchers of post-Soviet educational transitions point out, a common characteristic of all countries of this group is a quantitative and qualitative shortage of research in HE institutions, with the focus of institutions remaining exclusively on teaching [18]. We see this as a potential hindering factor for post-industrial learning transformations.

In addition to the broad socio-contextual factors, readiness for innovative digitally-mediated learning also contains an important gender aspect. Although the number of females obtaining post-secondary education in both countries are higher than that of males, only one-fifth of women work in high-tech industries in Ukraine, and about one quarter in Latvia [20]. This situation is typical around the world. Women largely tend to avoid ICT-related studies and are less likely to choose digital careers. In the EU, only 17,2% of IT-students and 16,7% of the employed IT-specialists in 2016 were females. The highest level of female participation in IT is in Bulgaria—31% of women work in technology-related positions. At the same time, by 2020, the EU will lack 900,000 skilled ICT professionals [21], which makes the broad gender imbalance in IT a critical issue.

Importantly, our previous studies of general digital competencies in Ukraine and Georgia [6], [5] did not reveal notable gender differences. Therefore, the gender divide in the IT community most probably relates to other socio-cultural factors [22]. Some researchers find that distinct gender attitudes toward technology is reproduced and institutionalized within educational systems themselves [23]. That is, females are largely enculturated to avoid working in a male-dominated IT sector. This disbalance in the labor market, which is bound to have negative consequences over the long term, might be partially addressed by incorporating more digitally focused learning models that emphasize human-computer interaction aligned with visions of a "smart-technology" society.

To summarize, the educational transformations in both Ukraine and Latvia should be *democratically and digitally* reformed with a focus on strengthening innovative research skills, self-directed (life-long) learning and emancipatory values such as those related to perspectival diversity, gender equality, interpersonal trust and freedom of expression.

III. From traditional to transformative learning: What changes are needed?

Mezirow, the founder of transformative learning theory, emphasized that adult learning must include reinterpreting the world rather than simply acting on prior beliefs, judgements, and feelings. He suggests, that learning goals should be considered in relation to both short- and long-term perspectives in a manner similar to the popular saying: giving a person a fish feeds her for a day (short-term objectives), teaching her how to fish can feed her for a lifetime (long-term goals). An short-term objective may be to obtain job-related competencies, but the life-term goal could extend is to become a socially active critical thinker and an engaged global citizen [24].

Traditional (content-focused and highly structured and teaching models) [25] tend to emphasize the transfer of discipline-specific knowledge. However, for successful functioning in a transitioning and digitalized economy marked by disruptive innovations, this content-focused system of higher education has limited value. University-level education would do well to focus on longer-term goals, and facilitating the development of digital skills, creative thinking, collaborative problem solving cognitive flexibility and emotional intelligence—the things that are difficult to automate with increasingly intelligent machines [26] Consistent with this perspective, Harkins [27] contrasts traditional learning to future-oriented digital education across five dimensions as shown in Table I.

Table I Differences in Traditional and Future-Oriented Models of Education

Distinguishing characteristics	Traditional ("download" and "open-access") models of digital education	Future-oriented ("knowledge and innovations producing") models of education
Meaning	Dictated or socially constructed with limited internet aid	Socially constructed and contextually reinvented knowledge, built through selective individual and team-driven embodiments in practice
Knowledge creating process	Transfer teacher-to-student, using learning management systems (LMS) or MOOCs	Technology-enabled co-construction of knowledge in learning community, amplified by critical creative feedback loops 24/7
Schooling location	In buildings or online, through hybrid or full internet courses with pre-recorded content	Everywhere, emerging in the globally networked human body
Who teachers are	Licensed professionals	Everybody, who is an innovation producing source, backed up by intuitive software "partners" and human collabourators
How employers view graduates	"Line workers", who must be trained to follow instructions	Knowledge- and innovation-producing co-workers and entrepreneurs who can initiate the new ways and support sustainable development

Some leading collaborative learning models, adopting a problem-based learning (PBL) approach are well aligned with this vision of future-oriented learning. For example, the Fully Online Learning Community (FOLC) [28] (Fig.1), developed by EILAB researchers at the University of Ontario Institute of Technology (UOIT), Canada, which has several notable characteristics. It was initially conceived as an offshoot of the Community of Inquiry (CoI) model [29], one of the most thoroughly researched and widely used collaborative-constructivist theoretical models in the world [30],[31]. However, it removed "teaching presence" as a distinct dimension of learning so as to emphasize the need for (re)distributing educational control, reducing power distance between students and educators, and democratizing communication. Moreover, it inserted an emphasis on digital affordances and related digital competences—elements which the CoI considers extraneous. FOLC continues to evolve conceptually as it is applied daily at the Faculty of Education, UOIT as the basis for teacher education programs and deployed as a research framework for empirically studying the dynamics of fully-online community-based learning.

Fully Online Learning Community Model

Fig. 1. Fully Online Learning Community (FOLC) model

FOLC-based research also recognizes the importance of socio-emotional interaction [32], and the innovative use of open digital affordances [33] (e.g., using mainstream social media for supporting collaborative inquiry and community-building). Most importantly, FOLC is designed as a flexible model, adaptable to the needs of learners in a variety of socio-cultural contexts. It is for this reason that we have used it as a guiding model in our own international teaching and learning projects.

IV. Modeling Learning Transformations with Activity Theory

Engeström, building directly on Vygotsky, Leontiev and other well-known cultural-historical psychologists from the early Soviet period [34],[35],[36] offers a visual theoretical model that includes an emphasis on individuals, tools, and several social and cultural mediators of collective technologically-mediated activity (Fig.2). This multi-triangle model, drawn with arrows to emphasize the dynamic nature of activity and the interrelatedness of the elements, offers a framework for envisioning educational transformation. More specifically, it provides an apparatus for modeling both current and desired educational practices (understood as activity systems), and exploring tensions between elements and other activity systems. Within activity theory, tensions are catalysts for pursuing critical analyses and concrete changes. The assumption is that professionals (e.g., teachers) are, or can become, active and empowered agents capable of making evidence-based decisions and changing their practices without an official mandate. In this sense, activity theory emphasizes and enables radical change "from below" through both individual and collective action.

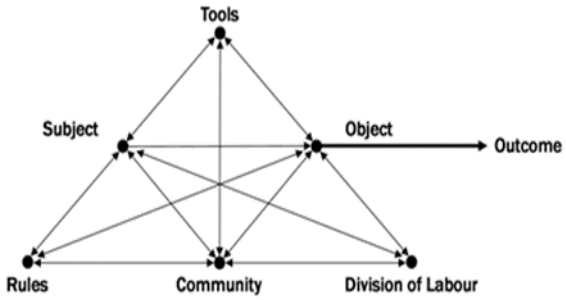


Fig. 2. Activity System model

The activity system model includes six constituent elements: the human agent, tools/artifacts, an object, a "community", rules, and division of labour. Each of these elements is modeled as a discrete entity which functions in a mediational relationship to the other elements. To more fully conceptualize these elements, they are presented in more detail.

- **Subject**: This is an active agent capable of taking purposeful action towards a goal by using the affordances of tools and technologies effectively, aligning her actions strategically with a driving collective objective (which is always flexible), and considering contextual opportunities and constraints.
- **Object**. The object or the "objective" represents the driving force of the activity. An object may be a need, desire, a vision of the future, a mission statement, a manifesto or a defined outcome. An activity without an object is meaningless!
- **Tools**. Tools include any instruments, artefacts or technologies used by participants to carry out an operation or action aligned with the activity. With today's digital society, a variety of digital hardware, software and systems are used to interact with each other and the world.
- **Community**. The community represents the relevant mediating collective, whether this is a team, workgroup, department, committee, class or something else. Collective activity always includes an individual and collective perspective, and there is no escape from the social nature of activity. Even actions taken alone are deeply connected to society and culture.
- **Rules**. These include both explicit and implicit ways for coordinating and mediating relations between the community members themselves and between them and the objects. Organization policies and values both function to mediate activities and participant perceptions.
- **Division of Labour**. From earliest recorded history, humans have coordinated their efforts to achieve desired outcomes more efficiently or effectively. Today, division of labour spans both human and machine agents in complex digitalized systems.
- **Outcome**. The outcome represents the end result of an activity. This outcome may become input for a new activity system. For example, a design or set of recommendations created from one group may be passed to another to implement.

Activity systems, by their nature, produce tensions, contradictions and innovations, which generate both resistance among participants and possibilities for transformations. With the assistance of human mediators, transformations can be guided through a cycle of "expansive learning" in which existing objectives and standard practices modelled, analyzed and reinvented. This cycle typically starts by identifying contradictions, and using the activity-system apparatus to map both the current and desired state of affairs. The next step is to gather motivation, resources and strategies for implementing the new model in practice. Achieving change requires communication, activism and ongoing learning. Indeed the most successful results will be achieved when change agents commit to democratic deliberation, critical reflection and divergent thinking [36].

To provide an example of how the activity system apparatus can be applied in practice, we have modelled both a (hypothetical) existing and target educational activity system based on our experience as educational-transformation consultants (Fig.3).

V. A Research Proposal

The activity-system apparatus and transformation process described, is ideally suited for exploring potential educational transformations in post-Soviet contexts. Moreover, this model can be supplemented with several additional conceptual tools to help develop a general readiness for change among all participants. In our previous research, we have identified digital competences and personal-cultural orientations as especially important readiness factors at the micro-level. To measure these dimensions of readiness at the individual and group-levels we rely on several valid instruments. These include:

- Digital Competency Profiler (DCP) [37]
- Personal Cultural Orientation (PCO) [38]
- Cultural Values Scale (CVScale) [38]
- Attitudes toward IT (AttIT) [39].

Data collected from these instruments can be analyzed using a number of statistical strategies that are designed based on the specific context of research and the guiding research questions. In any given study the focus may be on comparing different demographic segments, positioning participants in relation to established thresholds (that have been validated in relation to observed performance) or exploring individual changes over time. All of these situations require different analytical strategies.

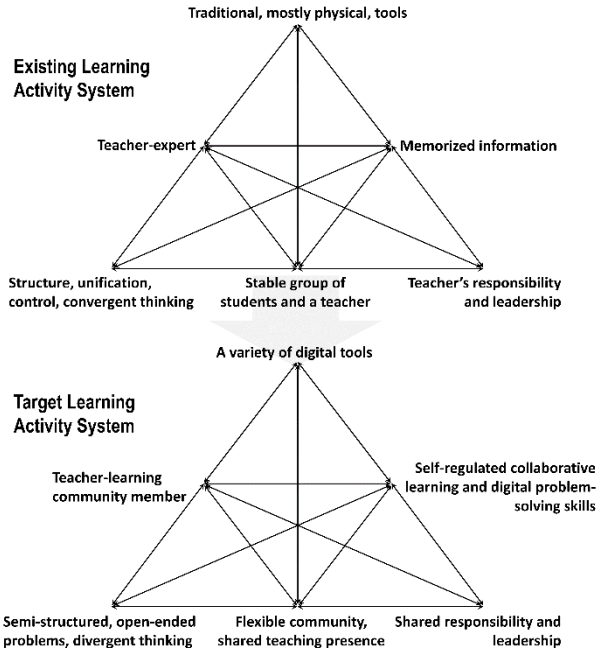


Fig. 3. The developmental model of the transition toward transformative learning

V. Discussion

The outcomes of a transformation process and accompanying readiness research can be used as a firm foundation for creating new online courses, improving the existing academic program in pedagogy and creating transformative teachers' professional-development programs. Because transforming education in the direction of collaborative digitally-mediated learning is an ongoing process all participants must remain open to new ideas.

When studying values, there is always the question as to what extent they can be changed as a result of transformative learning, and what time interval is needed for changes to become noticeable. Hofstede emphasizes that fundamental values are deeply entrenched within social practices and institutions, and they do not change quickly [40]. However, there are reasons to believe that practices can change. Some scholars have observed that when online communities come together to learn or organize social change, new sub-cultural dispositions emerge that, have a direct effect on forms of communication, democratic functioning, uses of digital technologies, and collective social commitments. Even if emergent community values cause (initial) internal tensions in some participants and trigger forms of subconscious resistance (or culture shock), caring and tolerant communities can move forward together by being open and honest about their feelings, respecting diversity as a core value and practicing understanding. In short, there is expressed commitment, there will be opportunities to hack our "software of the mind" [41].

VI. Conclusions

This theoretical proposal presented activity theory as a framework for modeling, analyzing and re-designing teaching and learning in higher education. Using the elements of Engeström's activity system, we suggested specific transformations toward digitalized and democratized learning that appear well-aligned with the needs of Latvia and Ukraine. Based on our research and experience we also noted that digital readiness and limiting gender perceptions are challenges to overcome. Our hope is that this proposal will provide learners and educators in higher-educational institutions in these countries with a practical apparatus for reforming university programs toward the requirements of a 21st century economy and society.

Acknowledgments

This article is supported by two funded projects: "Implementation of Transformative Digital Learning in Doctoral Program of Pedagogical Science in Latvia (DocTDLL) lzp-2018/2-0180" and "Gender aspects of digital readiness and development of human capital in regions" (Ukraine-Latvia) (Nr.LV-UA/2018/3).

References

[1] M. A. L. Luo, J; Boccardo, J, "New vision for education: Fostering social and emotional learning through technology," in "Industry Agenda," World Economic Forum, 2016, [Online]. Available: http://www3.weforum.org/docs/WEF_New_Vision_for_Education.pdf.

[2] T. Blayone, "Theorizing effective use of digital technology with activity theory," *Accepted for publication in Technology, Pedagogy and Education,* 2019.

[3] A. Till, V. Ratcheva, and S. Zakhidi, "The Future of Jobs Report " World Economic Forum, Centre for the New Economy and Society, 09/17/2018 2018, [Online]. Available: https://www.weforum.org/reports/the-future-of-jobs-report-2018.

[4] M. J. Elias *et al.*, Promoting social and emotional learning: Guidelines for educators. Alexandria, Virgina, USA: Association for Supervision and Curriculum Development, 1997.

[5] T. Blayone, O. Mykhailenko, R. VanOostveen, O. Grebeshkov, O. Hrebeshkova, and O. Vostryakov, "Surveying digital competencies of university students and professors in Ukraine for fully online collaborative learning," *Technology, Pedagogy and Education,* vol. 27, no. 3, pp. 1-18, 2017.

[6] T. Blayone, O. Mykhailenko, M. Kokhan, M. Kavtaradze, R. vanOostveen, and W. Barber, "Profiling the digital readiness of higher education students for transformative online learning in the post-soviet nations of Georgia and Ukraine," *International Journal of Educational Technology in Higher Education,* vol. 15, no. 37, pp. 1-22, 2018.

[7] O. Mykhailenko, T. Blayone, and R. vanOostveen, "Exploring democratized learning and dimensions of culture for educational transformation in Ukraine," in *RIDRU Conference: Higher Education Reforms in Post-Maidan Ukraine, October 25, 2016*, ed. Alberta, Canada, 2016.

[8] R. Morrar, H. Arman, and S. Mousa, "The fourth industrial revolution (Industry 4.0): A social innovation perspective," *Technology Innovation Management Review,* vol. 7, no. 11, pp. 12-20, 2017.

[9] B. Wang, "The Future of Manufacturing: A New Perspective," *Engineering,* vol. 4, no. 5, pp. 722-728, 2018.

[10] L. D. Xu, E. L. Xu, and L. Li, "Industry 4.0: State of the art and future trends," *International Journal of Production Research,* vol. 56, no. 8, pp. 2941-2962, 2018.

[11] A. Kuraev, "Soviet higher education: An alternative construct to the western university paradigm," *Higher Education,* vol. 71, pp. 181-193, 2016.

[12] B. Kutsyuruba, "Education in the period of post-Soviet transition in Ukraine," *Demokratizatsiya,* vol. 19, no. 3, pp. 287-309, 2011.

[13] D. V. Powell, S. Kuzmina, Y. Kuzmin, T. Yamchynska, and O. Shestopalyuk, "Using web-blended learning in Ukraine to facilitate engagement and globalize horizons: A pilot study," *The Online Journal of Distance Education and e-Learning,* vol. 2, no. 2, pp. 34-41, 2014.

[14] D. V. Powell, S. Kuzmina, T. Yamchynska, O. V. Shestopalyuk, and Y. Kuzmin, "Educational technologies for maturing democratic approaches to educational practices in Ukraine," *Procedia-Social and Behavioral Sciences,* vol. 176, pp. 378-385, 2015.

[15] S. Kvit, "The paradoxes of post-Soviet education," in *KyivPost*, ed. Kyiv, 2016.

[16] M. Hladchenko, D. Antonowicz, and H. de Boer, "Understanding the changes of the higher education governance in Poland and Ukraine: Institutional analysis," in *The University as a Critical Institution?*, D. R. and H. Eggins, Eds. Rotterdam, The Netherlands: Sense Publishers, 2017, pp. 55-74.

[17] T. Khavenson and M. Carnoy, "The unintended and intended academic consequences of educational reforms: The cases of Post-Soviet Estonia, Latvia and Russia," *Oxford Review of Education,* vol. 42, no. 2, pp. 178-199, 2016.

[18] A. Smolentseva, J. Huisman, and I. Froumin, *Transformation of Higher Education Institutional Landscape in Post-Soviet Countries: From Soviet Model to Where?* (25 Years of Transformations of Higher Education Systems in Post-Soviet Countries: Reform and Continuity). Palgrave Macmillan: Cham, Switzerland, 2018, pp. 1-43.

[19] A. C. J. McGuinness, "Reforms in the Baltics," *International Higher Education,* no. 25, pp. 13-14, 2001.

[20] A. Till, V. Ratcheva, and S. Zakhidi, "The Global Gender Gap Report " in "Insight Report," World Economic Forum, 2017, [Online]. Available: http://www3.weforum.org/docs/WEF_GGGR_2017.pdf.

[21] European Parliament. (2018, 08-03-2018). *More women in ICT: empowering women in the digital world* [Online]. Available: http://www.europarl.europa.eu/news/en/headlines/society/20180301STO98927/more-women-in-ict-empowering-women-in-the-digital-world.

[22] J. C. Blickenstaff, "Women and science careers: leaky pipeline or gender filter?," *Gender and education,* vol. 17, no. 4, pp. 369-386, 2005.

[23] P. C. Gorski, "What we're teaching teachers: An analysis of multicultural teacher education coursework syllabi," *Teaching and Teacher Education,* vol. 25, no. 2, pp. 309-318, 2009.

[24] J. Mezirow, "Transformative learning: Theory to practice," in *New directions for adult and continuing education*, vol. 1997no. 74). San Francisco, USA: Jossey-Bass Publishers, 1997, pp. 5-12.

[25] B. Levin, "Putting students at the centre in education reform," *Journal of Educational Change,* vol. 1, no. 2, pp. 155-172, 2000.

[26] C. B. Frey and M. A. Osborne, "The future of employment: How susceptible are jobs to computerisation?," *Technological Forecasting and Social Change,* vol. 114, pp. 254-280, 2017.

[27] A. M. Harkins, "Leapfrog principles and practices: Core components of education 3.0 and 4.0," *Futures Research Quarterly,* vol. 24, no. 1, pp. 19-31, 2008.

[28] T. Blayone, R. vanOostveen, W. Barber, M. DiGiuseppe, and E. Childs, "Democratizing digital learning: Theorizing the Fully Online Learning Community model," *International Journal of Educational Technology in Higher Education,* vol. 14, no. 13, pp. 1-16, 2017.

[29] D. R. Garrison, "Online Community of Inquiry review: Social, cognitive, and teaching Presence issues," *Journal of Asynchronous Learning Networks,* vol. 11, no. 1, pp. 61-72, 2007.

[30] P. P. E. Elicor, "The notion of pedagogical authority in the community of inquiry," *Kritike,* vol. 11, no. 2, pp. 80-92, 2017.

[31] W. Boston, S. R. Díaz, A. M. Gibson, P. Ice, J. Richardson, and K. Swan, "An exploration of the relationship between indicators of the community of inquiry framework and retention in online programs," *Journal of Asynchronous Learning Networks,* vol. 13, no. 3, pp. 67-83, 2014.

[32] R. vanOostveen, E. Childs, J. Clarkson, and K. Flynn, "Becoming close with others online: Distributed community building in online PBL courses," presented at the EdMedia 2015 World Conference on Educational Media and Technology, Montreal, Canada, June 22-25 2015, 2015.

