

Serhiy Shkarlet
Anatoliy Morozov
Alexander Palagin *Editors*

Mathematical Modeling and Simulation of Systems (MODS'2020)

Selected Papers of 15th International
Scientific-practical Conference,
MODS, 2020 June 29 – July 01,
Chernihiv, Ukraine

Advances in Intelligent Systems and Computing

Volume 1265

Series Editor

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences,
Warsaw, Poland

Advisory Editors

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India

Rafael Bello Perez, Faculty of Mathematics, Physics and Computing,
Universidad Central de Las Villas, Santa Clara, Cuba

Emilio S. Corchado, University of Salamanca, Salamanca, Spain

Hani Hagras, School of Computer Science and Electronic Engineering,
University of Essex, Colchester, UK

László T. Kóczy, Department of Automation, Széchenyi István University,
Gyor, Hungary


Vladik Kreinovich, Department of Computer Science, University of Texas
at El Paso, El Paso, TX, USA

Chin-Teng Lin, Department of Electrical Engineering, National Chiao
Tung University, Hsinchu, Taiwan

Jie Lu, Faculty of Engineering and Information Technology,
University of Technology Sydney, Sydney, NSW, Australia

Patricia Melin, Graduate Program of Computer Science, Tijuana Institute
of Technology, Tijuana, Mexico

Nadia Nedjah, Department of Electronics Engineering, University of Rio de Janeiro,
Rio de Janeiro, Brazil

Ngoc Thanh Nguyen , Faculty of Computer Science and Management,
Wrocław University of Technology, Wrocław, Poland

Jun Wang, Department of Mechanical and Automation Engineering,
The Chinese University of Hong Kong, Shatin, Hong Kong

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within “Advances in Intelligent Systems and Computing” are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink ****

More information about this series at <http://www.springer.com/series/11156>

Serhiy Shkarlet · Anatoliy Morozov ·
Alexander Palagin
Editors

Mathematical Modeling and Simulation of Systems (MODS'2020)

Selected Papers of 15th International
Scientific-practical Conference, MODS, 2020
June 29 – July 01, Chernihiv, Ukraine

 Springer

Editors

Serhiy Shkarlet
Ministerstvo Education and Science
Chernihiv National University
of Technology
Chernihiv, Ukraine

Anatoliy Morozov
Institute of Mathematical Machines
and Systems
Ukraine National Academy of Science
Kyiv, Ukraine

Alexander Palagin
VM Glushkov Institute of Cybernetics
of NAS of Ukraine
Ukraine National Academy of Science
Kyiv, Ukraine

ISSN 2194-5357

ISSN 2194-5365 (electronic)

Advances in Intelligent Systems and Computing

ISBN 978-3-030-58123-7

ISBN 978-3-030-58124-4 (eBook)

<https://doi.org/10.1007/978-3-030-58124-4>

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

In memory of
Professor Vitalii Lytvynov
and
Professor Valentyn Tomashevskiy
Conference founders, Teachers, and always good friends for all of us

The International Scientific-Practical Conference “Mathematical Modeling and Simulation of Systems” (MODS) was formed to bring together outstanding researchers and practitioners in the field of mathematical modeling and simulation from all over the world to share their experience and expertise.

It was established by the Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine (IMMSP of the NASU) in 2006. MODS is an annual international conference organized by IMMSP of the NASU and Chernihiv National University of Technology (CNUT). Since 2018, the State Scientific Research Institute of Armament and Military Equipment Testing and Certification became one of the conference organizers.

The XVth International Scientific-Practical Conference MODS’2020 was held in Chernihiv, Ukraine, during June 29 – July 01, 2020.

MODS’2020 received 77 papers submissions from different countries. All papers went through a rigorous peer review procedure including pre-review and formal review. Based on the review reports, the Program Committee finally selected 35 high-quality papers for presentation at MODS’2020, which are included in “Advances in Intelligent Systems and Computing” series.

This book contains papers devoted to relevant topics including tools and methods of mathematical modeling and simulation in ecology and geographic information systems, manufacturing and project management, information technology, modeling, analysis and tools of safety in distributed information systems, mathematical modeling and simulation of special purpose equipment samples. All of these offer us plenty of valuable information and would be of great benefit to the experience exchange among scientists in modeling and simulation.