[33] R. vanOostveen, M. DiGiuseppe, W. Barber, T. Blayone, and E. Childs, "New conceptions for digital technology sandboxes: Developing a Fully Online Learning Communities (FOLC) model," in *EdMedia 2016: World Conference on Educational Media and Technology*, G. Veletsianos, Ed. Vancouver, B.C., Canada: Association for the Advancement of Computing in Education (AACE), 2016, pp. 665-673.

[34] L. Vygotsky, *Mind in society: Development of higher psychological processes*. Cambridge, MA: Harvard University Press, 1978.

[35] A. N. Leontiev, *Activity, consciousness and personality*. Englewood Cliffs, NJ: Prentice-Hall, 1978. 1975.

[36] Y. Engeström, "Activity theory as a framework for analyzing and redesigning work," *Ergonomics,* vol. 43, no. 7, pp. 960-974, 2000.

[37] EILAB. (2017, October 29). *Digital Competency Profiler* [Online]. Available: https://dcp.eilab.ca/.

[38] P. Sharma, "Measuring personal cultural orientations: Scale development and validation," *Journal of the Academy of Marketing Science,* vol. 38, no. 6, pp. 787-806, 2009.

[39] A. A. Gokhale, P. E. Brauchle, and K. F. Machina, "Scale to measure attitudes toward information technology," *International Journal of Information and Communication Technology Education,* vol. 9, no. 3, pp. 13-26, 2013.

[40] G. Hofstede, *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*, Second ed. Thousand Oaks, California, USA: Sage Publications, 2001.

[41] G. Hofstede, "Culture and organizations," *International Studies of Management & Organization,* vol. 10, no. 4, pp. 15-41, 1980.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 228-232

Professional Competence of Future Engineers in the Process of Training of Computer Science Disciplines

Iryna Shumilova
the doctor of pedagogical sciences, the assistant professor of the chair of management and administration
Berdyansk State Pedagogical University
Berdyansk, Ukraine
azkur-shif@ukr.net

Olena Ionova
the doctor of pedagogical sciences, the professor of the chair of Natural Sciences and Mathematics
Kharkov National Pedagogical University named after G.S.Skovoroda
Kharkov, Ukraine
ionova_54@mail.ru

Oksana Syvak
Department of Cultural Studies and Information Activities
Mariupol State University
Mariupol, Ukraine
o.sivak@mdu.in.ua

***Abstract*—The theoretical principles of formation professional competences of future engineers are presented by means of informatics disciplines in this article. The issue of reforming of the higher system of education in Ukraine is considered by introducing a competent approach to the organization of the educational process. The role of teaching computer science disciplines is determined in the formation of professional competencies of future engineers, psychological and pedagogical background and methodical requirements as for the design of competence-oriented content of training of the informational disciplines of future engineers.**

The article reveals the ways of practical implementation of the competence approach during the training of the computer science disciplines by future engineers and describes the developed model of formation of professional competencies of future engineers in the process of studying computer science disciplines. The peculiarities of forming professional competencies are determined during classroom sessions. The means of development of professional competencies and formation of research skills of future engineers are revealed by means of the using of competence-oriented tasks in students' independent work and methodical peculiarities of organization of educational computer-technology practice in the context of implementation of the competence approach. The experimental review of the effectiveness of the developed model is described as for the formation of professional competencies of future engineers in the process of training computer science disciplines.

***Keywords*—future engineers, competency, competence, professional competence, competence-oriented content of training, computer science disciplines.**

I. INTRODUCTION

The main challenges that have to be taken into consideration when preparing modern engineers are the rapid development of science and technology, the need to implement strategy for sustainable development in all areas of human activity and a significant improvement in the practical training of engineers. The issue of the importance of competencies and their formation is one of the ways of studying the conditions of engineering education and it is necessary to create a set of competencies, that is associated with the effectiveness of teaching. The readiness of the university graduates to fulfill their professional duties at a high professional level is an indicator of their mobility in the present [6]. The modern practice of the educational system has its purpose as for strengthening the elements of competence that can include a set of main competences, skills, technologies that allow the university to provide students with educational services. Basic basis of engineering education should become not only academic disciplines, but also methods and forms of organization of educational activities [5].

In this article we consider of the concept of competence and identify the conditions of engineering education as the basis for the formation of a competency structure for effective teaching of computer science in achieving the methodological goals. As a result of the review of the literature, it is necessary to use different classifications of the set of competences that are determined for scientific and practical purposes as well as to determine their use and significance in connection with the concept of engineering education. In the conformity with the results of the study, the basic set which includes such elements as knowledge, skills and abilities, must also determine the readiness for the use of these elements that are interwoven with the concept of engineering education. The object of the research is the competence of the future engineer and its elements.

II. MATERIALS AND METHODS

The subject of the study is the interaction of elements of competence such as knowledge, skills and abilities. Methodology: review of literature, compilation and interpretation. The meanings and competency elements

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4133
© 2019 Iryna Shumilova, Olena Ionova, Oksana Syvak.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

in various fields related to engineering, management, social sciences and human resources are considered when studying sources.

The American Council for Professional Development Engineers (ECPD) defines the term "engineering" as follows: "Creative application of scientific principles for the design or development of structures, machines, equipment, production processes or work on their use separately or in combination; designing or managing them with full knowledge of their design; predicting their behavior in certain operational decisions"[1]. Thus, people who are prepared for engineering activities must have the necessary knowledge, abilities, creative potential and professional skills [3].

Professional competence is a dynamic structure that is shown via the management of the individual and it is a person's ability to meet social needs, to solve successfully professional and problem tasks.

Among the elements of professional competence of future engineers are: scientific and technical (confirmed knowledge, skills and abilities, experience in the field of engineering, knowledge of engineering activities); information and communication technology (confirmed ability of the person to use computer technologies for guaranteed information and getting the information to meet their own individual needs and social requirements; energy management in engineering activities); information and management (confirmed ability to perform information management functions, organization and management of engineering activities). The pronounced substantive and specialized nature of these elements confirms that the main factor in ensuring the up-and-coming and future professional growth of future engineers is their training in computer science disciplines that determines their role as a source of systemic getting professional competences by students.

Getting an engineering education at least means to realize the purpose of various types of engineering activities in the relevant subject area, see the whole range of means of actions, to know how and be able to use the main ones, possess their basic technologies and understand what is at the entrance and what is desirably to get at the exit [4]. In order to form the professional competence of the future engineer it is necessary to design a competence-oriented content of training for the teacher of the university, which is understood with the problem of applied character, the means of solving are modern computer and communication technologies and the content of tasks corresponds to certain official duties.

The basic psychological and pedagogical conditions for designing the competence-oriented content of teaching of computer science disciplines is the construction of the process of teaching the informatics disciplines on the basis of system, complex, activity-oriented, person-oriented, competence-oriented approaches; taking into consideration psychological and pedagogical peculiarities of the development of future engineers in the conditions of informational educational background providing with the content of training on the interest of future engineers in the profession; formation of knowledge as for the content of engineering activities and the skills to use the acquired knowledge to solve professional tasks; providing benefits to problem-based, research-based teaching methods.

The main methodological requirements for the organization of competence-oriented content of training of the computer disciplines of future engineers is aimed at: the formation of scientific and technical, information management, information and communication and technological competencies of all components of the educational process by establishing strong interdisciplinary links; training to transfer the acquired theoretical knowledge into the place of their direct use with maximum approaching to the real professional-engineering field and with the use of information and communication technologies; adherence to the principles of the development and use of competency-based tasks with the definition of their purpose and implementation in the educational process.

III. RESULTS AND DISCUSSION

In order to form the professional competencies of future engineers in the process of teaching computer science, we have developed a "Model for the formation of professional competencies of future engineers in the process of teaching computer science and its experimental verification" as an integral system of training.

The developed model contains three blocks: theoretical, operational-activity and criteria-estimating that in general covers all aspects of formation of professional competences of future engineers in the process of training of computer science disciplines.

Organizational-pedagogical conditions are direction of preparation of engineers for the complex formation of professional competencies; designing of competence-oriented content of computer science disciplines; step-by-step management of the process of formation of professional competencies; taking into consideration psychological and pedagogical and individual characteristics of students in the conditions of informational and educational environment.

The theoretical block of the model grounds the necessity of submission the training of the computer science disciplines to the formation of professional competencies of future engineers, training to transfer the acquired computer knowledge and skills into the field of their direct use that promote the competitiveness of the graduate-document scientist in the domestic and international labor markets.

In the theoretical block the following components are interwoven: preparation of specialists in engineers studies, formation of professional competencies of engineers in conformity with social order; the task: to master students with knowledge and skills in computer science disciplines; formation of skills in the use of information and communication technologies in solving competent-oriented tasks; development of students' motivation

for the active use of information and communication technologies in their professional activities; directing the student's personality to self-esteem, self-examination and professional self-improvement; psychological and pedagogical conditions, methodical requirements and principles: science approach, systemic, ensuring the developing nature of education, fundamentalism, variability and alternative, intersubject communications, educational interaction, consciousness and activity, professional orientation; approaches: systemic, complex, individual; components of professional competencies that need to be formed by students- engineers.

Operational-activity block of the model of formation of professional competences of future engineers in the process of training of the informational disciplines contains the stages of management in the process of formation of professional competencies (preparatory, basic, vocational); describes the classroom and extra-curricular learning resource, the content of training and methodical (technological) component (forms, methods and means of training that make students act, getting professional competencies in the process of teaching computer science disciplines). The educational content resource, criteria and indicators of the formation of professional competencies of future engineers are determined for each stage.

The main forms of organization of the educational process are lessons, independent educational and research work, practical training and control. We considered the type of interaction between the teacher and the student in determining the forms of students learning; the nature of the object and subject of study; the place and conditions of holding classes. Frontal, individual, group and individual-group forms of educational activity were allocated. Educational materials include printed educational materials (textbooks, manuals, methodical recommendations, charts), periodicals (newspapers, magazines, collections of scientific works), electronic educational materials (multimedia tutorials, presentations, disks), software, internet resources, multimedia projector, documents of various types, individual tasks, algorithms-tips, tests. Not only lecture, explanation and annotation were selected among the methods of organization of students educational activities for the formation the system of knowledge, but also the search for information with the help of the Internet, the analysis of periodicals, the conclusion of glossaries, preparation of abstracts, etc.

For the development of analytical skills it is necessary to write essays using different forms of argumentation and reflexive assessment, articles review, critical review of the sources, informative and argumentative presentation. If we want to develop skills of solving professional problems, we have to think about the problem of definition, collection, analysis and interpretation of data, situation, justification of the algorithm of action, assessment of alternatives; for demonstration of professional information and communication actions - thematic search, annotation, development of multimedia presentation support, database creation, site design, participation in the conference, thematic discussion and debate. If we develop skills of rational organization of educational activity, self-development, it is necessary to conduct diaries-organizers of time; create a personal bank of forms (technological cards), self-evaluation; mutual evaluation; mutual learning; group project.

The criteria-assessment block contains the following criteria: motivational-value that is based on the knowledge of the essence and structure of future professional activities and the possibilities of using computer science knowledge and skills, an interest and desire to execute competency-oriented tasks independently with the use of informational and communicational technologies; cognitive - the formation of a system of informative knowledge and skills in the performance of professional functions; activity - the ability to use information and communication technologies in their own professional activities; evaluative – the ability to evaluate their professional abilities, readiness for their development and indicators of the formation of professional competence of future engineers in the process of teaching computer science disciplines. The monitoring support for the implementation of the model of formation of professional competences of future engineers with the definition of the expected result, criteria of evaluation, diagnostic tools and types of control has been developed. An additional stage of adaptation in the primary position has been introduced.

The developed model of formation of professional competencies of future engineers in the process of teaching computer science offers a reorientation from the training of students of computer science disciplines to study the profession of engineer, theoretical knowledge is not formed verbally but through the active actions.

The model of the formation of professional competencies of future engineers in the process of teaching computer science disciplines according to the level of the system is the macro model, according to the level of formalization of information provision - the conceptual model, the place in the structure of scientific knowledge - epistemological, according to the orientation on the reproduction of certain aspects of the implementation of the competence approach - functional and structural with the way of communication between the presented data - hierarchical network according to the purpose and tasks of the modeling - descriptive prognostic and with the time factor - dynamic.

In order to implement the developed model of the formation of professional competencies of students, engineers analyzed the educational standards and curricula from the specialty 121 "Software Engineering", the sequence of studying of computer science and professional disciplines was formed, the connection of the content of the professional skills of engineers and content modules of computer science disciplines was determined. A bank of competence-oriented tasks for all computer science disciplines was created. These tasks were divided into groups in conformity with the levels of complexity and duration of implementation for using in the classroom, independent study and research work of students, as well as during student computer and technology practice.

Competent-oriented tasks during the computer-technological practice of students were formed in such a way that students had an opportunity to get acquainted with the document circulation of structural units, to form skills of using information and communication technologies in solving professional tasks, to create for themselves a database of forms (technological cards) for further professional activity. For example, the place of practice is the personnel management department. The content of the task is the development of an electronic form for employee registration, which allows to systematize information about the work experience, date of birth, information about the release of employees in the current year.

During the development and testing of the effectiveness of the model for the formation of professional competencies of future engineers with the help of computer science disciplines, a pedagogical experiment was held, consisting of three stages: summative, searching and formative assessment.

The test of the effectiveness of the developed model for the formation of professional competencies of future engineers in the study of the disciplines of the computer science cycle at the stage of the molding experiment was carried out according to the t-criteria of the Student. At the same time, all the samples were homogeneous and independent, and classes in the control and experimental groups of individual universities were held by the same teacher.

The only difference in learning in the experimental and control groups was the methodical system of teaching computer science disciplines. The experiment covered 13 experimental and 14 control groups of students. The total number of students who studied according to the experimental method was 343, and the control function of the experimental activity was carried out during the study of 351 students.

At the beginning of the experiment the hypotheses were put forward: zero H_0 - the level of academic achievements in informatics in experimental groups of students is not higher than in control groups, alternative H_1 – the level of educational achievements in computer science in experimental groups of students is higher than in control groups. By the t-criterion of the Investigator, the hypothesis H_0 was adopted, that is the level of educational achievements in informatics in the groups was the same. After the experimental study the results of the state certification of the same students are analyzed according to the same criteria. Among the used tasks were professional exercises, the solution of which should identify the students' ability to use information and communication technologies. The hypothesis was adopted – the level of academic achievement among students in the experimental group is higher than that of the control group students. Figure 1 shows the results of learning the students at the beginning and at the end of the experiment, reflecting the results.

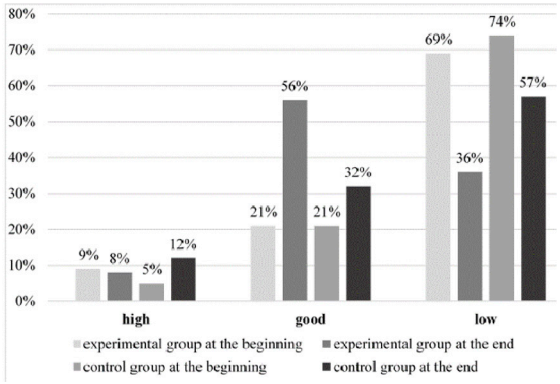


Fig.1 Diagram of student learning outcomes at the beginning and end of the experiment

In the process of forming an experiment, the criteria for the formation of professional competencies were also analyzed (Fig. 2).

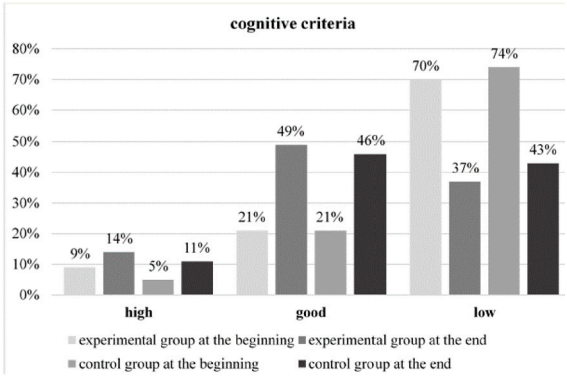


Fig.2 Diagrams of changes in the criteria for the formation of professional competencies (a)

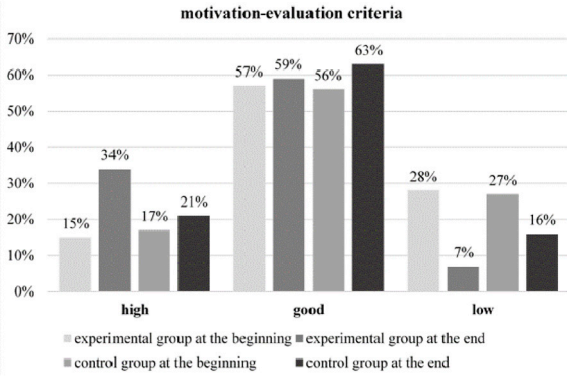


Fig.2 Diagrams of changes in the criteria for the formation of professional competencies (b)

Interviews made with university graduates who were working in their field found that those who studied in experimental groups adjusted more quickly to the workplace and some of the graduates held senior positions in one to two years.

IV. TESYINGAT THE UNIVERSITIES OF UKRAINE

Implementation of the research results was carried out during the experimental work at the Mariupol State University and it has its special value.

The developed model of formation of professional competencies of future engineers in the process of teaching of computer science disciplines has been introduced into the practice of teaching in Donetsk National University,

Mariupol State University, National University «Lviv Politechnik», Private Higher Educational University «European University», in Khmelnitsk institute of social researches [2].

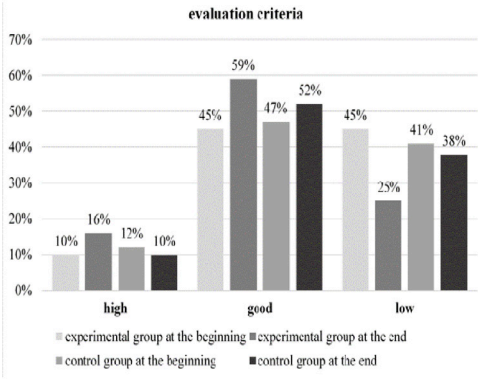


Fig.2 Diagrams of changes in the criteria for the formation of professional competencies (c)

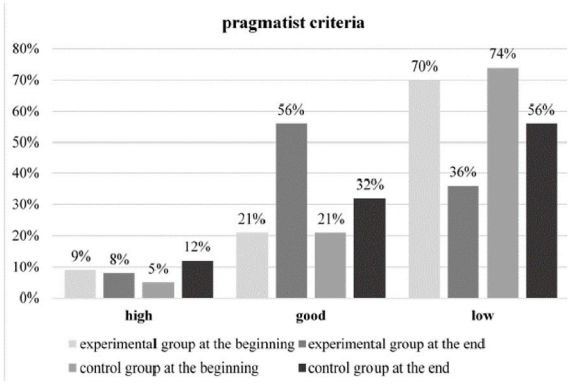


Fig.2 Diagrams of changes in the criteria for the formation of professional competencies (d)

V. CONCLUSION

The authors of the article conclude that the main dominant element of competence is the ability to act as well as knowledge and skills. The introduction of the model made it possible to determine the definite final results of the professional training of future engineers, positive changes in the quality of their preparation for work in the conditions of the information society, which is confirmed by the t-criteria of the Student.

ACKNOWLEDGEMENTS

According to the results of the experiment the effectiveness of the incorporating the integrity of the components of the model of professional competencies of future engineers during studying of educational subjects (theoretical, organizational-active and critical assessment block) that provide interaction of motivational, cognitive, active and evaluation criteria of forming professional competencies. This model presents the productivity, implementation of the competency approach not as discrete pedagogical phenomenon that changes certain aspects in the pedagogical process when studying computer science disciplines, but as a holistic system due to the transition of higher education to a new type of preparing specialists that contributes to lifelong learning.

REFERENCES

[1] Б. Бідій, "Інжинірингові послуги: нові можливості," Налоговый кодекс, 2011. [Online serial]. Available: www.buhgalter.com.ua

[2] О.А. Сивак, Формування професійних компетентностей майбутніх документознавців у процесі навчання інформативних дисциплін. Бердянськ: автореферат, 2014.

[3] А.В. Журавський, А.Я. Яцейко, Н.Б. Дьяченко, Основи технічної творчості та наукових досліджень, Львів: Львівська політехніка, 2012.

[4] М. Карпаш, О. Карпаш, Є. Крижанівський, Вища інженерна освіта в умовах сталого розвитку суспільства, Вища освіта України, №2, с. 55–60, 2015.

[5] В.С. Харченко, Ред., Основы зеленой ИТ-инженерии. Моделирование облачных систем. Практикум. Министерство образования и науки Украины, Нац. аэрокосмический ун-т им. Н.Е. Жуковского "ХАИ", 2016.

[6] І.Ф. Шумілова, Самоосвіта студентів з питань формування загальнокультурної компетентності, Наукові записки Бердянського державного педагогічного університету, Педагогічні науки, Вип. 1, с. 318–328, 2016.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 233-137

Learners' Interactions in Massive Open Online Courses: Analysis and Interpretation

Anna A. Shurtina
Department of Open Education
ITMO University
Saint Petersburg, Russia
a.shurtina@cde.ifmo.ru

Andrey V. Lyamin
Department of Open Education
ITMO University
Saint Petersburg, Russia
lyamin@mail.ifmo.ru

Elena N. Cherepovskaya
Department of Open Education
ITMO University
Saint Petersburg, Russia
cherepovskaya@cde.ifmo.ru

***Abstract*—The research is devoted to the study of data on learners' interactions in massive open online courses. Based on the logs of online-learning platforms, the following research was made: a comparison of the behaviour of motivated and unmotivated learners regarding of video lectures, identification of the most valuable for the successful completion of the course activities of learners, creating a model of going through time-limited assignments and identification of cheating approach based on this model.**

The following conclusions were made: motivated and unmotivated learners watch video lectures in different ways, motivated learners appeared to be 14 times more active, the most interesting and most viewable videos were revealed. When identifying the most valuable theoretical materials influencing the successful completion of the course, the following results were obtained: some of the videos have a strong influence on the successful completion of the final assignment. Some of the videos appeared to have weak effect, they can be interpreted as non-obligatory. Ungraded tests have a positive but moderate effect on learners' success, while communication via discussion forum has no effect at all. In addition, a model of going through time-limited assignments was built using the average passing time of reliable learners, the approach for identifying cheating with examples is presented in the study.

Keywords— learning analytics, MOOC.

I. Introduction

The rapid development of information technology has an impact on all spheres of human life, education is no exception. Already for a long time, traditional classroom lessons have been accompanied by the use of electronic educational resources and distance learning technologies and ultimately have gone online.

The technology of massive open online course allows hundreds of thousands of learners to study for free from the best teachers at leading universities in the world, regardless of time or location, provided they have a device with Internet access [1]. Modern online-learning platforms have powerful tools for presenting educational content [2], allowing students not only to gain theoretical knowledge, but also to immediately work it out in practice, providing a full range of knowledge and skills in the chosen field.

According to the Class Central portal (https://www.class-central.com/), which regularly provides MOOC statistics, as of January 2019, there are more than 900 universities-providers of MOOCs all over the world, and the number of MOOCs exceeds 11 thousand. The number of learners studying online is more than 101 million people.

Despite the high popularity of MOOCs and their lightning-fast development, students and developers of online courses encounter a number of issues. First, due to the high rate of development of the MOOCs, there are gaps in terms of the pedagogical design of the courses and the methodological and technical quality of the materials presented. Secondly, students are faced with a lack of self-organization and motivation to complete the learning: the world statistics says that on average less than 13% of enrolled learners successfully complete the MOOC [2, 3].

Learning analytics - a set of approaches for collecting, storing and processing online learning data - appears to the rescue of these problems [4]. Millions of learner watch videos, solve problems, do practical exercises, communicate via discussion forums of MOOCs and so on. Data on these activities is stored in the so-called logs of online-learning platforms — all events caused by learners' clicks on the various components of online courses are recorded. Such data fits into the concept of the so-called "Big data" - huge amounts of information, the study of which is aimed at finding latent patterns and regularities [5]. By now, none of the online-learning platform provides detailed statistics based on data from logs, so a lot of research has been done in order to prove that learning analytics could be helpful and needs to be embedded and used [6]. Different approaches of learning analytics are aimed to support students' motivation to complete learning [7, 8], personalize and individualize learning [9], monitor and improve the quality of the content provided and create feedback needed for both learners and course developers [10]. Several of approaches applied are presented in the following study.

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4047
© 2019 Anna A. Shurtina, Andrey V. Lyamin, Elena N. Cherepovskaya.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

II. MATERIALS AND METHODS

Data from online-learning platform logs of tree online courses on technical topics was chosen for this study. These courses are selected as they were mentioned in the list of recommended online courses of All-Russian Olympiads, which characterizes them as recognized by the academic community and qualitative in terms of teaching materials and their technical presentation. Logs contain data on all the events caused by the learners' clicks on MOOC components with a mention of a timestamp, a link to the course component and other meta-data of each event.

A. Unmotivated and motivated video-interaction patterns

For the video-interaction research the course with 1354 learners was selected, who all together initialized more than 54,000 events. It is known that some learners are enrolled in a course with the condition that successful completion guarantees them a positive assessment in the relevant discipline at the university. The logins of these learners on the platform are known; in this study, the data of learners' interaction is separated into two groups called motivated and unmotivated.

The tendency of launches and completed views of video lectures by motivated and unmotivated learners is being observed. 48 video lectures with theoretical materials of the course with different duration from 88 to 1333 seconds were taken into account. All of the video lectures contain a screencast of presentation slides and a voiceover. Some of the lectures contain scenes with the course author speaking. Downloading and launching a video lecture from its beginning on the platform triggers a "load" event with a unique identifier of the learner and the video lecture. A completed view of a video lecture, when the slider on the roller duration scale reaches the extreme position, triggers the "stop" event with the same parameters. Those events, when the video lecture was stopped at the initiative of the learner within 5 seconds before the end of the video, are also taken to account. This is necessary because some videos finish with a non-informative screensaver, or the informative part of the lecture just stops a few seconds before the end, therefore, the learner watched the content part completely, but was not patient enough to initiate the "stop" event. When a learner clicks on a pause, the platform triggers the "pause" event, which contains information about the learner's unique identifier, video lecture, and pause timestamp. In this case, we will take into account those "pause" events that have a timestamp different from the total duration of the video by no more than 5 seconds.

All the events described above are to be processed for both motivated and unmotivated students in order to compare their behavior patterns and determine the most popular and the most unwatched videos in the course.

B. Identification of the most important for the successful completion of the course activities of learners

As each MOOC is aimed to provide learners with useful resources to gain knowledge and skills, descriptive analytics such as counting times when learners referred to course components can be not enough. The more informative data describes if the provided material really does help learners to study or maybe it does not influence their performance at all. Using the logs of the online-learning platform on which the course was placed, data about the interaction of learners with various types of components of the course was obtained, such as watching video lectures, taking ungraded quizzes, mostly multiple-choice and checkbox problems with a detailed solution provided after passing, interacting on forums and, finally, performing weekly test.

To determine the most important components for successful completion of the course, data of activities within the first week of the course was used. 63 students initiated more than 2100 log events that were taken into account. All of the above activities were calculated until spending the maximum number of attempts when solving the graded weekly test. Before counting, each activity was measured by three states: when a learner did not initiate any events connected with the course component, when a learner started to interact but didn't complete (e.g. loaded the video but did not watch it till the end, took quiz but did not succeed, etc.), and when a learner fully completed the activity.

In order to assess the importance of interaction with different components, the Pearson correlation coefficients between the measured activities of learners and their performance in the weekly test are to be found. The maximum degree of correlation is taken as 1. There are the following types of correlation: strong (± 0.7… ± 1.0), moderate (± 0.5 ... ± 0.7) and weak (± 0.3 ... ± 0.5). If coefficient is <± 0.3, the connection is practically absent. The minus or plus sign of the correlation coefficient indicates the direction of the connection — the plus sign means that the connection between the activity and performance is direct (positive), the minus sign is the reverse connection (negative). These coefficients are to show how interacting with different types of components influence learners' performance.

C. Creating a model of going through time-limited assignments

Online-learning platforms today have various automated tools for assessing learners' knowledge gained. One of the common tools is the usage of time-limited sets of problems, randomly distinguished from libraries. Using data about the time spent on solving problems, a model of going through such assignments can be built. This method of research is possible in the case when the difficulty of the task correlates with its average solving time.

As part of the study, the events of the solution of 329 tasks by 578 learners was processed. At the beginning of the study, it is necessary to determine all the events in which the leaners solve specific problems, taking into account the time they spent on it. In this study, the course on a technical topic is being observed, it means that the solution of a problem implies the execution of several calculated actions, i.e. it is impossible to solve

the problem in one step just by reading its condition. In addition, information on the final grade of the learners on the course will be required in order to determine the group of students, which results can be considered as reliable: learners who received good and excellent final marks during the course and completed all the activities of the course and study all the theoretical materials.

Accordingly, for each set of problems with the above data a solving model could be build using the upper and lower expected time limits counted with the 95% confidence interval of reliable learners' solving time. If the course instructor has to validate some learners' results of solving a set of problems, the actual time of solving of this learner can be compared with the constructed model to check whether the actual time correlates with the expected time and falls into a certain confidence interval or not.

III. RESULTS AND DISCUSSION

A. Unmotivated and motivated video-interaction patterns

The number of motivated students of the chosen course is 109 people, i.e. 8% of the total. The number of events caused by the actions of learners in relation to video content was counted. In total, 54,303 events were recorded, of which 24,348 (or 45%) were caused by the actions of unmotivated learners, 29955 (55%) were caused by the actions of motivated learners. Consequently, one motivated learner interacts with video content approximately 14 times more active than the unmotivated one.

Figure 1 shows the graphs of the distribution of launches (black) and completed views (grey) of video lectures by unmotivated learners over the weeks of the online course.

Unmotivated learners show a high interest to watching videos only in the first week of the online course, by the second week the interest is reduced three times, on the subsequent weeks the number of launches is on average 20 times less than in the first week of the online course. This confirms the hypothesis that the overwhelming majority of learners drop out at the very beginning of the course due to lack of motivation.

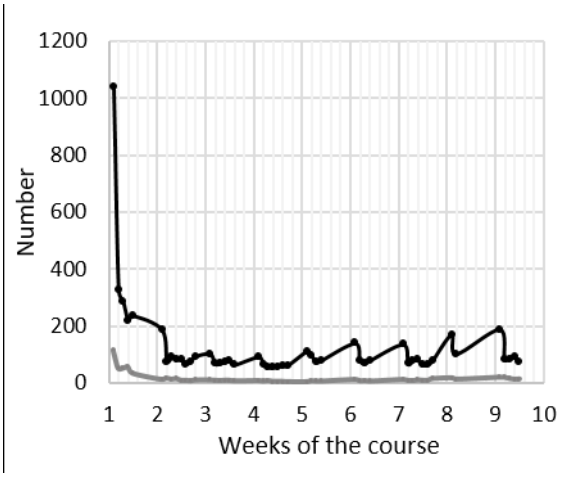


Fig. 1. Launches (black) and completed views (grey) of videos by unmotivated learners throug the course

In addition, it was revealed that at weeks 1–3 and 7–10 the percentage ratio of completed video lectures to launches is on average 11%. This means that 11% of started views were completed. At 4, 5 and 6 weeks of the online course this value is 6%, therefore, it can be assumed that at these weeks of the online course the video content was too complicated, too simple or of poor quality from a technical or methodological point of view, it is necessary to work to identify the causes of low interest of learners.

Figure 2 shows the graphs of the distribution of launches (black) and completed views (grey) of video lectures by motivated learners over the weeks of the online course.

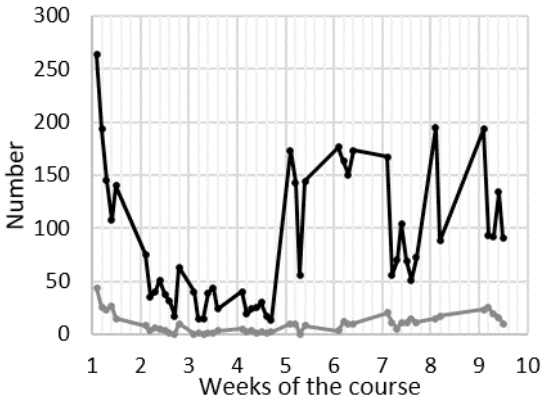


Fig. 2. Launches (black) and completed views (grey) of videos by motivated learners throug the course

Unlike the group of unmotivated learners, there is a different trend. The number of launches and completed views after the first week is reduced by half. In the period from the second to the fourth week of the online course, there is also a decrease in the activity of learners to launch video lectures, however, interest grows again in the fifth week and remains at a high level until the end of the course, except for materials of the seventh week. However, in relative terms, the seventh week of the course is one of the most popular in terms of completed views: 16% of the started views were completed, when for a fifth week, which is the most successful one in terms of the number of video loadings, this value is only 5%.

The main result of the observation of these data is that motivated and unmotivated learners do behave differently in interacting with content of a MOOC, so course instructors need to determine a reliable group of learners whose interaction data can be considered as informative feedback.

B. Identification of the most important for the successful completion of the course activities of learners

Results of calculating Pearson correlation coefficients between measured learners' activities and their performance on the first week of a course are represented in the Table I.

TABLE I. Correlation Coefficients Between Learners Activities and Peformance

	Weekly Test Performance
Watching Video 1	0,32
Watching Video 2	0,76
Taking Quiz 1	0,61
Watching Video 3	0,71
Taking Quiz 2	0,63
Watching Video 4	0,48
Taking Quiz 3	0,61
Watching Video 5	0,49
Taking Quiz 4	0,52
Discussing on Forum	0,02

The results can be interpreted as follows: important for successful completion of the first weekly test are watching video 2 and watching video 3. All the quizzes of the first week of the course have a moderate impact on the success of the final test. The first, fourth and fifth videos have a weak effect. Communication on the course forum does not correlate with the success in the weekly test at all.

As recommendations for the course developers, the following can be singled out: video 1 is to be placed in a separate section, since it seems to be introductory and has a weak effect on the successful completion of the first weekly test, the difficulty of quizzes in the first week of the course is to be increased by changing the problem type, e.g. from close to open response form, to make them more challenging.

As recommendations for learners, the following can be singled out: videos 2 and 3 must be viewed till the end and all surveys are recommended to be resolved, since these activities will most likely help them successfully complete the weekly test.

C. Creating a model of going through time-limited assignments

The correlation coefficient between solving time and problem difficulty for a group of students who successfully completed the course appeared to be 0.7034, which indicates a strong connection and the possibility of building a model for problem sets based on this data. For each problem, a 95% confidence interval of the solving time by reliable group of students was calculated. After that, every learners' solution can be placed on the graph with a created model.

For example, Figure 3 shows the model of passing the set of 13 problems, the green dashed line indicates the maximum expected solution time, the blue dotted line indicates the minimum expected solution time, the red solid line indicates the actual time of problem solving by the learner.

Fig. 3. Maximum expected time (green), minimum expected time (blue) and actual solving time (red) of a specific learner

The graph on Figure 3 shows that the solution of some problems falls within the confidence interval, the solution of some problems takes more or less time than expected. However, the graph clearly shows the dynamics of the time spent on the decision, which may indicate that the learner solved each task independently.

Figure 4 shows a graph of the solving time of another set of problems by another learner.

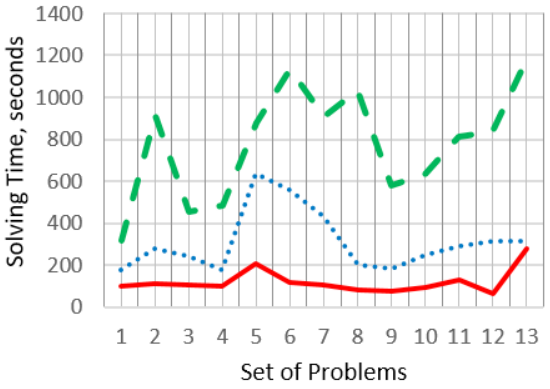


Fig. 4. Maximum expected time (green), minimum expected time (blue) and actual solving time (red) of another specific learner

The result from the Figure 4 can be interpreted unequivocally - the learner spends about the same amount of time on each problem, none of the problems' solving times do not fall into the expected interval. In case if all the problems were solved successfully, it can be concluded that the student was cheating.

Two of the most illustrative instances of applying the approach are shown in this study. Such identification can be required when a high score of solving a set of problems can be put by the course instructor into doubt and verified by comparing the actual solving time with the expected pattern.

IV. CONCLUSIONS

The implementation of the learning analytics methods presented in the study allows determining the activity and involvement of learners, finding a correlation between the activity of learners and their performance, forming recommendations for course developers and learners, as well as building models for passing problem sets and identifying cheating. Using the considered methods in e-learning systems can help authors of online courses assess the quality indicators of published content, and learners to maintain motivation and successfully complete learning on the course.

References

[1] Baturay, M. (2015) An Overview of the World of MOOCs. Procedia - Social and Behavioral Sciences, 174, Ankara: Elsevier, pp. 427-433

[2] Davis, D., Jivet, I., Kizilcec, R.F., Chen G., Hauff, C., Houben, G.J.: Follow the successful crowd: raising MOOC completion rates through social comparison at scale. In: Proceedings of LAK 2017, pp. 454–463. ACM (2017)

[3] Gomez-Zermeno, M. G., & Garza, L. A. (2016). Research analysis on MOOC course dropout and retention rates. Turkish Online Journal of Distance Education, 17(2), 3–14.

[4] Song, Donggil. (2018). Learning Analytics as an Educational Research Approach. International Journal of Multiple Research Approaches. 10. 102-111. 10.29034/ijmra.v10n1a6.

[5] D. Gibson, "Big Data in Higher Education: Research Methods and Analytics Supporting the Learning Journey," Technol. Knowl. Learn., vol. 22, no. 3, pp. 237–241, Oct. 2017.

[6] Banihashem, Seyyed Kazem & Aliabadi, Khadijeh & Ardakani, Saeid & Delavar, Ali & Nili Ahmadabadi, Mohammadreza. (2018). Learning Analytics: A Critical Literature Review. 9. 10.5812/ijvlms.63024.

[7] Alario-Hoyos, Carlos & Estévez-Ayres, Iria & Pérez-Sanagustín, Mar & Delgado-Kloos, Carlos & Fernández Panadero, Carmen. (2017). Understanding Learners' Motivation and Learning Strategies in MOOCs. The International Review of Research in Open and Distributed Learning. 18. 10.19173/irrodl.v18i3.2996.

[8] Onah, Daniel & Sinclair, Jane & Boyatt, R. (2014). Dropout Rates of Massive Open Online Courses: Behavioural Patterns. 10.13140/RG.2.1.2402.0009.

[9] Maseleno, Andino & Sabani, Noraisikin & Huda, Miftachul & Ahmad, Roslee & Jasmi, Kamarul Azmi & Basiron, Bushrah. (2018). Demystifying Learning Analytics in Personalised Learning. International Journal of Engineering and Technology. 7. 10.14419/ijet.v7i3.9789.

[10] Michos, K., Hernández-Leo, D., Jiménez, M., (2017) How educators value data analytics about their MOOCs, CEUR Proceedings of Work in Progress Papers of the Experience and Research Tracks and Position Papers of the Policy Track at EMOOCs 2017 co-located with the EMOOCs 2017 Conference (Vol-1841), Madrid, Spain, 77-82

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 238-242

Competency-Based Training in Hydrological Education

Larisa Timofeeva
Russian State Hydrometeorological University
Saint Petersburg, Russia
tilarisa@gmail.com

Zinaida Timofeeva
Peter the Great St.Petersburg Polytechnic University
Saint Petersburg, Russia
whoiszoe@gmail.com

***Abstract*—Hydrology is an interdisciplinary science, incorporating aspects of many Earth Sciences. Key hydrological tasks such as floods prediction, quantitative and qualitative assessment of water resources and the environmental status of water bodies become increasingly important and challenging. Latest advances in scientific and technological developments require update training. The World Meteorological Organization (WMO) encourages increasing capacity of its Training Centres to address the rapidly developing demand for improved services. However, adequate education is generally not yet available. Implementation of competency-based training is needed, including active learning approaches. The paper presents international experiences and results of the use of project-based and flipped learning while teaching hydrological curriculum at Russian State Hydrometeorological University, Saint Petersburg. Some pros and cons of these approaches and difficulties of their implementation are discussed. Usually, students improve both their competence to work independently and solve problems collaboratively. Tackling near real issues and dealing with case studies get them more engaged in educational process and enhance practical outcomes. The feedback from the students and exam results have proved the efficiency of these approaches.**

***Keywords*—competency, flipped classrooms, hydrology, project.**

I. Introduction

Competency-based training (CBT) focuses on learners' practical achievements. In the late 1990s, the use of new techniques and data in hydrometeorology caused an increasing demand for more practical training outcomes. The CBT era started. Such approach is encouraged by the World Meteorological Organization (WMO) and applied in its 28 Regional Training Centers, according to Competency Requirements for Education and Training Providers [1]. Russian State Hydrometeorological University (RSHU) has been one of them since 1994.