The organizers of MODS'2020 made great efforts to ensure the success of this conference. We hereby would like to thank all the members of MODS'2020 Advisory Committee for their guidance and advice, the members of Program Committee and Organizing Committee, and the referees for their effort in reviewing and soliciting the papers, and all authors for their contribution to the formation of a common intellectual environment for solving relevant scientific problems.

Also, we are grateful to Springer-Verlag and Janusz Kacprzyk as the editor responsible for the series "Advances in Intelligent System and Computing" for their great support in publishing these selected papers.

Serhiy Shkarlet
Alexander Palagin
Anatoliy Morozov

Organization

Organizers

Ministry of Education and Science of Ukraine
The National Academy of Sciences of Ukraine
Academy of Technological Sciences of Ukraine
Engineering Academy of Ukraine
State Scientific Research Institute of Armament and Military Equipment Testing
and Certification, Ukraine
Glyndwr University, Wrexham, UK
US Army Research Laboratory, USA
Defence Institute of Tsvetan Lazarov, Bulgaria
Lodz University of Technology, Poland
Riga Technical University, Latvia
Tallinn University of Technology, Estonia
University of Extremadura, Badajoz, Spain
Francysk Scoryna Gomel State University, Belarus
Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
National Technical University of Ukraine “Kyiv Polytechnic Institute”, Ukraine
Yuriy Kondratiyk Poltava National Technical University, Ukraine
The Bohdan Khmelnytsky National University of Cherkasy, Ukraine
Chernihiv National University of Technology, Ukraine

Chairs

Ireneusz Zbiciński	Lodz University of Technology, Poland
Morozov A. A. (Corresponding Member of the NASU)	Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
Onishchenko V. A.	Yuriy Kondratiyk Poltava National Technical University, Ukraine
Enrique Romero-Cadaval	University of Extremadura, Badajoz, Spain

Shkarlet S. N.	Chernihiv National University of Technology, Ukraine
Bashynskiy V. G. (Chief Sc.)	State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Ukraine
Vasiliev A. I.	A. N. Podgorny Institute for Mechanical Engineering Problems of the NASU, Ukraine
Dmitri Vinnikov	Tallinn University of Technology, Estonia
Ilya Galkin	Riga Technical University, Latvia
Demidenko O. M.	Francysk Scoryna Gomel State University, Belarus
John Davies	Glyndwr University, Wrexham, UK

Program Committee

Adamchuk V. V.	The National Academy of Agrarian Sciences of Ukraine, Ukraine
Azarov O. D.	Vinnitsia National Technical University, Ukraine
Bushuev S. D.	Kyiv National University of Construction and Architecture, Ukraine
Dmytriiev V. A. (Chief Sc.)	State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Ukraine
Dorosh M. S.	Chernihiv National University of Technology, Ukraine
Gitis V. B.	Donbass State Engineering Academy, Ukraine
Holub S. V.	Yuriy Fedkovych Chernivtsi National University, Ukraine
Gryshko V. V.	Yuriy Kondratiyk Poltava National Technical University, Ukraine
Janis Zakis	Riga Technical University, Latvia
Kazymyr V. V.	Chernihiv National University of Technology, Ukraine
Kharchenko V. S.	National Technical University of Ukraine “Kyiv Polytechnic Institute”, Ukraine
Klimenko V. P.	Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
Kovalevskiy S. V.	Donbass State Engineering Academy, Ukraine
Kovalets I. V.	Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
Kraskevych V. E.	Kyiv National University of Trade and Economics, Ukraine