Hydrological curriculum has been taught at RSHU since 1930s and great experiences were gained. However, the content of former five-year Engineering Programmes, even dramatically compressed, can't be included in four-year Bachelor Programmes. They stay lecture-based, don't guarantee practical outcomes and are abundant with out of date content, which doesn't correspond to the students' professional future. Unfortunately, trainers often teach students in a traditional face-to-face lecture style, nearly the same way they were taught and are most familiar with. However, this strategy is not the most effective for engaging students. Recent research has shown a potential for great variability within the hydrological curriculum [2]. Implementing active-learning approaches based on available open resources trainers can serve as more of a "coach" or "mentor" to the students. Teaching this way requires fulfilling modern competencies.

According to [2], personnel should be competent to: (1) Identify learning needs and specify learning outcomes; (2) Determine a learning solution; (3) Design learning activities and produce learning resources; (4) Deliver training and manage the learning experience; (5) Assess learning and evaluate the training process.

The WMO supports building training capacity through courses and workshops delivering by the Education and Training Office (ETR) on- and offline. The WMO Global Campus concept encourages individuals and institutions to move away from considering only what they can develop or deliver themselves to how they can benefit from or contribute to the wider WMO ETR community. Useful resources are available at the Training Portal in the WMO web-site [3]. There are resources on active learning strategies, used within the CBT framework.

Project-based learning (PBL) and Flipped Learning (FL) have become quite popular in engineering education, since they allow more than just gaining required professional competencies, which are likely changing during the professional development of an individual. These approaches benefit the ability to learn and develop continuously (lifelong learning), to work in teams and to be able to integrate between sciences studied at University.

II. Materials and methods

At RSHU, Basic Hydrology is taught for two semesters to the second and third year students. The number of them in a group vary from 15 up to nearly 30. There is one lecture and one practical class a week.

Traditionally, the content is presented topic by topic. Such format does not always demonstrate how the course

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4067
© 2019 Larisa Timofeeva, Zinaida Timofeeva.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

topics fit together as a whole and how they can be used to solve real professional problems and help obtain practical outcomes from the Course.

Since the course is completed with an exam and a course project, it was decided to start preparing the students for such challenges during the first semester. For the last two years, FCl and PBL were applied to develop deep content knowledge as well as critical thinking, creativity, and skills communication, e q professionals skills, which can make current students more competitive at job market.

A. Flipped Classrooms

The simplest definition of a Flipped Classrooms approach is expressed as "what is done at school done at home, homework done at home completed in class". This approach was introduced in the 1990s by Harvard Professor Eric Mazur, but came into general use in the early mid-2000s when it was popularized by chemistry teachers Jon Bergman and Aaron Sams [4] and the founder of the Khan Academy, Salman Khan. In FL, part or all of direct instruction is delivered through videos and other media; and the class time is used for engaging students in collaborative, hands-on activities [5]. The traditional FCl model is presented in Fig. 1.

As seen in Fig. 1, FCl training needs changes in content, teacher's role and delivery mode.

It should be noted, that reading and studying the theory before class is not a novelty, but it used to be more popular while teaching humanitarian disciplines. The terms Flipped Classrooms and Flipped Learning are not interchangeable. FCl does not necessarily lead to blended FL, which requires full incorporating of the following four pillars: flexible environment, learning culture, intentional content, professional educator [7].

Since the entire Basic Hydrology course has not been flipped yet, we can only talk about FCl. The topics, taught during the first semester, are all about rivers and processes on their catchments. Studying some of them requires solid theoretical basis, so they are particularly suitable for FCl. Successful implementation of FCl in engineering education requires relevant resources available (including online ones), which were rather limited recently. Unfortunately, due to the lack of such resources in Russian only few topics can be taught as FCl.

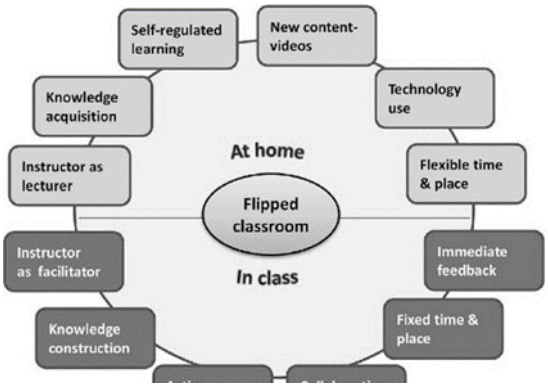


Fig. 1. The components of the traditional flipped classroom model [6].

Snowmelt processes, the most complex topic, was chosen for the experiment, since understanding the physical nature of snowmelt and methods for its intensity calculation is time consuming and needs advanced preparation. Luckily, in 2011, the authors were introduced to the Community Outreach Missions Education Team/ University Corporation for Atmospheric Research (COMET/UCAR) program, which offers e-lessons and courses on Earth Sciences. Some lessons were translated into Russian and are used for teaching at RSHU. Original lessons are available on the web-site [8]. After completing the Snowmelt Processes lesson, students have an on-line Quiz and the results are sent to the teacher automatically. Passing the Quiz is a kind of a formative assessment, which is often neglected, but should not be, since a proper assessment enhances learning, and lack of assessment weakens training process. Assessment can be carried out via MOODLE.

Besides completing the e-lesson, later on students are asked to explore the data base, available on the web-site of the World Data Center of Roshydromet (www.meteo.ru), and try to get the data necessary for carrying out further snowmelt intensity calculations.

Thus, such pre-lesson activities save time for a more detailed discussion of the process being studied and dealing with possible difficulties of data access and calculations. Thankful to applying FCL, it has become possible to complete this practical work in a near real mode. Earlier, this work, due to being extremely time and efforts consuming, was done in a simplified form, based on the data and templates, prepared by trainers. This couldn't make students more competent in their future profession. Obviously, applying blended flipped classrooms has resulted not only in deeper theoretical understanding (Comet lessons, video resources and other media), but in important practical outcomes: students become more prepared for independent professional activities (use of data bases, technical means, ability to communicate). FCL can be considered as a powerful CBT approach, which can alter even Bloom's Taxonomy (Fig. 2). Thus, the trainers' role changes from being a main information source (lower Taxonomy levels) to coaching the students while working together on analyzing and evaluating of the training content (upper Taxonomy levels).

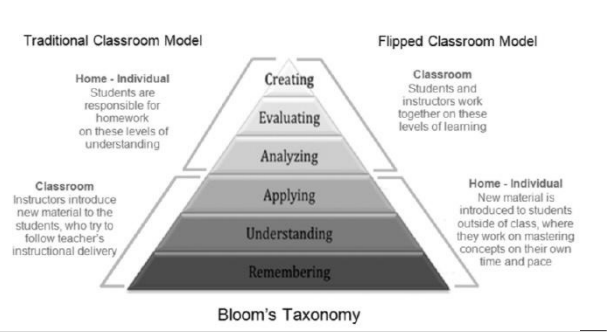


Fig. 2. Connection between Traditional and Flipped Classroom to Bloom's Taxonomy [9].

B. Project-based training

Educators agree that students learn better, when they engage in complex problems and projects [10]. So, if there is a project planned in the Course, it is wise to make the best of it! If a project is not included, a trainer can turn one or several practical works into small projects.

While teaching Practical Basic Hydrology different ways are used. During the first semester, three practical works are combined into a small project (P1), during the second semester there is a planned project (P2). In both cases, students work on authentic, meaningful projects for several weeks. It allows self-pacing and more time for sorting out real problems or answering complex questions. In our case, intended outcomes from both P1 and P2 were enhanced competencies to collect, analyze and apply data for solving hydrological tasks.

P1 consists of three practical works (including flipped Snowmelt Processes, described above), which are carried out for a selected river and for a particular year. Thus, younger students work on the same tasks but with different objects (data). They have to learn how to collect hourly, daily, monthly and annual data on a number of hydrometeorological parameters from traditional (Monthly weather reviews) and modern (e-databases) sources. The experiences gained will make them competent data users. Additionally, application the data for studying water regime and catchment processes (analytically and graphically) ensure students' deeper hydrological understanding. When students collect all the three works together and analyze them, they can realize how melting water forms high flow and rains become runoff. This might be their first professional finding! Besides, such collection of small projects might start a further research.

Having successfully completed P1, students are positively aimed towards more complex P2. For years, course projects on hydrology were typical and covered only a few topics. Such situation was explained with data and techlogies limitations. It resulted in lack of original researches: students often tried to copy earlier completed projects with little change. Nowadays, former limitations are easily coped, and working on projects have become a powerful CBT means. The key problem is to choose a project topic, which would engage students and, possible, become a part of their Bachelor thesis. The bulk of the work is done independently, based on the experiences gained earlier, but the trainer's help is needed. At the end, students demonstrate their theoretical achievements and practical skills by developing presentations for their peers, who assess them according to specific rubrics.

It should be noted that carrying out P1 might not be enough to prepare less advanced students for independent P2. If so, it might be useful to complete an additional pre-project, which can be done in small groups of two to four. Thus, the less advanced students are, the more project-based training is required.

As a result, not only intended outcomes are achieved, but students also develop deeper content knowledge as well as critical thinking, creativity, and communication skills. Project based training forms friendly environment among students and teachers.

III. Results and discussion

The first author's experiences of applying CBT through Project Based Training and Flipped Classrooms in hydrological education at RSHU have been quite positive. However, there are more evidence on these approaches pros and cons for both, students and trainers.

As for students, it should be taken into account that they learn differently and not all of them accept active learning, which requires to be engaged in the learning activity in order to realize any learning benefit. When they're not engaged or not willing to work at home, those approaches don't work perfectly.

The main benefit for students is enhanced professional competencies and success at exams. Even though the students who were taught through FCl and PBT don't do better at exams than the students, who were taught traditionally, they are more exposed to further teamwork, leadership and communication skills development. As a surprising benefit, replacing lecturing with in-class work and discussion has become personally rewarding. The variability in student performance has decreased [11], since they could learn from peers, class discussions and trainers' immediate feedback. Except training undergraduate students, individual FCl seem to be useful while teaching Master students, who often have different background.

As for trainers, there are certain difficulties. First, Competency Requirements for hydrologists are not approved by the WMO yet. So learning outcomes should be thought out thoroughly and discussed with stakeholders, if needed. Second, trainers' competencies include choosing technology and software required for delivering training, preparing presentations and learning resources, applying learning activities that include authentic tasks, build upon the prior knowledge of students. No doubt, that implementing FCl and PBT leads to additional workload on trainers. So, they need help from University Administrations, experts in pedagogy and colleagues. Obviously, flipping entire courses is possible only if responsible authorities approve and support it. The more flipped classrooms are there, the higher students' learnability and better entire learning outcomes will be.

Luckily, a lot of resources on FCl and PBT are available now. Among them are: web-sites, which offer numerous theoretical and practical materials [5], [12], the American Society for Engineering Education [13], specialized resources for the geoscience community provided by The United States Geological Survey [14] and a free collection of numerous training resources, designed by COMET UCAR [8].

Conclusions

Opportunities in hydrologic science have never been greater, and the challenges that lie ahead have never been more compelling [15]. While there is a rising interest in and demand for civil engineering and hydrology educa-

tion, some have suggested a widening gap between how students are instructed in hydrology, and the subsequent professional skill set required for a career as a hydrological engineer [16]. According to the WMO requirements, hydrological training should be delivered by competent personnel, be competency based and include modern active learning approaches. Based on the experiences of teaching Basic Hydrology to undergraduate students in RSHU, it is possible to conclude:

1. Flipped-classrooms are easier applied to younger students, since all their stages are fully controlled by teachers and few specific skills are required from the students. They study some theoretical content, prepared by the teacher, in advance. This saves time for mastering practical skills at classes (data bases, Excel, GIS, AutoCAD). There is more time to develop students' competency to work independently.

2. Project-based training better suits older students, since it is a cooperative, rather time and efforts consuming approach. A real on-job environment can be created while solving authentic problems. Students can improve their competency in face-to-face communication and sorting out problems collaboratively to get practical learning outcomes.

The feedback from students and exam results have proved the efficiency of CBT.

References

[1] Competency Requirements for Education and Training Providers. WMO, 2013. [Online]. Available: https://library.wmo.int/pmb_ged/etrp_competency-requirements_en.pdf, [Accessed Feb. 10, 2019].

[2] T. Wagener, C. Kelleher, M. Weiler, B. McGlynn, M. Gooseff, L. Marshall, T. Meixner, K. McGuire, S. Gregg, P. Sharma, and S. Zappe, "It takes a community to raise a hydrologist: the Modular Curriculum for Hydrologic Advancement (MOCHA)," Hydrol. Earth Syst. Sci., 16, 3405–3418, 2012, [Online]. Available: Hydrology and Earth System Sciences https://www.hydrol-earth-syst-sci.net/16/3405/2012/ [Accessed Mar. 01, 2019], https://doi:10.5194/hess-16-3405-2012

[3] WMO Education and Training Programme, [Online]. Available: https://etrp.wmo.int/course/index.php?categoryid=49 [Accessed Mar. 01, 2019].

[4] J. Bergmann and A. Sams, "Flip Your Classroom: Reach Every Student in Every Class Every Day", Washington DC: International Society for Technology in Education, pp. 120-190, 2012.

[5] FLIP Learning, [Online]. Available: https://flippedlearning.org/board/ [Accessed Feb. 21, 2019].

[6] I. Blau and T. Shamir-Inbal, "Re-designed flipped learning model in an academic course: The role of co-creation and co-regulation", Journal Computers & Education archive, vol. 115, Issue C, pp. 69-81, Dec. 2017. [Online]. Available: https://www.researchgate.net/publication/318831616_Re-designed_flipped_learning_model_in_an_academic_course_The_role_of_co-creation_and_co-regulation [Accessed Feb. 21, 2019], DOI: 10.1016/j.compedu.2017.07.014.

[7] The Four Pillars of F-L-I-P, 2014. Flipped Learning Network. [Online]. Available: https://flippedlearning.org/wp-content/uploads/2016/07/FLIP_handout_FNL_Web.pdf [Accessed Feb. 20, 2019].

[8] COMET MetED, [Online]. Available: https://www.meted.ucar.edu/index.php [Accessed Feb. 21, 2019].

[9] A.P. Lopes and F. Soares, Flipping a mathematics course, a blended learning approach. 12th International Technology, Education and Development Conference INTED2018, Mar. 5-7, 2018, Valencia, Spain. [Online]. Available: https://www.researchgate.net/publication/323900654_FLIPPING_A_MATHEMATICS_COURSE_A_BLENDED_LEARNING_APPROACH, [Accessed Feb. 15, 2019]. DOI: 10.21125/inted.2018.0749.

[10] T. Wagener, M. Weiler, B. McGlynn, L. Marshall, M. McHale, T. Meixner and K. McGuire, "Taking the pulse of hydrology education", Hydrol. Proc., 21, 1789-1792, 2007. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.6766, [Accessed Feb. 20, 2019]. https://doi.org/10.1002/hyp.6766

[11] S. C. Shah-Fairbank and K. Lamb, Comparison of traditional, hybrid and flipped classrooms for water resources design courses: American Society for Engineering Education, Pacific Southwest Section Meeting, April, 2017, Tempe, Arizona. [Online]. Available: https://peer.asee.org/29209, [Accessed Feb. 25, 2019].

[12] Buck Institute for Education PBLWORKS. [Online]. Available: https://www.pblworks.org/, [Accessed Feb. 20, 2019].

[13] American Society for Engineering Education. [Online]. Available: https://www.asee.org/, [Accessed Feb. 15, 2019].

[14] U.S. Geological Survey. [Online]. Available: https://www.usgs.gov/, [Accessed Feb. 15, 2019].

[15] W. J. Gutowski Jr., G. M. Hornberger, E. Bernhardt, W. E. Dietrich, D. Entekhabi, G. E. Fogg, E. Foufoula-Georgiou, W. B. Lyons, K. W. Potter, S. W. Tyler, H. J. Vaux, C. J. Vorosmarty, C. Welty, C. A. Woodhouse and C. Zheng, "Challenges and Opportunities in the Hydrologic Sciences". Geological and Atmospheric Sciences Reports, 4, 2012. [Online]. Available: https://lib.dr.iastate.edu/ge_at_reports/4, [Accessed Feb. 21, 2019].

[16] M.M. Lombardi, "Authentic Learning for the 21st Century: An Overview", ELI Paper 1, 2007. https://www.researchgate.net/publication/220040581_Authentic_Learning_for_the_21st_Century_An_Overview, [Accessed Feb. 16, 2019].

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 242-246

Triggering the Students' Positive Attitude for the Studies of Engineering Graphics Courses through the Augmented Reality Content

Zoja Veide
Dep. Of Computer Aided Engineering Graphics
Riga Technical University
Riga, Latvia
Zoja.Veide@rtu.lv

Veronika Strozheva
Dep. Of Computer Aided Engineering Graphics
Riga Technical University
Riga, Latvia
Veronika.Strozeva@rtu.lv

***Abstract*—Contemporary researches show that triggering interest can enhance students' thinking skills, creativity, joy of learning and improved performance. The digital technologies can be beneficial for students' learning engagement and motivation. The digital context helps students to keep positivity in education process that allows to activate attention, memory, and makes learning skills better. Augmented Reality is one of technologies that lets you interact with the real world and virtual objects at the same time. Augmented Reality application enables faster comprehension of complex spatial problems and relationships, which makes it useful during the learning process of engineering graphic courses. In this paper the application of AR technologies in RTU engineering courses is described. The results of this approach will be useful for vocational school teachers, college and university teachers in the successful integration of AR technologies in the teaching process.**

***Keywords*—Augmented Reality, Engineering Education, Trigger.**

I. Introduction

The modern era is a period when the Internet and computers occupy a central place. The current generation is actively using innovative Internet technologies and digital equipment. Today's students spent their entire lives surrounded and using computers, video games, camcorders, mobile phones and all other toys and tools of the digital age. Modern students today are "native speakers" of the digital language of computers, video games and the Internet. Children first start playing games, and only later do they start learning to read and write, or these processes occur in parallel. Now it became clear that because of this omnipresent environment and the huge amount of their interaction with it, modern students think and process information in a fundamentally different way than their predecessors. They would like to get the necessary information quickly. They love the parallel process and multitasking. They prefer graphics rather than text and the game to "serious" work.

A lot of studies provide insight into the pedagogical use of digital technologies and how to trigger today's generation interest in learning using contemporary technologies. The trigger is a factor that motivates students to study and be passionate about learning. Many researchers have noted that the trigger can improve the learning motivation of students and lead to good activities during classroom [1], [2], [3]. In the educational process, a trigger is applied to promote student interest in study and support their learning. A trigger can be represented in various ways, such as videos, slides, games, animation, puzzles, textual information, problems, etc. The purpose of the triggers is to help students adapt to the courses, meet their expectations and in this way enjoy learning.

New hardware and digital tools introduced by technological developments have significantly changed the education methods, teaching process, access, use and present information. The rapid development of information technologies, including mobile devices and innovative technologies, has stimulated researches in the field of education and educational technologies since the methods of the 21st century are technology-oriented methods [4], [5]. Specific aspects of engineering graphics teaching and key principles of the elaboration of educational computer games for developing spatial imagination are studied just occasionally [6]. An attempt to measure an engineering graphics literacy was made by means of more entertaining "technological toy" – a contemporary 3D parametric modelling CAD software [7]. The students were required to prove their spatial comprehension and the skills of the interpretation of the assembly drawing by modelling in a limited period as many individual parts as possible.

The use of unattached devices and mobile applications in the educational process is becoming more common. The application of augmented reality interests as a technology that allows interacting with digital objects in a real environment through mobile applications [8]. The interactive environment between the virtual and the real world is created by supplementing virtual objects, such

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4184
© 2019 Zoja Veide, Veronika Strozheva.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

as computer text, video, graphics, and other information, into the real world. The real environment is visualized using cameras such as computers, smartphones, and tablets [9]. In this way, augmented reality is an environment where virtual objects and the real world exist together. Augmented reality technology can have a significant impact on the teaching process because it makes it possible to combine the real world with virtual learning materials and gives students the opportunity to manage these materials [10], [11]. Using this technology, a person can interact with the virtual environment and has an opportunity to actively develop the cognitive process, enrich himself with experience and have the opportunity to learn by the method of discovery. From this point of view, augmented reality provides an expanded learning environment and supports creative learning principles [12], [13].

In this paper the results of a study on the use of AR technology in teaching courses "Descriptive Geometry and Engineering Graphics" and "Civil Engineering Computer Graphics" for the students of RTU are presented. A particular attention was paid to the study of the students' feedback concerning AR application and the attitude of students to the new methodological approach introduced.

II. Augmented Reality in Learning Environments

A didactic toolkit AR-DEHAES [14] was used as an additional learning material for the course "Descriptive Geometry and Engineering Graphics" in first-year studies. The main aim was to trigger their interest in the studies they do not enjoy in a traditional way with paper and pencil, which was used for decades before in engineering studies. This was supposed to improve their spatial abilities and to help them for a better understanding of the course content and motivate students to study more deeply as well [15]. This toolkit is based on the application of the principles Augmented Reality (AR) and requires a computer and a webcam. These two "technological toys" were supposed to act as a triggering aspect. During the operations, virtual elements are visualized on the monitor. The AR-DEHAES toolkit consists of software and a book that contains exercises that must be solved by students. Each task has a marker for visualizing virtual 3D objects. The registration of virtual elements in the real world requires accurate tracking of the position and orientation of the marker (using a marker-based method). The main element of the marker is a black square containing symbols. For the program, the camera to capture the real environment and track the marker is required. The integration of the real world with a 3D virtual model occurs when the main marker is fixed by the camera. To recognize virtual objects and to display them on the screen, a marker is used, which is placed with a specific exercise.

By turning and moving the marker or changing the distance from the marker to the webcam, you can see different perspectives of the virtual model and get additional information for the exercise performance (Fig. 1). The didactic material in its structure has five levels of complexity, each of which contains several types of exercises such as: defining of surfaces and vertexes on orthographic and axonometric views; construction of orthographic views, selection of the minimum number of views for construction of an object; sketch a missing orthographic view using two views of an object; drawing of all orthographic views. To verify the correctness of the solution and the correspondence of their sketches by hand to 3D virtual models, students can visualize the 3D model in the AR environment and view the result.

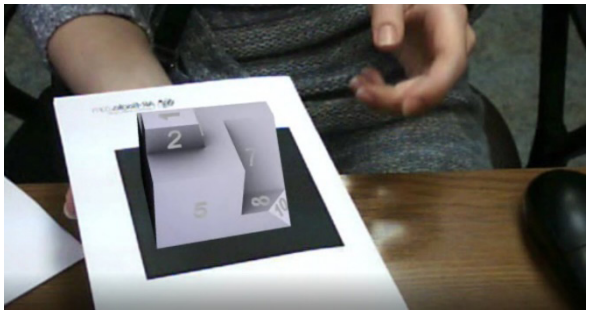


Fig. 1. Study process with application AR-DEAHAES.

During the academic semester, students became familiar with the augmented reality technology tool - the AR-DEHAES software and worked with them performing corresponding AR-BOOKS tasks. At the end of the semester, students conducted a self-assessment of spatial perception before and after using the AR-DEHAES software, using a qualitative research method – a test form. As well as spatial abilities of engineering students were measured before and after training through Mental Rotation Test. Engineering students who learned the subject with AR-training and control group mechanical engineering students having regular course performed test. Table 1 shows the scores obtained by students in the MRT test.

Table 1 Mean pre- and post-test and gain test scores (standard deviation) for experimental and control groups.

	Pre-test	Post-test	Gain
Experimental group n=48	18.12 (5.91)	23.45 (4.05)	5.33 (4.31)
Control group n=24	17.42 (5.39)	21.83 (5.08)	4.41 (4.26)

For the statistical analysis we used a Student's t-test, taking as the null hypothesis the fact that mean values for spatial visualization abilities did not vary after the end of the course. The t-test for paired series was applied and the ρ values are $\rho = 0.00000035 < 0.001$. Hence the null hypothesis is rejected, and we can conclude, with a significance level of higher than 99.9 percent, that the mean scores for the experimental group underwent a positive variation. An analysis of variance (ANOVA) was performed to determine the effect of the course type (regular or with AR training) on MRT. The analysis shows there was no significant differences between groups ($F = 0.598$, $\rho = 0.442$).

These students worked with this software for the first time, so their self-assessment objectively deserves attention. The following data was obtained from the study,

the most important of which are:

• 53% of students at the beginning of the PG and IG courses had prior knowledge of engineering graphics acquired in primary or secondary school. All students rated their spatial perception as good (50%) and mean (50%);

• 19% of students can always present the result starting to perform the graphical task. Others can do this only occasionally;

• All students recognize that working with AR-DEHAES, one can better understand the position of invisible points and edges of an object in space and allows keep interest on learning;

• 81% of respondents admit that after working with AR-DEHAES it is easier to perceive and visualize spatial objects, from which it can be concluded that augmented reality technologies help students develop spatial perception.

• 62.5% of students indicate that there are gaps in the software that should be eliminated, for example, increasing the grey contrast on surfaces and the language used should be English;

• All students, except one, recommend using this software for other students and assessed tool as entrancing and useful in the education process;

Results show that all students expressed a positive attitude to the AR material. Most students considered it interesting and they were satisfied with the additional study material technology and methodology. According to results in the Mental Rotation Test, the course "Descriptive Geometry and Engineering Graphics" enhanced with AR technology-based training material improves the spatial abilities of students (5.33 points in MRT compared with 4.41 points, obtained in a "regular" engineering graphics course).

To help students understand the training materials of the course "Descriptive Geometry and Engineering Graphics", the staff of our department prepared 3D objects from graphical tasks for placement in the AR environment. Also, according to our experience in application AR-DEHAES, the use of AR models triggers interest in the study of the graphical subject.

Augmented Reality scenes were created using the Envisage AR software. Virtual 3D models were superimposed on the environment and visualized through a computer's webcam so that they looked like part of real space. During the operation of this software, marker-based tracking is used, which means that 3D models are displayed attached to physically printed markers (Fig2, 3).

For each object, the individual marker and the 3D model were constructed, and thus the AR scene was created. To make a 3D-model was used SolidWorks. Those models were saved as STL files for subsequent import into the AR scene.

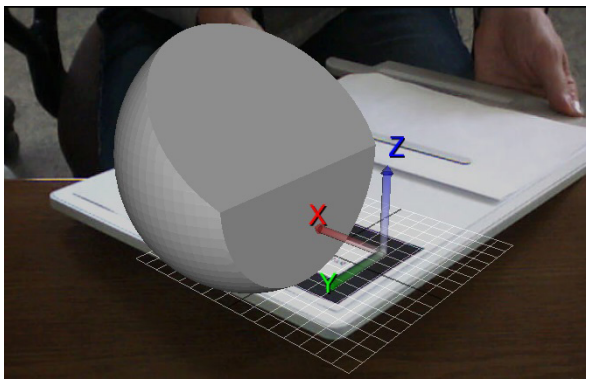


Fig. 2. AR model for exercise on Descriptive Geometry topic.

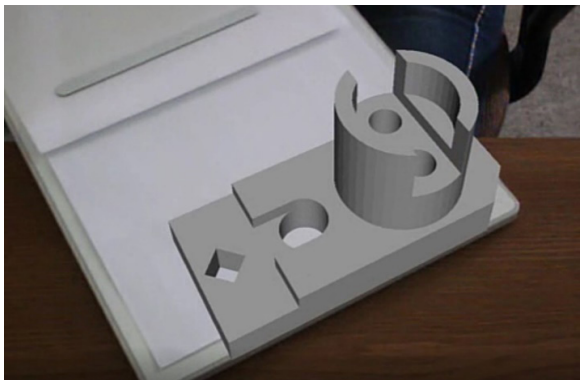


Fig. 3. AR model for exercise on Engineering Graphics topic.

To assess the effectiveness and convenience of the use of AR models in the course, at the end of the semester there were surveys in which students shared their views on AR technology. All the students found this approach very useful in solving graphic exercises. It was recognized as very interesting in general and in particular very useful for the topic of construction of multi-view projections from 3D geometric objects. In latter case it was possible to observe the transformation of the 3D AR model into 2D projections, which could be interactively manipulated in real time by observing the result on the monitor screen. The overall response of the students about the use of AR technology in the "Descriptive Geometry and Engineering Graphics" course was very positive.

Fast development of computer technology provides the possibility of the use of mobile Augmented Reality systems which are running on the smartphones. The application AR technology in the learning process of the course "Civil Engineering Computer Graphics" is based on the participation of students in the creation of AR scenes and the use of mobile devices for AR objects visualization. The students visualize their own 3D models, created by means of ArchiCAD software, using Augment – an AR SaaS platform that allows visualizing objects in the real world and in real-time through tablets or smartphones. In the quality of AR objects, the 3D digital model of a building and 3D model of the roof are used. "Roof construction" and "Building model creation" – individual exercises for students of this course.

Platform Augment has a mobile application for visualization of 3D models in AR and as well as the web interface called Augment Manager using for content

management. Augment Manager allows you to upload 3D models and trackers, edit them and share models. According to the task, students use the tracker-based method in this work. Tracker – the 2D image that should be attached to the 3D model and printed for use in the visualization process of the model. After download the Augment application to the smartphone and scan of the tracker, the corresponding 3D model of roof or building appears as part of the surrounding environment on the mobile device screen (Fig. 4).

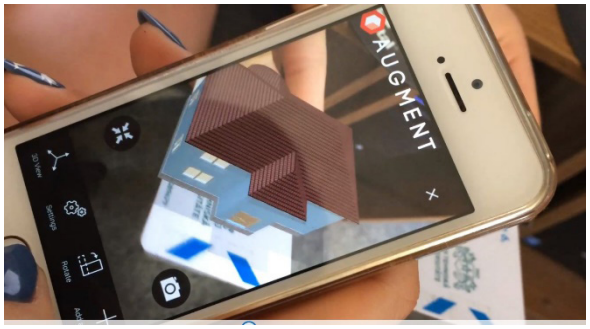


Fig. 4. Visualized AR 3D model of the building.

For visualization of 3D models, students use as tracker the top view of the corresponding object. Images should be saved in .jpg, .bmp, .png or .tga formats (Fig. 5).

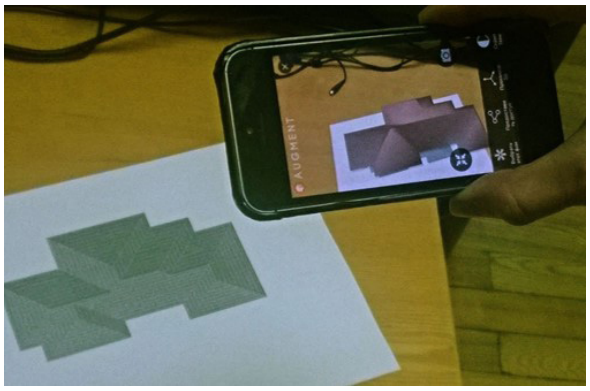


Fig. 5. Visualized AR 3D model of the roof.

According to the opinion of students, the presentation method in which performed task is shown as part of the real environment makes the learning more enjoyable. Work with AR models enhances students' imagination and creativity, helps the transfer of knowledge and skills acquired in the virtual environment to the real world. Augmented reality technologies create a realistic simulation and testing environment by transferring computerized applications to the real world and make the training and teaching process more interactive, influential, powerful and enjoyable for students, provide the students with their own unique knowledge discovery path.

III. Conclusions

In the field of education, trigger can successfully be used to stimulate student motivation to study. In the initial courses it is important to interest students, to create a positive atmosphere during the lesson. The introduction of AR technology shows positive results in the study of graphic disciplines, where students have difficulties.

The use of AR increases the effectiveness of student independent work. The amount of study hours for independent study of the course is increasing, therefore, students must learn individually, and technologies make it possible to facilitate the understanding the topic of the course significantly. In particular, it is effective in studying the geometric part of the course, where abstract thinking is necessary.

AR usage improves the quality of training, which is directly related to the quality of engineering education in general. Tests showed a positive trend in the quality of drawing performed and the spatial skills of students. A real model of an object facilitates the understanding of errors made by students in the graphical exercises.

The use of digital technologies attracts people of the today's generation, which motivates them to learn. It is important to interest students and to prevent many students from being expelled during the first year of study.

The use of AR technologies also motivates teachers to change the kind of teaching towards students and makes lessons more attractive and interesting.

This study can help university instructors, college and vocational school teachers in development of new methodological approaches with support of AR technologies.

References

[1] S. Brauer, P. Siklander., S. Ruhalahti, "Motivation in digital open badge-driven learning in vocational teacher education", Ammattikasvatuksen aikakauskirja, vol. 19(3), pp.7–23. 2017.

[2] L. P. Sun, P.Siklander, H. Ruokamo, "How to trigger students' interest in digital learning environments: A systematic literature review", Seminar.Net, vol. 14(1), pp. 62-84. Retrieved from https://journals.hioa.no/index.php/seminar/article/view/2597.2018. [Accessed February 4, 2019].

[3] S. Kirkley and J. Kirkley, "Creating next generation blended learning environments using mixed reality, video games and simulation", TechTrends, vol. 49(3), pp. 42-89. 2005.

[4] A. Arslan and A. T. Ünal, "Examination of cell phone usage habits and purposes of education faculty students", International Journal of Human Sciences, vol. 10(1), pp. 182-201. 2013.

[5] F. Wang and M. J. Hannafin, "Design-based research and technology-enhanced learning environments", Educational Technology Research and Development, vol. 53(4), pp. 5-23. 2005.

[6] I. Jurane, Engineering Education Improvement Opportunities using Computer Games. Society, Integration, Education, Book Series: Society Integration Education, Vol. I, pp. 100-109. 2013, International Scientific Conference on Society, Integration, and Education Location, May 24-25, 2013, Rezekne, Latvia.

[7] T.J. Branoff, M. Dobelis, "The Relationship between Spatial Visualization Ability and Students' Ability to Model 3D Objects from Engineering Assembly Drawings", Engineering Design Graphics Journal of the ASEE, vol. 76(3), pp. 37-43. 2012.

[8] W.Chen, Historical Oslo on a handheld device – a mobile augmented reality Application. 18th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems - KES2014. Procedia Computer Science, vol. 35, pp. 979 – 985. 2014.

[9] M. da Silva, J. Teixeira, P. Cavalcante, V. Teichrieb, "Perspectives on how to evaluate augmented reality technology tools for education: a systematic review". J Braz. Comput. Soc., vol. 25, no. 3. 2019. Available: Springer Open, https://doi.org/10.1186/s13173-019-0084-8. [Accessed February 4, 2019].

[10] Y. H. Wang, "Exploring the effectiveness of integrating augmented reality-based materials to support writing activities", Computers & Education, vol. 113, pp.162–176. 2017. Abstract

available: https://doi.org/10.1016/j.compedu.2017.04.013. [Accessed February 4, 2019].

[11] M. Dunleavy and C. Dede, "Augmented reality teaching and learning", in Handbook of research on educational communications and technology J. M. Spector, M. D. Merrill, J. Elen. and M. J. Bishop, (forth edition), London: Springer. 2014, pp. 735-745.

[12] M. Kesim, and Y. Ozarslan, "Augmented reality in education: current technologies and the potential for education", Procedia-Social and Behavioral Sciences, vol. 47, pp. 297-302. 2012.

[13] A. D. Serio, M. B. Ibanez, C. D. Kloos, "Impact of an augmented reality system on students motivation for a visual art course", Computers and Education, vol. 32, pp. 586-596. 2013.

[14] J. Martín-Gutiérrez, J.L. Saorín, M. Contero, M. Alcaniz, AR_ Dehaes: An Educational Toolkit Based on Augmented Reality Technology for Learning Engineering Graphics. 10th IEEE International Conference on Advanced Learning Technologies. July 5-7, 2010, Sousse, Tunisia, 2010.

[15] J. Martin-Gutiérrez, R.E. Navarro, M.A. Gonzalez, Mixed Reality for Development of Spatial Skills of First-Year Engineering Students", The Proceedings of the 41st Frontiers in Education Conference, October 12-15, 2011, Rapid City, SD, USA, pp. T1A-1- T1A-6, 2011.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 247-252

Electronic Information and Educational Environment of the University as a Means of Organizing Independent Work of Students

Elena Zhestkova
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Address: K.Marx 36
Arzamas, Russia
ezhestkova@mail.ru

Natalia Fomina
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Address: K.Marx 36
Arzamas, Russia
ekvidi@mail.ru

***Abstract*—In the context of the modernization of Russian higher education, training should be built on a differentiated basis, have a problematic and developing character. The main purpose of higher education is to prepare a graduate who is not just knowledgeable but who knows how to apply his knowledge. In this regard, there is a need for continuous self-education and self-development, in the development of selection skills useful for students of professional development and improvement of information. At the same time the efforts of the teacher should be aimed at creating an environment focused on independence, interactivity, productivity of students' activities, an environment that provides an opportunity to form an individual educational experience of a student moving along their own educational path. To organize effective independent work of students it is advisable to create an integrated system for organizing this work based on the information and educational environment of the university. The article presents the experience of organizing independent work of students of the Arzamas branch of Nizhny NovgorodStateUniversity through the electronic information and educational environment of the university. The article describes the developed electronic courses in the disciplines of the psychological-pedagogical cycle, shows the importance of such areas of independent work of students when using the electronic educational environment of the university as the search for the necessary information; preparation for practical, laboratory, seminar classes; preparation for the survey, colloquium; test preparation; preparation for tests and other types of written test works; preparation for various competitions, contests and more.**

***Keywords*—electronic information-educational environment, higher education, independent work, student.**

I. Introduction

The implementation of state educational standards at all levels of education implies increased requirements for the quality of education. High school is no exception. Here, the problem of improving the quality of the educational process is central. In the modern information world, the quality of education cannot be achieved only by filling the student with more and more knowledge. In the modern world it is impossible, as in the times of Ya.A. Kamensky, to collect all the knowledge in one book and pass it on to the younger generation since the accumulation of information is proceeding at an accelerated pace. A modern student should be able to navigate information flows and use it with conscience [1]. "Knowledge is knowledge only in the case when it is acquire by the thought and not by memory" [2] – this thesis of L.N. Tolstoy fills the student's learning activities with a strategy of independent activity and personal activity. Therefore, it is necessary to find new approaches, means and opportunities to improve the quality of the educational process in higher education.

In the context of the modernization of the Russian school, a competence paradigm has been adopted. The main goal of higher education is to prepare a graduate who not only possesses knowledge, but knows how to dispose of his knowledge [3]. It is necessary to prepare a professional who would be able to choose the most optimal among a multitude of solutions, a professional who is ready for self-education, self-determination and self-development [4], [5].

The need to prepare a university graduate for continuous self-development and self-education makes the independent work of students one of the basic forms of their preparation for successful life in the information society.

Independent work of students in the scientific literature is a multidimensional concept. Various approaches in defining independent work by various authors systematizes and classifies by I.V. Kharitonova. Independent work is defined as independent exercise activity (V.A.Kozakov [6]);work on the instructions of the teacher (R. Mikelson, B. Esipov [6]); type of

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4091
© 2019 Elena Zhestkova, Natalia Fomina.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

learning activity (L.I.Ruvinsky) [6]); teaching method (L.V. Zharkova, A.V. Usov [7]); form of organization of educational activities (T.I. Shamov [8], M.I. Moro [9], I. E. Unt [10]); a means of organizing independent activities (P.I.Pidkasisty, R.A.Lozovskaya [11]) and others. Following definition was most prevalent: "independent work of students is an active, purposeful acquisition by the student of new knowledge for him without direct participation in this process of teachers" emphasizes I.V. Kharitonova.

Increase the role of independent work of students in conducting various types of training involves:

- optimization of teaching methods, introduction of new technologies into the educational process, increasing teacher's productivity, active use of information technologies, allowing the student to master the teaching material at a convenient time for him;
- widespread introduction of computerized testing;
- improvement of the practice of conducting practices and research work of students, since it is these types of work that primarily prepare them for the independent fulfillment of professional tasks;
- modernization of the course and diploma design system which should enhance the role of the student in the selection of material, finding ways to solve research problems.

So, the formation of the future specialist skills and skills of independent work is one of the most important tasks of higher education. Successful solution requires an integrated approach which includes scientific substantiation of the essence and content of independent work, development of an acceptable for the student planning system and organization of independent work in the educational process, development of methods of independent activity during the classroom and extracurricular time, management and self-management at different stages of preparation young specialist.