Lande D. V.	Institute for Information Recording of NAS of Ukraine, Ukraine
Liakhov O. L.	Yuriy Kondratiyk Poltava National Technical University, Ukraine
Lytvynov V. V.	Chernihiv National University of Technology, Ukraine
Maderych V. S	Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
Mozharovskiy V. V.	Francysk Scoryna Gomel State University, Belarus
Myronenko V. G. (Member of the NAASU)	The National Academy of Agrarian Sciences of Ukraine, Ukraine
Nosovskyi A. V. (Corresponding Member of the NASU)	Institute for Safety Problems of Nuclear Power Plants of the NASU, Ukraine
Sagayda P. I.	Donbass State Engineering Academy
Snyiuk V. E.	Taras Shevchenko National University of Kyiv, Ukraine
Stetsenko I. V.	National Technical University of Ukraine “Igor Sikorsky Kiev Polytechnic Institute,” Ukraine
Tarasenko V. P.	National Technical University of Ukraine “Igor Sikorsky Kiev Polytechnic Institute,” Ukraine
Tarasov O. F.	Donbass State Engineering Academy (DSEA)
Tomashevskiy V. M	National Technical University of Ukraine “Igor Sikorsky Kiev Polytechnic Institute,” Ukraine
Verlan A. F.	National Technical University of Ukraine “Igor Sikorsky Kiev Polytechnic Institute,” Ukraine
Voloshyn O. F.	Taras Shevchenko National University of Kyiv, Ukraine
Zatserkovnyi V. I	Taras Shevchenko National University of Kyiv, Ukraine
Zhelezniak M. Y.	Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine

Steering Chairs

Igor Skiter	Chernihiv National University of Technology, Ukraine
Igor Brovchenko	Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine

Local Organizing Committee

Alla Hrebennyk	Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
Mariia Voitsekhovska	Chernihiv National University of Technology, Ukraine
Valentyn Nekhai	Chernihiv National University of Technology, Ukraine

Sponsors

«S&T Ukraine», Kyiv, Ukraine



Cyber Rapid Analysis for Defense
Awareness of Real-time Situation–CyRADARS
(Project SPS G5286)



About the Book

The development of complex multicomponent systems requires an increase in accuracy, efficiency, and adequacy while reducing the cost of their creation. This book contains works on mathematical and simulation modeling of processes in various domains: ecology and geographic information systems, IT, industry, project management.

The studies presented in the book will be useful to specialists who involved in the development of real events models-analog, management and decision-making models, production models, and software products. Scientists can get acquainted with the latest research in various decisions proposed by leading scholars and identify promising directions for solving complex scientific and practical problems. The chapters of this book contain the contributions presented on the 15th International Scientific-practical Conference, MODS, June 29 – July 01, 2020, Chernihiv, Ukraine.

Contents

Mathematical Modeling and Simulation of Systems in Ecology and Geographic Information Systems	
Use of Nuclear Emergency Response System for Assessment of Transboundary Transfer and Radiological Risks of the Potential Accidental Releases at Khmelnytsky NPP	3
Ivan Kovalets and Oleksandr Romanenko	
Evaluation of Spectral/Grid Nudging Methods for Weather Analysis and Forecasting in Kyiv Region with the Use of WRF Mesoscale Meteorological Model	13
Aleksander Khalchenkov and Ivan Kovalets	
Simulation Modeling for Predicting the Formation of Municipal Waste	24
Viktoriia Khrutba, Tetiana Morozova, Iryna Kotsiuba, and Volodymyr Shamrai	
Modelling Radionuclide Scavenging in the Ocean by a Particle Tracking in Multicomponent Medium with First-Order Reaction Kinetics	36
Igor Brovchenko and Vladimir Maderich	
Information System of Ecological Monitoring “Small Mammals as Bioindicator”	47
Mariia Talakh and Serhii Holub	
Mathematical Modeling of the Summarizing Index for the Biosystems Status as a Tool to Control the Functioning of the Environmental Management System at Business Entities	56
Olena Barabash and Ganna Weigang	