II. Vital question

Education in higher education should be built on a differentiated basis allowing interest and maximum use of strong trained students while not suppressing those students who are not sufficiently prepared and cannot keep up with the pace of work of their peers. The fundamental principle for this should be an independent choice by students of their level of mastering educational material. The nature of multi-level should be done so that unobtrusively, but constantly it was possible to teach a lower level student to constant improvement. Training should be problematic and developing, it should activate students' cognitive activity. It is necessary for students to form the need for continuous self-education and self-development, to develop selection skills useful for their professional development and improvement of information.

III. Problematics

The transformation knowledge type of education to activity type makes the main system forming factor not the component of obtain knowledge but the method to solve various educational tasks.

Function of teaching must be transformed to function of supporting study. Thelearner's position should change from a passive learning object, a recipient of ready-made learning information,a learning object and educating influences to an active subject of the learning, independently extracting the necessary information and constructing the necessary ways of doing it [12]. The efforts of the teacher should be aimed at creating an environment focused on autonomy, interactivity, productivity of students, the environment providing the possibility of forming an individual educational experience of the student, moving along their own educational trajectory. To organize effective independent work of students it is advisable to create an integrated system for organizing this work based on the information and educational environment of the university [13], [14]. The teacher should actively use the capabilities of the information and educational environment of the university for the organization of all types of independent work of students, to be an active participant in its creation. The degree of use of the capabilities of the information-educational environment depends on the position of the teacher. The teacher is the initiator of engaging in the educational process of electronic resources.

Analysis of the intensity of use by teachers of the capabilities of the electronic information and educational environment of the Lobachevsky State University of Nizhny Novgorod (Arzamas Branch) shows that more than half of teachers do not seek to actively use it. The study of such a passive attitude of higher education teachers allowed us to identify a number of barriers in the use of the possibilities of the information educational environment of the university:

- conservatism of the faculty of universities (especially older teachers) which is manifested in the fact that these teachers are reluctant to master the possibilities of electronic resources and the capabilities of the university's e-learning system;
- necessity to have long timeto create your own electronic courses;
- unwillingness of students, especially the correspondence department, to work in the e-learning system;
- necessity to monitor the implementation of students' tasks in the e-learning system (verification, adjustment of completed tasks requires considerable time which are not taken into account in the teaching load of teachers);
- lack of organized management of the activities of students and teachers when working in the e-learning system by the administration of educational departments of the university, faculties and departments;
- not a well-established system of incentives for teachers who actively use the capabilities of e-learning system.

Therefore, at present, due to the prevailing

contradiction between the need to use the electronic information educational environment of the university for organizing and effectively managing the independent work of students and a number of identified barriers, which together determine the research problem, a range of tasks to be addressed should be defined.

The conservatism of teachers is largely based on the fact that many of them are not familiar with the capabilities of e-learning system, are not sufficiently aware of the fact that e-learning environment makes it possible to put into practice practically all the principles of organizing students' independent work, that is, it is in this sense universal means. The system of interrelated principles of the organization of independent work of students is represented by the following:

1) the motivation and interest of students in improving the level of professional training;
2) meaningfulness, activity and initiative in mastering subjects;
3) mutual respect and trust of students and teachers;
4) regularity and consistency (division of work into logical parts, determining the time for each of them);
5) expediency (the content of independent work should correspond to the main goals and objectives of certain specific subject areas);
6) practical orientation, that is, the orientation of the tasks of independent work on what is necessary in the practical activity of a specialist;
7) optimization of actions to achieve the most significant results, minimization of costs for the implementation of secondary actions;
8) individualization (independent educational work is based on individually-typological (temperament, character, abilities) and personal (features of perception, memory, thinking, will) properties;
9) instrumentalization of actions, that is, the involvement of technical means to search for information, solve typical problems, and assimilate knowledge;
10) availability of information and teaching aids which presupposes a high level of equipment in the material, technical and information base, the possibility of its use by students;
11) assessability (each type of independent work must be timely evaluated);
12) diagnosability (any independent work should contribute to personal, mental and professional development; diagnosis should be carried out with the control carried out by means of diagnosed events).

IV. Methods, organization and research results

The article presents the materials obtained through the theoretical analysis of the scientific and methodological literature and the description of experimental work on the use of the capabilities of e-learning system by teachers of the Lobachevsky State University of Nizhny Novgorod (Arzamas Branch)for the organization of independent work of students of the faculty of preschool and primary education.

The experiment was attended by students of 1 and 2 courses in the amount of 100 full-time and part-time departments for 2 years (2016-2017 academic year, 2017-2018 academic year).

The experiment took place in three stages.The first stage is the ascertaining experiment.The purpose of the ascertaining experiment is to identify the level of formation of skills of independent work of students of the 1st course.

One of the factors determining the effectiveness of independent work is knowledge of how to work, what skills to apply. In connection with the five-component structure of educational activities by N.V. Kuzmina all the skills of independent work can be divided into the following blocks: gnostic, design, construction, organization, communication [15].

Gnostic skills include skills related to the process of learning and mastering.Design skills constitute the skills of goal-setting and design goals of independent activities.Constructive skills are the ability to determine the content, logic, sequence of performance of various types of independent study work.

The organizational unit highlights the ability to implement the goals and plans formulated in the process of designing and constructing educational activities.

Communication skills is the ability of the individual to create and implement communicative relations in the conditions of independent work.

Formation of self-study skills begins in school, but in practice, the level of their formation does not meet the requirements of university education. In high school, the need for independent work skills increases dramatically. Therefore, the first-yearstudent has no such skills.

To determine the level of formation of skills of independent work, students were offered a questionnaire in which it was necessary to evaluate themselves for five blocks of skills.

The gnostic skills block contains seven points. Gnostic skills are the skills that:

- highlight the most important thing in the text, in the task, the ability to read the contents of the monitor screen, to see the screen entirely, and not fragmentary, to be able to focus attention on the necessary areas;
- highlight the relationship between the studied phenomena;
- use reference books;
- involve additional sources in the study of the material; formulate a hypothesis;
- outline ways to test it; apply knowledge in changed conditions;
- critically perceive the material, find contradictions;
- use quick reading techniques;
- make analysis;
- make synthesis, generalizations, comparisons,

be able to concretize, give examples of common phenomena;
- plan the pace and rhythm of independent learning activities; be able to break work into stages and plan each of them;
- select the desired method or method for solving the problem;
- plan independent work for a long term;
- plan independent work for the near future and now.

The constructive skills allow to:

- briefly and concisely jot down the material;
- draw up annotations to the texts;
- select and memorize vivid quotes; draw up a plan for the upcoming message, abstract;
- present the studied material in charts, tables, making clusters;
- draw conclusions and generalizations based on the studied literature.

Organizational skills are aimed to:

- organize time, not to waste it on unimportant and unnecessary things;
- organize independent activities in order to acquire the skills and abilities necessary for future profession;
- organize workplace, the means for independent work; organize the information accumulated during the assignment in such a way that it can be used upon completion of work.

Communication skills are the skills that enable to:

- express, highlight own attitude to the facts;
- formulate value judgments and argue own opinion;
- establish contacts with the teacher, fellow students on educational issues.

For each item, students were asked to rate themselves on a scale: high level, medium level, low level.The results of the survey are presented in Fig. 1.

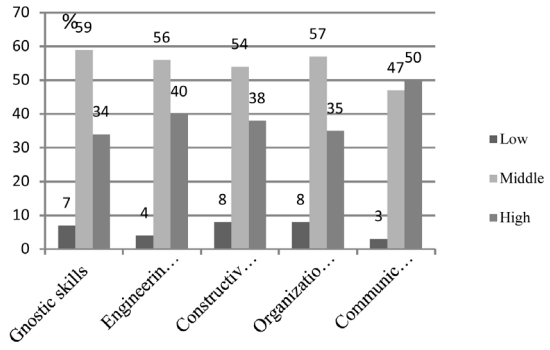


Fig. 1. The level of formation skills of independent work1st year students,% of respondents

The diagnostic results presented in the diagram indicate that in the first year all skills, with the exception of communicative ones, are formed at an average level. It can be assumed that the relatively high degree of development of communicative skills is related to the fact that at all levels of school education a sufficient amount of manpower and resources was devoted to the development of students' communication skills. However, gnostic, design, constructive and organizational skills were not sufficiently formed for the productive independent work of students.

The result of the ascertaining stage of the experiment is the conclusion that the formation of skills for independent work should be continued in higher education, it largely depends on the organizing activity of the teacher, on the methodological support of independent work and maximizing the potential of the electronic information educational system of the university.

The second stage of experimental work assumed that, starting from the first year, students of experimental groups were centrally registered in the electronic information system of the university. Each of them received an individual login and password to be able to log in to this system.

For the organization of independent work with students, electronic courses have been developed in the disciplines of psychological and pedagogical cycle ("Introduction to pedagogical activity", "Methodology and methods of psychological and pedagogical research", "Pedagogy (two parts – Theoretical pedagogy and Practical pedagogy).

According to the discipline "Pedagogy" (theoretical and practical parts), the structure of the courses consists of several sections. Each section contains a plan for studying a separate topic of the course supporting theses on the fundamentals of the theory of the studied question. Each topic is presented on several pages of the electronic manual, that is, students study theoretical material "in portions". After each mini-portion of the theoretical material, students are asked questions and given tests to consolidate and verify their mastery of part of the educational material. The software capabilities of the electronic system allow usingvideo and audio content in addition to the text. Students at a convenient time for them turn to the implementation of tasks on the studied topic. The teacher has the opportunity to remotely check the progress of the student in studying the materials of a section, to give recommendations for its implementation using such forms of communication as electronic correspondence, remote video, as well as audio circulation. Working with an electronic course allows students to vary the course content. The teacher can independently fill in the e-course, supplement it, update its content. The student and the teacher do not lose touch with each other even when the part-time student leaves home after the sessions. Supervision of independent work of students is conducted purposefully, continuously and consistently.

In order to exercise strict control over the activities of students, the teacher sets the time for the performance of a particular task, puts a limitation on the duration of testing.

The fact that the teacher has the ability to quickly verify the test, answer questions, carry out the verification of students' independent work on various creative tasks

(check the essay, abstract, composition, answer recorded in audio and video formats, etc.) can be considered to be positive. In this case, the student promptly receives a response from the teacher. A part-time student has the opportunity to receive qualified assistance. The teacher pre-specifies the time of personal communication with the student, gives the schedule of consultations which is always available to the student. The student learns the content of training coursesat his own pace.The implementation of distance learning, control over the educational activities of students is possible only if certain requirements are met which educational information must meet. Compliance with these requirements contributes to the optimization of the educational process.

When composing e-learning disciplines, the teacher should consider the following requirements:

1. The completeness of information that is supported by past experience, the inclusion of a variety of perspectives on the issue under consideration, compliance with the interests and goals of the student, a focus on practical application, compliance with the profile of education.
2. Relevance of information. The amount of information necessary for the successful organization of vocational training should include not all the informationfrom the sources used, but only the part of it that is relevant to the objectives of the training, the study of a specific issue. Therefore, it is necessary to differentiate meaningfulbasic and additional information. The basic information must be in the reference points of the studied question, placed on the main page since this is where the student starts to study this information. Additional information may be the content of hyperlinks, the contents of which are primary sources, pages of various sites, video clips, audio recordings, etc.
3. Objectivity and accuracy of information. The information that is given for the development must be objective and accurate. This is provided by the presented comparative analysis of various points of view on the studied question and evidence (logical evidence, the results of diagnostic methods, the results of experimental work).
4. Structured information. The multidimensional nature of information coming from various sources makes it difficult to master it, especially under time pressure. Moreover, the structure of each block is also determined by the logic of the studied question. The material is presented in the form of orderly structured blocks.
5. The specificity of information. An example of such a requirement is the non-repeatability of studying basic concepts in various disciplines, a clear distribution among the disciplines of the material being studied (which saves time, covers a greater amount of material in each particular discipline, and makes reasonable connections between different subjects and within each subject).
6. Availability of information. Information can develop the educational and training function when its content is clear to all students. It is achieved by ensuring that every student has access to work in the electronic information environment of a higher education institution, and access to training in e-courses for teachers. In addition, the student should have the opportunity not only to work in a specific e-course, developed by his teacher, but also to have access to the funds of the main electronic libraries of the University "Student Consultant", "Urayt", "Znanium", "Lan". Accessibility is also provided by the organization of a clear conversational communication, execution of the schedule of direct and indirect contact of the student and teacher in an electronic environment.
7. Timeliness and continuity of information. Information should be timely, as any delayed information either becomes useless, or even harm the student, because it will not correspond to real events and harm the learning situation.

The main areas of independent work of students using the electronic educational environment of the university can be:

- search for necessary information;
- preparation for practical, laboratory, seminar classes;
- preparation for the survey, colloquium;
- preparation for testing;
- preparation for tests and other types of written test papers;
- preparation for variousolympiads, competitions;
- performance of home examinations and tasks;
- writing essays, reports of articles;
- drawing up annotations on text documents and watched video and audio recordings;
- implementation of various projects, research works;
- preparation for tests;
- preparation for course and final exams.

The positive fact is that all the information, the products of the student's learning activities remain in electronic form for many years on the student's personal page. A student and graduate can refer to it at any time and clarify forgotten information taking advantage of their work.

Active use of the electronic information system of the university by students contributes to the development of their skills in independent work. This conclusion is confirmed by the results of re-diagnosis (the third stage of the experimental work is the control experiment). The same students participated in the second survey but a year later (2nd year students), that is, in the 2017-2018 academic year (Fig. 2).

The positive dynamics of the level of formation of the skills of independent work of students are presented in Fig. 3.

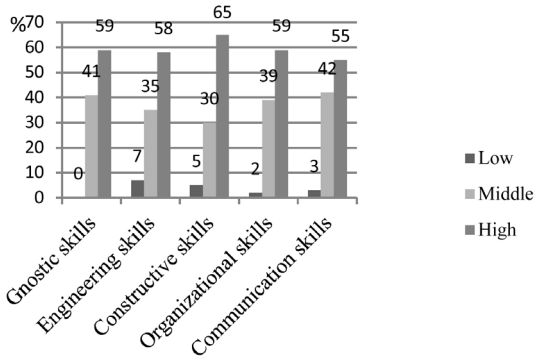


Fig. 2. The level of formation skills of independent work 2nd year students,% of respondents

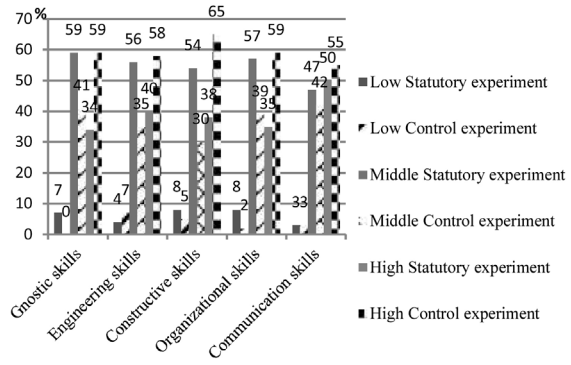


Fig. 3.Comparative characteristics of the results of the ascertaining and control stages of experimental work to identify the dynamics of the level of formation of skills of independent work of students

V. Conclusions

A qualitative analysis of the results of the ascertaining and control experiments allows us to draw the following conclusions. The greatest positive dynamics can be traced by the skills of the gnostic, design, constructive and organizational groups. The least change is experienced by the skills of the communicative group. Students have no difficulty in communicating with fellow students and with teachers, however, they find it difficult to argue their answers in oral and written forms. Gnostic skills in second-year students are more pronounced at a high level. Students forming individual techniques and ways of studying the material, they are able to identify the main and peripheral concepts of the topic. Achieve engineering, constructive and organizational skills a high level of development in more than half of the students. Students are much better at planning the pace of independent work, skillfully managing their time resources, rationally distributing their strength in the study of complex material and already familiar issues [16].

Analysis of the results of diagnostics confirmed the assumption about the positive impact of the active use of the electronic information-educational environment of the university on the development of independent work skills among students.

References

[1] E.I. Mashbits, Psychological and pedagogical problems of learning computerization. Moscow: Pedagogy, 1988, 191 p.

[2] L.N. Tolstoy, Education in freedom: fav. Ped Art. Moscow: Izdat. House "Karapuz", 2005, pp.124-125.

[3] A.P. Lobanov and N.V. Drozdov, Independent work of students in the context of the competence approach: Organization of independent work of students at the university faculty: materials of the International Scientific and Practical Conference, Minsk, November 16-17, 2006. Minsk: BSU, 2006. pp. 35-37.

[4] I.V. Kharitonov, Influence of strategies of educational work of a university on students' self-organization. Human Capital. 2011. № 3 (64), pp. 33 - 38.

[5] I.V. Kharitonov and S.N. Fomina, Organization of independent work as a means of self-organization of students in a student-centered approach to learning. Social Policy and Sociology, 2010, № 7 (61), pp.70 - 77.

[6] I.V. Kharitonov, The main directions and psychological and pedagogical foundations of the organization of independent work of students: monograph. Saransk: Publishing House Mordov. University, 2012. pp. 9-10.

[7] L.V. Zharkova, Organization of independent educational and cognitive activity of students: studies. Manual. Leningrad State. Ped, in-t, 1986, pp.67-68.

[8] T.I. Shamova, Activization of schoolchildren's teaching. Moscow: Pedagogy, 1982. pp.84-85.

[9] M.I. Moro, Independent work of students in the lessons of arithmetic in primary school. Moscow: Publishing house of the APS of the RSFSR, 1963, pp. 14-15.

[10] I.E. Unt, Individualization and differentiation of training. Moscow: Pedagogy, 1990, pp.135-136.

[11] P.I. Pidkasisty, Requirements for students in universities. Pedagogy, 2005, №3, pp. 47-52.

[12] E.O. Ivanova and I.M. Osmolovskaya, Learning Theory in the Information Society. Moscow: Prsveshchenie, 2011.

[13] G.M. Kiselev and R.V. Bochkova, Information technologies in pedagogical education. Moscow: Publishing and Trading Corporation "Dashkov and K", 2013.

[14] I.V. Robert and S. V. Panyukova, A. A. Kuznetsov and A. Yu. Kravtsova, Information and communication technologies in education: teaching aid; by ed. I.V. Robert. Moscow: Drofa, 2008, 312 p.

[15] N.V. Kuzmina, Professional activities of a teacher and a master of vocational school vocational training. Moscow: Higher school, 1989.

[16] E. Zhestkova, Subject information and educational environment as means of formation of information and communication competence of future professionals: Environment. Technology. Resources: Proceedings of the 11th International Scientific and Practical Conference June 15-17, 2017. Volume II. Rezekne: Rezekne Academy of Technologies, 2017. pp. 180-186.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 253-258

Social Networks as a Component of Educational Area of the High School

Elena Zhestkova
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
ezhestkova@mail.ru

Dmitriy Gusev
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
dimigus@rambler.ru

Natalja Kudakova
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
kkudakova@mail.ru

Elvira Maklaeva
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
mak_ela@mail.ru

Svetlana Fedorova
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
sveta_fedorov@mail.ru

Liudmila Filippova
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
lyda0968@bk.ru

Abstract – **Currently, there is an active process of expanding the educational space, including the higher education of the Russian Federation, at the expense of informatization. Informatization of educational area of higher education is impossible without using web-technologies and, such as social networks.**

The authors of the article suggest social networks to be understood as a working version of an "interactive multi-user website that implements a network social structure consisting of a group of nodes — social objects (groups of people, communities) and links between them (social relationships), on the basis of which participants can establish relationship with each other.

Currently, there are several main functions of social networks: educational; adaptive (acts as a resource of adaptation); compensatory (replacing institutional mechanisms of adaptation); informational (supports communication between authors of social interaction); transit (allows an individual to make the transition along the social ladder); coordination; social support functions (strengthen communication within and outside the network) and the function of a sociocultural marker.

Social networks, speaking as a special social space of the Internet, have become the sphere in which traditional forms of socialization and social relations are transformed, and communication as a type of leisure activity becomes possible not in the traditional form of direct live communication, but acquires the features of simple communication.

The authors state that at present using social networks in the educational area of the higher school of the Russian Federation is minimal. The network educational community on the basis of a social network - a virtual educational environment - is necessary, first of all, for students who have difficulties in communicating directly or need additional knowledge and skills that an educational institution cannot provide. From the point of view of education, social networks can be: freely available (non-specialized networks for which professional communities are not paramount and purely professional communities of practice) and in a corporate format (free-access networks; not specialized ("general profile" network)).

The advent of Web 2.0 has expanded the possibilities of using social networks in education, has changed the attitude to the Internet as a whole, and teachers have begun to more actively use the Internet services for educational and educational purposes, in extracurricular activities and creative activities.

It has already been experimentally proven that network communities can serve as pedagogical practices for development: co-thinking, tolerance, mastering decentralized models, critical thinking.

Keywords – High school, informatization, social networks, web-technologies.

I. Introduction

Currently, there is an active process of expanding education area, including higher education of the Russian Federation, due to informatization. "The term" informatization has become generally accepted and spread in all spheres of society. Under it today they understand both politics, and processes, and technologies, and mechanisms [1].

Informatization of educational area of the higher education without using of web technologies and, such as social network.

In modern native literature, despite of the presence of publications, which reveal the essence and describe the content of already used information educational technologies, works dedicated to social networks as an information educational web-technology, obviously not enough. Therefore, the purpose of this article is to disclose essence of the concept and phenomenon of

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4092
© 2019 Elena Zhestkova, Dmitriy Gusev, Natalja Kudakova, Elvira Maklaeva, Svetlana Fedorova, Liudmila Filippova.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

social networks and their using in educational area. Its main tasks were: identifying keeping the concepts of "networks" and "social networks", studying of social networks as educational technology and resource for development (personality, group, vocational school, etc.), studying of the dynamics of social networks and their using in education [2], [3].

The works on the analysis and popularization of concrete experience in the using of social networks in education have already appeared in literature. For example, in the work "The development of social networks and their integration into the system of education in Russia", the authors consider the genesis, development and modern various interpretations of the concept of "social network", describe the most popular social networks of Russia (Vkontakte, Odnoklassniki, Facebook, Twitter) and the problems of their using in educational procession. Summarizing several definitions, the authors propose to understand that social network is a working version of "interactive multiplayer website that implements a network social structure consisting of a group nodes - social objects (groups of people, communities) and links between them (social relations), on the basis of which participants can establish relationship with each other [4, p. 475].

According to A.S. Duzhnikova, the site may be called a social network when it is possible to create individual profiles (name, status, interests, etc.); user interaction (through browsing each other's profiles, internal mail, comments, etc.); achieving joint goals through cooperation (for example, finding new friends, group blogging, etc.); sharing of resources (for example, information); meet the needs through the accumulation of resources [5, p. 240].

Noting the author's sociological approach and considering it incomplete we consider the functionality of social networks as a web-technology; and such distinctive features as "the creation of personal products by a file in which you often need to provide real personal data and information about yourself (place of study and work, hobbies, life principles, etc.); providing a full range of opportunities for the exchange of information (sharing photos, videos, placing text records (in the blogs or micro blogs), the organization of thematic communities, the exchange of personal messaging, etc.); ability to set and maintain a list of others users (for example, friendship, kinship, business and work relationships, etc.).

Currently, there are several basic functions of social networks: educational; adaptive (acts as a resource of adaptation); compensatory (replacing institutional mechanisms of adaptation); information (support communication between the authors of the social actions); transit (allows an individual to make the transition through social ladder); coordination; social support functions (strengthened connections inside and outside the network) and the function of sociocultural marker. Social networks, speaking as a special social space of the Internet that became the sphere in which the traditional forms of socialization and social relations, and communication as a type of leisure activity it becomes possible for a person not in the traditional form of direct living communication, but acquires the features of simple communication: in a network a person can to be what he wants to appear, but he is not. According to analytical company Strategy Analytics, in the next five years, various about 1 billion people will use social sites. Currently using social networks in the educational area of higher education of the Russian Federation is a minimum. Apparently, network educational resource will unite teachers, students, programmers, methodologists. It is a complex and multifaceted task. Network social networking community - virtual education environment is necessary, first of all, for students who have difficulties communication in direct communication or in need of additional knowledge and skills that the school cannot provide.

Actualize the problem of using social networks in education area of the higher school of the Russian Federation allows running experience of their application [6], [7].

In particular, it shows that social networks have functional scrap, which allows you to share information quickly, discuss important problems, create collaborative learning content. Using social networks, they master new means and ways of communicating with other people all over the world, recognize each other's interests, learn to effectively search and analysis of information [8]. Thanks to the experience of using social networks for solving educational and research tasks, changes convert students into a resource as solely entertaining understanding of its powerful uses in professional activities.

The possibility of interaction between students and teachers in the network in a convenient time ensures the continuity of the educational process and the possibility of detail-planning of educational and research work of students (tasks and consultations every day). It allows students who miss lessons, take part online. It is possible for teachers to conduct classroom lectures in interactive mode [9]. A virtual classroom created in a social network can be accessible to students and parents everywhere via the mobile Internet. You can show uploaded videos, discussions.

The presentations that begin in classroom can be continued in a social setting. It allows students to spend more time in active learning through discussion. When students create their social profile they indicate different details about their personality: views, interests, favorite music, movies and books, favorite quotes, etc. Teachers have an opportunity to learn more about the student's personality, his individual features [10].

II. General Regulations

With the emergence of a large number of new social services and networks, teachers and students have wide opportunities to use them in the educational process. These web services and networks are gradually becoming the social and information environment within which you can solve many of the pedagogical tasks of new

educational standards that require using of radically new methods and forms of education.

As is known, the interest in social networks among young people in modern society is becoming one of the main features of the 21st century. Most of the life of modern youth goes on-line. That is why using the Internet and social networks so popular among young people is becoming a very important tool for increasing the efficiency of the educational process, a way to create an interactive educational environment and a reliable means of communication between teachers and students. In this regard, an active search is being carried out for the possibilities of applying social networks in the teaching and educational area of the university, taking into account the needs of modern student youth [11].

According to research, the most relevant social networks in Russia are Odnoklassniki, VKontakte, Moi Mir and others. Moreover, VKontakte is the most popular social resource for the student community, and it can be considered as an educational resource in organizing educational activities of students [12].

The communication of the teacher with the students with the help of this interactive resource makes them equal participants in the communication process, the communication process itself becomes more interesting, operational and efficient, and its goal is clearer and more realistic. In addition, as practice shows, the fact that the teacher is an active and experienced user of this interactive resource increases interest in him as an individual and teacher, encourages communicating and cooperating with him and, as a result, significantly affects the psychological climate of the educational process and its effectiveness [13].

In organizational terms, using of social networks in the educational process of a university creates opportunities for:

1. Organizations of individual, collective and group work of students. The combination of group and individual forms of work contributes to a better understanding and assimilation of educational material. The advantage of using social networks is that there is the possibility of a collective assessment of the results and processes of work. An indisputable fact is that with the help of social networks you can build a student's individual educational route and use it as a space for organizing students' independent work and for distance learning.

2. Organization of educational and extracurricular activities. The possibility of holding various contests, presentations, videos using social services and networks can help the teacher in his work, if it is used as information medium in which information about various events will be placed.

3. Organization of project activities of students. Social networks can become a platform for creating individual and network projects, which will allow participants to implement their projects, gain knowledge from practitioners, find like-minded people and the necessary resources, and gain motivation for further self-development.

The ability to attach text documents, images, video and audio recordings to VKontakte messages is another important advantage. Working with students in a created group allows you to centrally post information in the news feed, upload documents, video, audio, photographic materials, where students can watch video lectures, subject films, video tutorials on various subjects of primary school, listen to lectures by leading teachers and psychologists, as well as music that can be used for educational purposes [14], [15].

Lobacevsky State University of Nizhny Novgorod (Arzamas Branch) has accumulated experience of using the VKontakte social network in order to organize students' independent work when studying methodological disciplines [16]. This work is used as a supplement to classroom studies in the study of pedagogical and methodological disciplines, such as "Universal educational activities and technologies of their formation", "Theory and technologies for the development of mathematical representations of children", "Theory and technologies of teaching mathematics", etc.

Since the first training session, the orientation of the student audience to interaction in social networks begins. The teacher tells how to find his page on the social network and how the group will work together. It is explained that this group will systematically place the necessary educational material - lecture materials, tasks for performing extracurricular independent work, etc. In addition, a specific time is specified in advance that the teacher is willing to devote to communicating with students and answering their questions. Thus, this group is a social-educational environment that allows a group of people to communicate, in our case, students and teachers united by a common interest [17], [18], [19].

Interactive communication of the teacher with students and students among themselves is carried out in the following areas:

- placement of lecture materials in the social network, as well as text, video and audio materials aimed at expanding and deepening the knowledge gained during the classroom (for example, video tutorials posted at https://www.youtube.com, https: / /videouroki.net);

- implementing of tasks for independent work (search and processing of information, preparation of project tasks, messages, reports, discussion and correction of which are carried out quickly and efficiently);

- editing by the teacher of materials prepared by students, discussing them online, both with the teacher and among themselves, which significantly reduces the preparation time;

- individual tasks for students with the ability to monitor their implementation and additional counseling during extra-curricular time, which, respectively, significantly saves classroom time;

- joint work of students on creative projects with the

possibility of getting online advice from the teacher;

- the ability to control the self-study activities of students who, due to illness or for some other reason, do not attend classes, which allows them to prevent a significant lag in students from the curriculum and to flexibly regulate the course of the educational process;

- application of the presented material in the group allows all participants to independently or jointly create online learning content (glossaries, articles, discussions, multimedia libraries, etc.),

- solving organizational issues, etc. [20], [21], [22].

The experience of using such a group makes it possible to highlight the following benefits: students, communicating in an informal setting, have the opportunity to gain new knowledge, leading to a more effective development of the discipline; the student and the teacher, interacting in the social network, behave more freely, which allows the student to ask questions about the discipline without fear of critical assessment from the teacher and classmates; the student has the ability to communicate online not only with the teacher, but also with his classmates, which makes it possible to organize conferences and dialogues, especially before the session; the teacher for the student becomes a member of the social network, that is, the interaction at the vertical level is replaced by the interaction at the horizontal level and this inspires confidence on the part of the student and improves the process of mastering the material; the teacher has the opportunity to quickly inform students about upcoming plans [23].

Thus, using social networks in pedagogical activity ensures the continuity of the educational process, stimulates students' independent cognitive activity, and contributes to the improvement of the skills of comprehensive assessment and comparison of the information received; ensures the continuity of the educational process; enables students to build an individual educational trajectory.

III. Results and discussion

The scientific laboratory of didactics of the rural school of Lobachevsky State University of Nizhny Novgorod (Arzamas Branch) is working to identify effective ways of organizing educational activities with using ICT by future and current teachers of rural schools. One of the innovative scientific projects of the laboratory is a model of network interaction of teachers of rural schools in the implementation of the pedagogical potential of folk applied creativity, aimed at studying the possibilities of using ICT in the implementation of artistic and aesthetic education and spiritual and moral development of rural students by means of folk applied creativity [24], [25].

The project is developed on the basis of the social network "creative teachers Network" (Innovative Teachers Network) – an Internet portal for communities of teachers working in secondary education and using or planning to use ICT in the learning process. The largest teacher Internet project in Russia the educational Internet is the resource of Federal importance, it is intended for network interaction of teachers in order to implement new technologies in the organization of the educational process [26], [27].

The organization of work on the network communication of teachers can be presented in the form of a Community of teachers of rural schools "from folk crafts – to the moral ideals of cultural heritage", which will be another web site of the network of creative teachers. Its activities will focus on:

- the development and support of new technologies in the organization of the educational process of rural schools;

- exchange of advanced pedagogical experience;

- the preservation and development of Russian folk culture;

- creating a database of author's materials prepared on the basis of ICT for the organization of work for the student and the teacher of rural school on arts and crafts.

A holistic view of the structure and sections of the main web page of the community of teachers of rural schools "from folk crafts-to the moral ideals of cultural heritage" is given in the form of a model in Fig. 1.

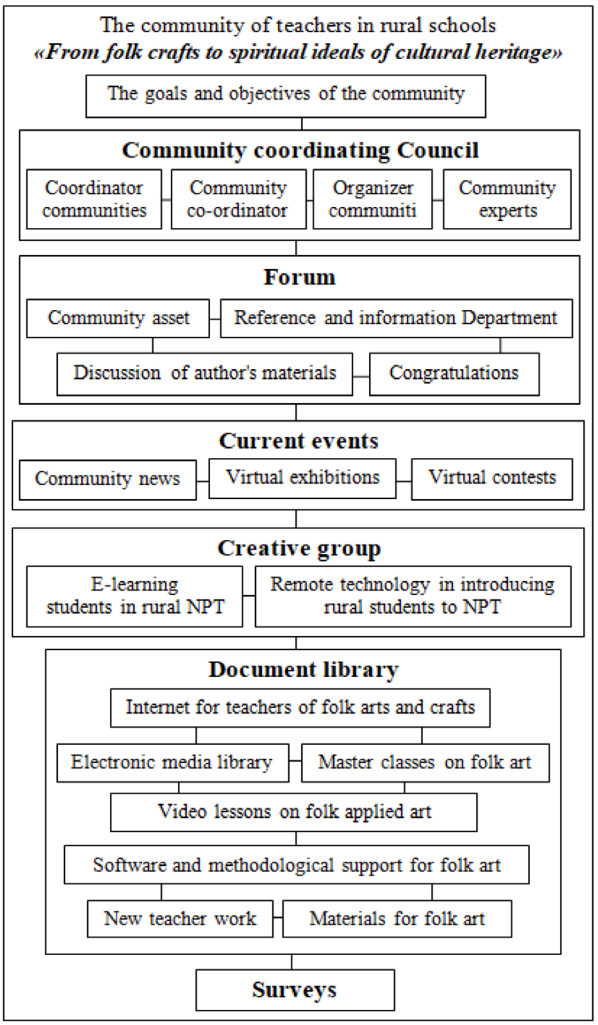


Fig. 1. Model of network interaction of teachers of rural schools in realization of pedagogical potential of national applied creativity

(NPT)

IV. Conclusion

First of all, the clarity of the ideology and interface of social networks of the majority of the Internet audience saves time by passing the students' adaptation to the new communicative space. Placing information in a social network a person not only indicates the range of his professional interests, but also informing additional information, non-verbally invites interested people to his surroundings.

The network educational community based on the social network is a virtual educational environment that allows students not only to carry out vertical communication, but also to gain knowledge, share them, draw logical conclusions, share their reasoning with others.

Networking communities allow students and teachers to form groups for more in-depth study of the material, create models together, replicate ideas, and develop joint thinking. Proper using a social network by the teacher allows you to intuitively feel the audience being trained and select a suitable training platform and tools for it.

Social networks make it possible not only to transfer information in a mobile way and exchange data, but also to conduct various kinds of sociological surveys that provide the basis for monitoring issues of any social strata.

The possibility of constant interaction of students and teachers in the network at a convenient time for them ensures the continuity of the educational process, there is an opportunity for a more detailed organization of work individually with each of the students. In addition, discussions, dialogues, started during the classroom, can be continued in the social network, which allows students to spend more time in the process of discussing educational issues, which provides a more thorough development of the material and the active position of the student in the learning process. Information support of the educational site in the social network allows students, who have missed a lesson not to drop out of the educational process, take part in discussions and complete tasks from home.

Of course, social networks are not the main means of network education, but their ability to solve educational problems today is underestimated by the professional community.

References

[1] S.A. Atroshchenko and S.V. Napalkov, The use of thematic educational Web-quest in the development of cognitive independence of students. Web-technologies in the educational space: problems, approaches, prospects: a collection of articles of the participants of the International Scientific and Practical Conference; Arzamas branch of UNN. N. Novgorod: LLC «Raster-NN», 2015.

[2] E.A. Zhestkova, Subject information and educational environment formation of information and communication competence of future professionals: Environment. Technology. Resources: Proceedings of the 11th International Scientific and Practical Conference, june 15-17, 2017, vol. II. Rezekne: Rezekne Academy of Technologies, 2017, pp. 180-186.

[3] S.V. Fedorova, and E.V. Maklaeva, "Ways of organizing the reflective activity of younger schoolchildren in the implementation of educational projects," Problems of modern pedagogical education. Ser.: Pedagogy and Psychology, vol. 57, Part 1, pp. 239-245, 2017.

[4] G.A. Vorobiev, Technology web quests in the teaching of foreign languages and cultures: Theoretical and experimental linguodidactics: Coll. scientific articles. Pyatigorsk: PSLU, 2003.

[5] E.N. Voronova, "Organization of independent educational activities of students as a factor in improving the quality of the pedagogical process," Language and world of the studied language, vol. 1, pp. 125-127, 2011.

[6] O.G. Miller, "Web quest in teaching students," Proceedings of the Southern Federal University. Technical science, vol. 111, No. 10, 2010. [Access mode] Available: http://cyberleninka/article/n/veb-kvest-v-obuchenii-studentov#ixzz3N03bptTr..

[7] E.S. Poleshchuk, "Web quest as a way of organizing students' search and research activities in teaching a foreign language," Personality, family and society: issues of pedagogy and psychology, vol. 44, pp. 34-37, 2014.

[8] E.R. Dikikh, On the use of social networks in education: Personality, family and society: issues of pedagogy and psychology: Sat. Art. on mater. XVI Intern. scientific-practical conf. Part I, Novosibirsk: SibAK, 2012.

[9] O.A. Klimenko, Social networks as a means of learning and interaction of participants in the educational process: Theory and practice of education in the modern world: materials of the Intern. scientific conf. (St. Petersburg, February 2012), SPb.: Renome, 2012. pp. 405-407, Available: https://moluch.ru/conf/ped/archive/21/1799/ [Access April, 3, 2019].

[10] A.A. Dragunova, "Educational Internet resources as a means of forming professional foreign language communicative competence of students," Yaroslavl Pedagogical Journal, vol. 1, pp. 163-165, 2013.

[11] L.V. Rusina, The modern science: 2013, Budapest. [Access mode] Available: http://scaspee.com/contacts.html.

[12] B. Dodge, Some thoughts about Web Quests, 1995, [Online] Available: [Access mode] http://webquest.sdsu.edu/about_webquests. html

[13] T. March, "Are we there yet? A parable on the educational effectiveness of technology," Multimedia Schools Magazine, vol. 7(3), pp. 54-55, 2000. [Access mode]. Available: www.infotoday.com/MMSchools/may00/march.htm.

[14] E. Zhestkova, E. Minaeva, N. Ivanova, O. Kolesova, and N. Lapin, Teachers from different countries: Revista ESPACIOS, vol. 38(25), 2017. Page 8. [Access mode] Available: http://www.revistaespacios.com/a17v38n25/17382508.html.

[15] E. Zhestkova, E. Safekiova, N. Ivanova, T. Luchina, and E. Minaeva, "Specifics of Educational Activity Anti-motivation in Future Teachers Subject to the Training Period International," Review of Management and Marketing, vol. 3, pp. 265-259, 2016.

[16] S.V. Fedorova, and E.V. Maklaeva, "Technological aspects of the development of critical thinking of students in teaching mathematics," Modern problems of science and education, vol. 4, 2015. Available: http://science-education.ru/ru/article/view?id=21125 [Access May, 3, 2019].

[17] E.A. Zhestkova, Problematic approach to the study of literary disciplines in a pedagogical university: Society. Integration. Education. Proceedings of the International Scientific Conference, may 27-28, 2016, Rezekne: Rezeknes Academy of Technologies, 2016. pp. 286-299.

[18] E.V. Maklaeva, and S.V. Fedorova, "Test technologies in teaching students of pedagogical universities," Elementary School, vol. 12, pp. 11-15, 2016.

[19] L.V. Filippova, "Project technologies in teaching of philological disciplines," Volga Scientific Herald Arzamas Pedagogical Readings, vol. 8(36), Part 2, Izhevsk, pp. 44-48, 2014.

[20] E. Zhestkova, E. Gubanichina, S. Oparina, V. Sidorskaya, and D. Gusev, Interactive technologies as a means of forming general cultural and professional competencies among students of a pedagogical university: Society. Integration. Education: Proceedings of the International Scientific Conference, vol. I, may 26-27, 2017, Rezekne: Rezeknes Academy of Technologies, 2017, pp. 454-469.

[21] N.S. Kudakova, Media education of younger schoolchildren:

Media and information literacy of a modern teacher: at 3 pm Part 1. Materials of the All-Russian Scientific and Practical Conference. Orenburg, Orenburg State Pedagogical University, october 30-31, 2018, Orenburg: Publishing house "Orenburg book", 2018, pp. 167-171.