Mathematical Modeling and Simulation of Systems in Manufacturing	
Theoretical Studies of the Rectilinear Motion of the Axis of the Locked Wheel After Braking the Vehicle on the Uphill	69
Serhii Shuklinov, Dmytro Leontiev, Volodymyr Makarov, Victor Verbitskiy, and Anton Hubin	
Analytical Methods for Determining the Static and Dynamic Behavior of Thin-Walled Structures During Machining	82
Sergiy Plankovskyy, Vitalii Myntiuk, Yevgen Tsegelnyk, Sergiy Zadorozhnyi, and Volodymyr Kombarov	
Modelling and Simulation of Metal Construction Stress-Strain Behaviour When Designing Road-Building Machines	92
Oleksandr Rieznikov, Dmytro Kleets, Anton Kholodov, Leonid Khmara, Mykhailo Kholodov, and Natalia Didenko	
Transport Category Airplane Fuselage Master Geometry Parametrical Modeling Method	101
Anton Chumak, Liliia Buival, Andrii Humennyi, Oleksandr Grebenikov, and Dmytro Konyshchuk	
Circular Layout in Thermal Deburring	111
Sergiy Plankovskyy, Olga Shypul, Yevgen Tsegelnyk, Alexander Pankratov, Tatiana Romanova, and Igor Litvinchev	
Mathematical Simulation of a Buck Quasi-resonant Converter as a Part of Autonomous Voltage Inverter	121
Alexey Gorodny, Andrii Dymyrets, Roman Yershov, Yuriy Denisov, and Serhii Boiko	
Tensor Modeling of Spatial Rigidity of Metal-Cutting Machine Spindles	132
Volodymyr Chupryna and Andriy Rudyk	
Mathematical Modeling and Simulation of Systems in Information Technology and Information Security	
Hybrid Convolutional Neuro-Fuzzy Networks for Diagnostics of MRI-Images of Brain Tumors	147
Yuriy Zaychenko and Galib Hamidov	
Method for Automatic Analysis of Compliance of Settlements with Suppliers and Settlements with Customers by Neural Network Model of Forecast	156
Tatiana Neskorodieva and Eugene Fedorov	

Mathematical Modeling of Non-Gaussian Dependent Random Variables by Nonlinear Regression Models Based on the Multivariate Normalizing Transformations 166
 Sergiy Prykhodko and Natalia Prykhodko

Generation of Information Impacts Scenarios in Management Decision Support Systems 175
 Andriy Boychenko and Dmytro Lande

Performance of the Reverse Load Balancer Method in Cluster and Cloud Infrastructures 186
 Oleksandr Khoshaba, Vitalii Lytvynov, Viktor Grechaninov, and Kostiantyn Zaverailo

Building of Directed Weighted Networks of Terms for Decision-Making Support During Information Operations Recognition 197
 Dmytro Lande, Oleh Dmytrenko, Oleh Andriichuk, Vitaliy Tsyganok, and Yaroslava Porplenko

Personalization of E-Learning Process Using AI-Powered Chatbot Integration 209
 John N. Davies, Mariya Verovko, Oleksandr Verovko, and Iryna Solomakha

Methods for Determining the Group Ranking of Alternatives for Incomplete Expert Rankings 217
 Hrygorii Hnatiienko, Nataliia Tmienova, and Alexander Kruglov

Profiling of Clusters in Information Technologies of Intellectual Monitoring 227
 Serhii Holub and Svitlana Kunytska

Proof of Stake for Blockchain Based Distributed Intrusion Detecting System 237
 Ivan Burmaka, Nikolai Stoianov, Vitalii Lytvynov, Mariia Dorosh, and Svitlana Lytvyn

Perfect Reconstruction Condition for Rational Wavelet Transform with Reducible Rational Dilation Factor 248
 Oleg Chertov and Volodymyr Malchykov

Principles of Adaptive Corporate Network Security Management 255
 Vitalii Lytvynov, Alla Hrebennyk, Elena Trunova, Igor Skiter, and Yurii Lysetskyi

Petri-object Simulation Two Level Visual Programming Language 266
 Inna V. Stetsenko and Anton Dyfuchyn

Detection of SQL Injection Attack Using Neural Networks 277
 Oleksandr Hubsnyi, Tetiana Babenko, Larysa Myrutenko,
 and Oleksandr Oksiuk

**Mathematical Modeling and Simulation of Special Purpose
 Equipment Samples**

**Improving Accuracy of Ground Targets Determination by Unmanned
 Aerial Systems** 289
 Maksym Gerashchenko, Oleksandr Isachenko, Sergey Nesterenko,
 and Serhii Rudnichenko

Unmanned Aerial Vehicle Mass Model Peculiarities 299
 Dmitriy Kritskiy, Sergiy Yashin, and Sergiy Koba

**Investigation of the Model of Testing for Weapons
 and Military Equipment** 309
 Ihor Korniienko, Svitlana Korniienko, Volodymyr Dmytriiev,
 Anatolii Pavlenko, and Dmytro Kamak

**Evaluation of the Spatial Resolution of Digital Aerospace Image
 by the Bidirectional Point Spread Function Parameterization** 317
 Sergey A. Stankevich

**Mathematical Modeling and Simulation of Systems
 in Project Management**

**Model for Human Capital Increasing via Motivation for General
 Training in Company Under Economic Crisis Overcoming** 331
 Liliia Nikiforova and Anatolii Shyian

**Modelling of Emotional Infection to the Information System
 Management Project Success** 341
 Sergey Bushuyev, Denis Bushuiev, and Victoria Bushuieva

**Modeling of Project Portfolio Management Process
 by CART Algorithm** 353
 Nataliia Yehorchenkova and Oleksii Yehorchenkov

**Modeling of Information Security Management System
 in the Project** 364
 Serhiy Shkarlet, Mariia Dorosh, Oleksandr Druzhyinin,
 Mariia Voitsekhovska, and Iryna Bohdan

Author Index 377

**Mathematical Modeling and Simulation
of Systems in Ecology and Geographic
Information Systems**



Use of Nuclear Emergency Response System for Assessment of Transboundary Transfer and Radiological Risks of the Potential Accidental Releases at Khmelnytsky NPP

Ivan Kovalets^{1,2}  and Oleksandr Romanenko³ 

¹ Institute of Mathematical Machines and System Problems NAS of Ukraine,
Prosp. Glushkova, 42, Kiev 03187, Ukraine

ivkov084@gmail.com

² Ukrainian Center of Environmental and Water Projects, Prosp. Glushkova,
42, Kiev 03187, Ukraine

³ Rivne Nuclear Power Plant, Rivne Reg., Varash 34400, Ukraine

Abstract. The JRODOS nuclear emergency response system was applied for estimation of risks caused by transboundary transport of possible accidental release from the newly designed units 3&4 of Khmelnytsky NPP. The total effective potential dose for 1-year children, received during 1 year after release from all pathways was calculated and used as measure of impact. About 4000 calculations of the dispersion scenarios were conducted using JRODOS atmospheric dispersion model and food dose model, with the start times of the releases falling within the time period of June–November, 2018. The numerical weather prediction data of WRF-Ukraine forecasting system operating in the Ukrainian Hydrometeorological Center were used to drive the JRODOS simulations. For each dispersion scenario maximum dose reached at the territory of every considered neighbor country (Belarus, Poland, Slovakia, Hungary, Romania, Moldova) was calculated. The simulated series of effective doses were fitted with GEV distribution. Taking into account probability of the accident the total probability of effective dose exceeding 1 mSv in at least one of neighboring countries is $9.74E - 07$ yr⁻¹, which corresponds to the frequency of occurrence of such event 1 time in 1 million years.

Keywords: RODOS · Dose · Transboundary transport · Risk assessment

1 Introduction

Nuclear emergency response systems (ERS) like RODOS [1], NARAC [2], Argos [3] and others are widely used for assessment of short-term and long-term consequences of accidental releases from nuclear power plants (NPPs) and in emergency preparedness trainings. Such systems typically contain models of different processes, such as atmospheric transport and deposition models, food chain and dose models, combined in the so-called ‘model chains’. This allows use of ERS for risk assessment of the consequences of potential accidental releases from NPPs. Hence when the scenario of the