[22] N. Kudakova, E. Gubanihina, E. Klyueva, T. Naumova, and E. Zhestkova, The pedagogical potential of information and communication technologies in the formation of ethnic and cultural awareness of younger schoolchildren: Vide. Tehnologija. Resursi Environment, Technology, Resources 11. Cep. «Environment. Technology. Resources - Proceedings of the 11th International Scientific and Practical Conference», 2017. pp. 36-41.

[23] L.V. Filippova, Using active teaching methods in the teaching of the course "Theory and technology of speech development of children," Modern high technologies, vol. 12-3, pp. 561-564, 2015.

[24] E.M. Shulgina and M.A. Bovtenko, "The didactic potential of the web quest technology in the formation of foreign language communicative competence of students of non-linguistic tourist departments," Language and Culture, vol. 1, pp. 76-84, 2013.

[25] J.G. Laborda, "Using Educational Technology & Society", vol.12 (1), pp. 258-270, 2009.

[26] D. Gusev, E. Zhestkova, N. Kudakova, L. Filippova, and F. Podshednaya, Pedagogical potential of folk-applied creativity by means of ICT in the process of formation of educational and cognitive interest of younger students: Society. Integration. Education. Proceedings of the International Scientific Conference, May 26-27, 2017, Rezekne: Rezeknes Academy of Technologies, 2017, vol. II, pp. 93-108.

[27] D. Gusev, S. Aksenov. R. Arifulina, T. Belyaeva, E. Ilaltdinova, and I. Lebedeva, "In the search of national ideal of russian pedagogue," Journal of Entrepreneurship Education, vol. 3, 2017.

Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 12th International Scientific and Practical Conference. Volume II, 259-263

Web-Qwest as a Form of Organization of Independent Work of Students

Elena Zhestkova
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
ezhestkova@mail.ru

Dmitriy Gusev
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
dimigus@rambler.ru

Natalja Kudakova
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
kkudakova@mail.ru

Elena Gubanihina
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
gubanihiny@yandex.ru

Elena Klyueva
Faculty of preschool and primary education
Lobachevsky State University of Nizhny Novgorod (Arzamas Branch).
Arzamas, Russia
klyueva.lena@list.ru

Faina Povshednaya
Faculty of preschool and primary education
Nizhny Novgorod State Pedagogical University K. Minina.
Nizhny Novgorod, Russia
povshedfv@yandex.ru

***Abstract*— The article is devoted to the organization of independent work of students enrolled in the direction of teacher training based on web quests. An example of an educational web quest for the 3rd year is the "Inclusive School".**

One of the leading directions of development of the system of higher pedagogical education is the strengthening of the practice-oriented orientation of the study of disciplines, including the educational cycle. This concerns not only an increase in the share of active and interactive methods in teaching, an increased attention to organizing and conducting internships in educational institutions, but also a change in the approach to organizing independent work of bachelor students. Much attention is increasingly given to increasing the level of information and communication competence of future teachers (electronic libraries and educational Internet portals are the main source of educational information for modern students), it is advisable to use information and communication technologies as an organizational and substantive basis for independent work.

This approach to the organization of independent work involves the development of tasks by teachers using Internet technologies, which, depending on the type of task, can be performed online or offline. Tasks for independent work can be directed to an independent search in the network with the subsequent processing of information about a specific problem; on the organization of interaction in the network; create web pages; conduct educational research, social or educational projects. The forms for presenting the results may also vary: text documents (with hyperlinks), databases of multimedia presentations, web pages, web quests.

Particular attention, in our opinion, should be paid to teaching real web quests. A web quest is considered as an activity-oriented didactic model that provides for students' independent search work on the Internet.

As a form of organizing independent work, a web quest has several advantages, including: increased student motivation; activation of individual or group activities; more efficiently use the understanding of the time allotted for independent work of students; the ability to not only integrate different types of assignments and forms of providing results, but also to vary the level of difficulty of the task for independent work, the duration of implementation, the number of performers.

***Keywords*— Educational web-quest, independent work of students, teacher training.**

I. Introduction

In the modern educational situation, one of the qualification requirements for the future teacher is the ability to independently master knowledge in various fields, to engage in self-education.

The most important means of forming students of all types of cognitive activity is the performance of various types and types of self-learning by them. Independent work in the modern educational process is considered as a form of training organization that is able to provide an independent search for the necessary information, creative perception and understanding of the educational material obtained during the classes. It involves the use of various forms of cognitive activity of students in the classroom and after school hours, the development of analytical skills, control skills and planning training time, the development of skills and rational organization of educational work [1]. Thus, independent work is a form of organization of the educational process, stimulating the activity, independence and cognitive interest of students.

Independent work of students using information

Print ISSN 1691-5402
Online ISSN 2256-070X
http://dx.doi.org/10.17770/etr2019vol2.4193
© 2019 Elena Zhestkova, Dmitriy Gusev, Natalja Kudakova, Elena Gubanihina, Elena Klyueva, Faina Povshednaya.
Published by Rezekne Academy of Technologies.
This is an open access articleunder the Creative Commons Attribution 4.0 International License.

technology, Internet technology is an active method of learning, in which students, in accordance with the instructions of the teacher and under his guidance, not only solve the learning task, but also receive the necessary computer skills with various types of information [2], [3]. The activity of this method is determined, first of all, by the goal that is realized by the student during his independent work, it becomes relevant and significant for him, there are motives for the activity:

- the need to expand knowledge, learn new things;
- master the ability to work using a computer;
- the desire to show independence, to perform the task without outside help;
- the need to test your knowledge;
- the possibility of public presentation of the results of activities.

In order to improve students' self-study, constant work is being carried out to select the most effective self-learning organization technologies that, based on the interests, individual characteristics, type of thinking and educational needs of students, will contribute to mastering general and professional skills, including information and communication [4].

Considering that the majority of students spend most of their time in the virtual space, on the Internet, it is advisable to organize the execution of their own tasks in their usual environment. When organizing extracurricular independent work, it is truly independent, since the student himself determines the place and time of inclusion in the educational process [5], [6].

One of these technologies is a web quest. This technology contributes to increasing students' interest in the subjects being studied, their ability to work with information placed on various media, as well as the ability to plan their activities independently [7].

Analysis of scientific sources suggests that the web quest is considered as a form, tool and method of education [8], as a technology, problem task [9] and even as an Internet site, with which students work, solving a particular educational task. According to A. Shestakova, a web quest is an authentic learning technology that allows students to build their own concepts based on the study of the information provided, to create practice-oriented projects [10]. This is explained by the following signs of web quests: unlimited educational resources, the connection of educational material with real life, the active involvement of students in research activities, the advisory role of the teacher in the process of independent work of students.

In our opinion, the web quest is a problem-based learning technology that combines a set of specific methods and tools, including a game and a project method, and involves the active use of information and communication technologies to solve interdisciplinary educational problems [11], [12].

II. Materials and methods

The main research methods used in the work should include: a theoretical analysis of psychological, pedagogical and methodological literature; comparison, synthesis, systematization of the activities of educational institutions in the field of arts and crafts; modeling the content of the educational web-quest, the study and synthesis of the pedagogical experience of teachers of rural educational institutions; observation of the surrounding educational environment in order to design the use of folk arts and crafts.

III. Results and discussion

In the context of the socio-economic reorganization of Russian society, the introduction of innovative high-precision production technologies into various spheres, the system of national education is being modernized. Special changes have occurred in the system of primary and basic general education, firstly, due to the entry into force on September 1, 2013 of the Federal Law "On Education in the Russian Federation", and, secondly, due to the implementation of the new federal state standards general education. One of the strategic objectives of the new Law and the standard can be considered - the continuity of the implementation of educational programs of primary, basic and secondary education using e-learning and distance learning technologies.

Among the basic ideas embodied in the exemplary curriculum for school subjects for the first stage of general education, based on the requirements of the new standard, the rethinking of the essence and meaning of the fine arts in elementary school draws attention. Fine art is a basic subject now, aimed mainly at the formation of the emotional and imaginative, artistic type of thinking, which is a condition for the development of the intellectual and spiritual activity of a growing personality.

As educational practice shows, the results of school monitoring, as well as the opinions of many scientists, one of the favorite activities in elementary school that significantly affect the amplification of child development, is decorative and applied, implemented in art classes in accordance with the GEF using electronic educational and information resources. Despite the fact that electronic resources began to appear in a rural school at the end of the 20th century, the full-fledged organization of the educational process with the use of information technologies began to be carried out quite recently, and not everywhere.

It should be noted that the use of electronic resources, according to researchers in the modern educational process of both urban and rural schools, is actively implemented mainly on the subjects of the natural-mathematical cycle, unfortunately, they are not fully utilized in humanitarian subjects of the aesthetic cycle. Today, in a village school, more than ever, it is important to apply modern electronic resources on aesthetic-oriented subjects, which will allow solving several important problems in the district educational environment [13].

First, schoolchildren's training in fine arts, crafts, music and technology based on electronic and distance learning will make it possible to solve the personnel

problem in rural schools. But there are no many teachers specializing in a specific field of art in these schools and, based on the number of students in schools, they will not be presented for a long time.

Secondly, the rural school is far from the social and cultural opportunities of big cities, in the village the sources of culture were exhausted for a number of reasons, that`s why the great difficulties cause in educational work with children. Today there is one of the few jewels in the village, the secret of which is inherited in every family. This invaluable "phenomenon" is popular applied art created by a simple local resident without a special education for this purpose. Almost in every rural school, house of creativity, club there is a museum where valuable information on applied arts and folk culture is stored, which characterizes the identity and nationality of the settlement. In modern conditions, it is simply necessary to "bring" this information into the school's electronic resource, so that in the future it will be possible to exchange experiences of folk art through the local district network or in any other way, to fully educate children in the village.

With the use of electronic and distance learning resources, it is possible to organize work on folk applied arts with students of rural schools in various forms, going beyond the school organization by connecting to the information network of other schools. So, at the faculty of preschool and primary education of the Arzamas branch of UNN, much attention is paid to preparing future teachers for work in rural areas, one of the directions of such training of students is mastering ICT competence in the field of applied arts. Students learn to independently organize work with students of rural schools in the field of folk arts and crafts through various forms of electronic and distance learning resources [14].

One of such forms of work with rural schoolchildren in applied creativity can be a web-forum, in which the information base accumulates, the students, using this form of work, organize inter-school communication in a particular area of art.

Another form of students' work with students in the field of folk art can be a virtual conference on the exchange of experience, on summarizing the results of work in a specific form of productive activity.

Conducting virtual presentations, exhibitions, competitions and contests are effective forms of work of students with schoolchildren of great educational value.

Another form of work with rural schoolchildren in applied creativity using electronic resources can be a virtual lesson, conducted at a certain time, as a teacher of visual arts at the same time in computer classes at several regional schools.

The organization of out-of-class work in a rural school can also be carried out in the form of virtual excursions on the species diversity of folk art and much more.

Let us dwell in more detail on the innovative project - the web quest "Folk applied arts in a rural school", developed in the scientific laboratory of the didactics of the rural school of the Arzamas branch of UNN, aimed at exploring the possibilities of using electronic and distance learning technologies to improve the organization of the educational process of the rural school.

A web quest is a form of organizing self-study with students; they implement this project as part of writing course works, while undergoing various types of practices, while preparing final qualifying works, as part of studying individual disciplines in preparation for laboratory and practical classes in various disciplines.

Structurally, the web-quest "Decorative and Applied Creativity (DPT) in a Rural School" includes several sections, its visual presentation is shown in Figure 1, the first of which should be considered the section on the regional traditions of decorative and applied creativity in the village. The main content of this section is the text with the names of the types of applied art, which are typical for the villagers, photo of needlework and text relating to the authors of the products.

In the system aspect, the quest includes the following components: targeted, structural, informative and functional.

The quest is intended, above all, to accumulate experience in the use of folk arts and crafts in working with the children of the municipality: district, region, republic, federation. Its purpose is multifaceted.

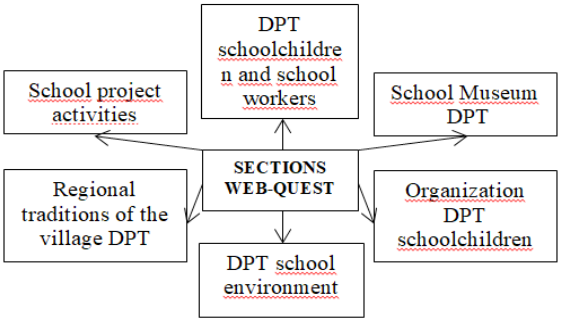


Fig. 1. The structure of the web-quest

In general cultural terms, the quest should contribute to the preservation and further development of the traditions of folk applied arts of rural residents. Peasant art, partly lost by our contemporaries, will receive additional impulses for further improvement of work, and in some cases, the revival of many half-forgotten or non-functioning "centers" of folk culture in rural society.

In socio-economic terms, the quest will contribute to the involvement of villagers in socially useful productive work, bringing personal benefits and material well-being. Having conducted a historical insight into the essence and purpose of folk crafts in the life of a rural resident, it can be safely concluded that this kind of practical activity has not lost its significance to this day since the emergence and development of settlements in the Russian hinterland. Indeed, for many settlements, this primordial occupation was sometimes and is the only source of both material

and spiritual benefits.

In the educational terms, access to the quest materials will allow to learn a new area of artistic and aesthetic knowledge, knowledge about the species diversity of folk art. Applied creativity is a natural occupation of people for almost every village. The teacher in the village, in his historical mission, is the custodian and repeater of local unique traditions. However, in practice, for several reasons, the teacher is not always ready to fulfill this mission. In such cases, the electronic resource of the album will provide indispensable assistance to teachers in the resumption of the school traditions of arts and crafts.

In the didactic aspect, quest resources can be used as visual illustrative material for classes in subjects. On the lessons of the Russian language and literature, relying on visualized electronic images of folk art can be given tasks for schoolchildren: composing a story about a folk product, describing it verbally, analyzing special words such as "matryoshka", "weaving", "potter", etc., on the composition, the drafting of a proposal on any kind of applied arts. At fine art lessons an electronic album can be used as a means of visual-illustrative communication of students with works of applied art, during this process they become acquainted and draw various types of folk art. At the lessons of artistic work, you can use the content of the album as technological charts, explaining the sequence of creating folk products. Even in the lessons of mathematics, images of applied creativity can be used, for example, when schoolchildren become acquainted with arithmetic, when they explain the essence of the operations of addition and subtraction, and when teaching them how to solve various subject problems. In this case, the teacher or schoolchildren speak the content of the task using special terminology, which in itself ensures the involvement of elements of decorative and applied art in the educational process in the classroom.

From the spiritual and moral point of view, the use of quest materials will contribute to the improvement of rural society. Education or the healing of the human soul is not a momentary action with an obvious result, but a long process of systematic and consistent communication of a person with goodness, beauty and creation. The integration of these three components lies in the works of decorative art. Decorative scenes are filled by goodness, artistic images are the embodiment of true human values. Beauty is reflected in the subordination of the form, color and content of the folk product. The idea of creation, in fact, is the leitmotif of all creativity, including decorative and applied arts. A small person surrounded by beauty, goodness and creation, for the most part, will never be a manifestation of ugly, evil and destructive in adult life. That is precisely why the value of folk art in the artistic and aesthetic education of rural schoolchildren is so invaluable.

IV. Conclusion

Constructively, the web-quest "Decorative and Applied Creativity in a Rural School" includes several sections, in the first of which it is advisable to consider a section on regional traditions of decorative and applied creativity in the village. Another important component of the web quest is the section of decorative and applied arts of students and school staff, the content of which is: product photos; texts with the name of the type of creativity, a list of materials necessary for its implementation; with the name and surname of the author.

The section on the organization of arts and crafts of schoolchildren, as a component of the quest, contains information relating to: circles, exhibitions, competitions of arts and crafts creativity.

Another important component of the quest can be a section on the project activities of rural schoolchildren in the field of arts and crafts. The content of the quest should contain information about the school museums of arts and crafts. An essential attribute can be a photo gallery of works from previous years.

Structurally, one should adhere to the fact that the web quest should be the same for all schools in terms of form and content. The amount of information is presented depending on the degree of study in a rural school of each described direction.

The presented web quest can be created on various online platforms. The content of such web quests will be the result of the work of students in the framework of certain types of work at the university. The creation of web-quests will, firstly, contribute to the improvement of the educational activities of the rural school on humanitarian subjects; secondly, the preservation, development and revival of a multinational folk culture; thirdly, generalization and popularization of a rural school in a separate subject area.

References

[1] E.A. Zhestkova, Subject information and educational environment formation of information and communication competence of future professionals: environment. Technology. Resources: Proceedings of the 11th International Scientific and Practical Conference, june 15-17, 2017, vol. II, Rezekne: Rezekne Academy of Technologies, 2017, pp. 180-186.

[2] E. Zhestkova, E. Minaeva, N. Ivanova, O. Kolesova, and N. Lapin, "Teachers from different countries" Revista ESPACIOS, vol. 38(25), 2017, page 8. Available: http://www.revistaespacios.com/a17v38n25/17382508.html

[3] E. Zhestkova, A. Akpayeva, N. Ivanova, T. Luchina, E. Minaeva, "Specifics of Educational Activity Anti-motivation in Future Teachers Subject to the Training Period," International Review of Management and Marketing, vol. 3, pp. 265-259, 2016.

[4] E. Zhestkova, Problematic approach to the study of literary disciplines in a pedagogical university: Society. Integration. Education. Proceedings of the International Scientific Conference, may 27-28, 2016, Rezekne: Rezeknes Academy of Technologies, 2016, pp. 286-299.

[5] E.R. Dikikh, On the use of social networks in education: Personality, family and society: issues of pedagogy and psychology: Sat. Art. on mater. XVI Intern. scientific-practical conf. Part I, Novosibirsk: SibAK, 2012.

[6] O.A. Klimenko, Social networks as a means of learning and interaction of participants in the educational process: Theory and practice of education in the modern world: materials of the Intern. scientific conf, St. Petersburg, February 2012, SPb.: Renome, 2012, pp. 405-407. Available: https://moluch.ru/conf/ped/archive/21/1799/ [Access April, 3, 2019].

[7] N.S. Kudakova, Media education of younger schoolchildren: Media and information literacy of a modern teacher: at 3 pm Part 1.

Materials of the All-Russian Scientific and Practical Conference. Orenburg, Orenburg State Pedagogical University, october 30-31, 2018, Orenburg: Publishing house "Orenburg book", 2018. pp. 167-171.

[8] Y.S. Bykhovsky, Educational web quests [Electronic resource]: Proceedings of the international conference "Information technologies in education. ITO-99". Available: http://ito.edu.ru/1999/III/1/30015.html [Access June, 3, 2016].

[9] M.V. Andreeva, Technologies web quest in the formation of communicative and sociocultural competence: Information and communication technologies in teaching foreign languages: theses of reports of the I International Scientific and Practical Conference, 2004, pp. 20-25.

[10] A.Y. Shestakova, "Web quest as a new educational technology in high school," Scientific and technical statements of St. Petersburg GPU. Humanities and social sciences, vol. 2, pp. 64-68, 2012.

[11] E. Zhestkova, E. Gubanichina, S. Oparina, V. Sidorskaya, and D. Gusev, Interactive technologies as a means of forming general cultural and professional competencies among students of a pedagogical university: Society. Integration. Education. Proceedings of the International Scientific Conference, vol. I, may 26-27, 2017, Rezekne: Rezeknes Academy of Technologies, 2017. pp. 454-469.

[12] N. Kudakova, E. Gubanihina, E. Klyueva, T. Naumova, and E. Zhestkova, The pedagogical potential of information and communication technologies in the formation of ethnic and cultural awareness of younger schoolchildren: Vide. Tehnologija. Resursi - Environment, Technology, Resources 11. Cep. "Environment. Technology. Resources - Proceedings of the 11th International Scientific and Practical Conference", 2017, pp. 36-41.

[13] D. Gusev, S. Aksenov, R. Arifulina, T. Belyaeva, E. Ilaltdinova, I. Lebedeva, "In the search of national ideal of russian pedagogue," Journal of Entrepreneurship Education, T. 20. vol. 3, 2017.

[14] D. Gusev, E. Zhestkova, N. Kudakova, L. Filippova, F. Podshednay, Pedagogical potential of folk-applied creativity by means of ICT in the process of formation of educational and cognitive interest of younger students: Society. Integration. Education. Proceedings of the International Scientific Conference, vol. II, may 26-27, 2017, Rezekne: Rezeknes Academy of Technologies, 2017, pp. 93-108.