potential release from NPP is available, it could be used for generation of calculation sets of atmospheric transport following release and finally of the potential doses received by people. The reliable risk estimates are possible only if generated sets cover wide range of meteorological conditions of atmospheric dispersion which include wind, precipitation and stability patterns. When the region of interest exceeds about 30 km [4] the conditions of constant in time and in space wind are unrealistic. Therefore the commonly used practice is to calculate a set of atmospheric dispersion scenarios following meteorological data representative of an extended period in the area of interest (e.g. [5, 6]). The JRODOS system, which is redesigned Java-version of RODOS [7], contains special functionality called ‘Statistic Output Tool’ which is specially designed to solve such task. To do this the JRODOS requires the input meteorological data generated by the numerical weather prediction (NWP) models on a relatively fine grid (less than 0.25 dec. deg. spatial resolution, at least 1 h time resolution). Thus generated input meteorological data are then further downscaled with the aid of the RODOS meteorological preprocessor on even finer grid [8] and the atmospheric dispersion is then performed with the aid of one of the atmospheric transport models available in Local-Scale Model Chain of RODOS [1], most of which are applicable to the distances of at least 800 km. Finally the statistical processing of the obtained results remains on responsibility of user. Therefore the goal of the present paper was to perform assessments of the transboundary transfer of the potential accidental releases at Khmelnytsky NPP in Ukraine with the JRODOS system which was recently implemented in Ukraine and to apply suitable statistical processing tools for estimation of the corresponding radiological risks.

2 Problem Description and Solution Methodology

2.1 Khmelnytsky NPP

Khmelnytsky Nuclear Power Plant (KhNPP) is operated by the Ukrainian “National Nuclear Energy Generating Company” (Energoatom). It is located in North-Western part of Ukraine (Fig. 1) in 150 km from the Ukrainian border with Belarus, 180 km from Poland, 210 km from Moldova, 230 km from Romania, 320 km from Slovakia, 380 km from Hungary. Presently two units VVER-1000/V-320 are in operation at KhNPP. Construction of the third and fourth units of KhNPP started in 1985. During a long period in 1990-th this construction was halted and renewed only after 2010. Presently the newly constructed units 3&4 of KhNPP undergo environmental impact assessment (EIA) procedures [9]. For that reason one of the tasks of this EIA is to evaluate possible radiological risks related to transboundary transfer of radioactivity following potential accidental releases at Khmelnytsky NPP units 3&4.

2.2 Meteorological Data and Source Term

Simulations of atmospheric transport following potential accidental releases from KhNPP were calculated using the meteorological data generated in operational runs of the WRF mesoscale meteorological model (www2.mmm.ucar.edu/wrf/users) which is

installed and used in Ukrainian Hydrometeorological Center (UHMC) as a part of WRF-Ukraine numerical weather prediction system. This system supply the RODOS system operated by UHMC and other governmental organizations in Ukraine with numerical weather prediction (NWP) data. The NWP data used in this work covered time period from 1 June to 11 November 2018. The computational domain covered territory of Ukraine and neighbor countries with grid resolution 0.25 dec. deg. In this work the meteorological calculations were performed using global final analysis data with 1 dec. deg. grid resolution of the Global Forecasting System (GFS) operated by the US National Centers of Environmental Prediction (NCEP) with 6 h update time period.

The source term for the potential accidental releases from KhNPP units 3&4 was estimated in [10] for accident with a ‘guillotine rupture’ of the main circulation pipeline. Accordingly the total released inventories are presented in Table 1. Furthermore the release duration was assumed 8 h. The release height was considered 300 m which is realistic value taking into account large heat content of the possible release. On the other hand, since in this work we are primarily interested in dose effects beyond 150 km, the results are not expected to be very sensitive to this parameter. As additional sensitivity tests had shown, the variation of the maximum calculated dose at distances from the source beyond 150 km and less than 500 km, for the release heights varying from 50 to 500 m does not exceed 15% as compared to average value of the maximum dose at the corresponding distance calculated for different release heights.

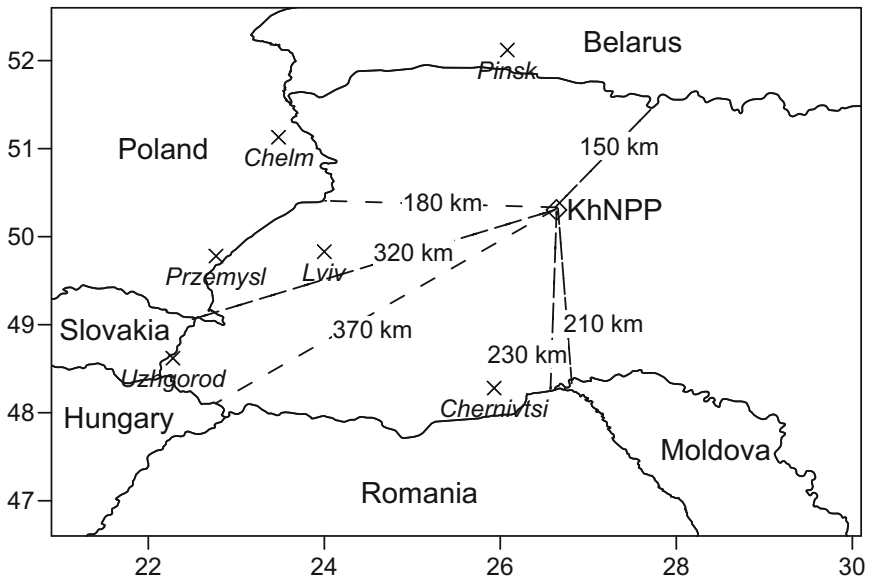


Fig. 1. Location of KhNPP and closest distances to the borders of neighbor countries.

Table 1. Release inventories for the potential accidental scenario [10].

Nuclide	Inventory, Bq	Nuclide	Inventory, Bq
I-131	8.792E+13	Zr-95	2.600E+10
I-132	7.610E+13	Nb-95	8.880E+10
I-133	2.052E+14	Ru-103	8.130E+11
I-135	4.687E+13	Ru-106	8.040E+10
Kr-85 m	1.920E+14	Cs-134	7.210E+11
Kr-85	1.170E+13	Cs-137	4.480E+11
Xe-133	2.180E+15	Ba-140	8.740E+11
Xe-135	4.670E+14	Ce-144	6.120E+11
Sr-90	4.090E+10		

2.3 Other Settings

For the estimate of transboundary impact of the potential release from KhNPP about 4000 calculations of the dispersion scenarios were conducted using JRODOS RIMPUFF atmospheric dispersion model and food dose model FDMT [1], with the start time of the release located between 01.06.2018 to 08.11.2018. The JRODOS Statistic Output Tool was used to automatically generate and perform dispersion scenarios within the considered time period. For each dispersion scenario atmospheric transport during 96 h following start of release was simulated and maps of total effective potential dose for 1-year children, received during 1 year after release from all pathways (hereafter refereed briefly as ‘effective dose’) were calculated. The 1-year children were chosen as a reference age group because preliminary testing had shown that for all distances from the point of release total effective doses for this group exceeded respective doses received by other age groups (5, 10, 15-years children and adults). Thus for each country the series of maximum effective doses reached in each dispersion scenario at the territory of the respective country were created.

For the statistical processing of the obtained series it is important that the results obtained for different dispersion scenarios are statistically independent. This may be violated start times of the releases of the neighbor dispersion scenarios are close. Therefore decorrelation procedure was applied for the simulated series of effective doses. The selected values were deleted from series until correlation coefficient of the obtained series of simulated effective doses and shifted series by one element was less than 0.05. Depending on distance from the release point this required separation between start times of the releases from 4 to 8 h. Thus for each country the decorrelated series of maximum effective doses corresponding to different dispersion scenarios were obtained. After this from the calculated series only elements with non-zero doses for the respective countries were chosen. The probability of the non-zero dose (or equivalently – probability of transport into the given country) was immediately calculated. This finally reduced series vector with nonzero doses was then statistically processed as described below.

3 Results of Calculations

3.1 Calculated Doses

Figure 2 presents as an example series of calculated maximum doses at distance 200 km from KhNPP. It is interesting that variability of this parameter created exclusively by meteorological conditions (source term is fixed) is very large and spans three orders of magnitude. As additional analysis had shown more than 99% of effective dose at this distance during the first year after the accident is created by food ingestion containing iodine isotopes.

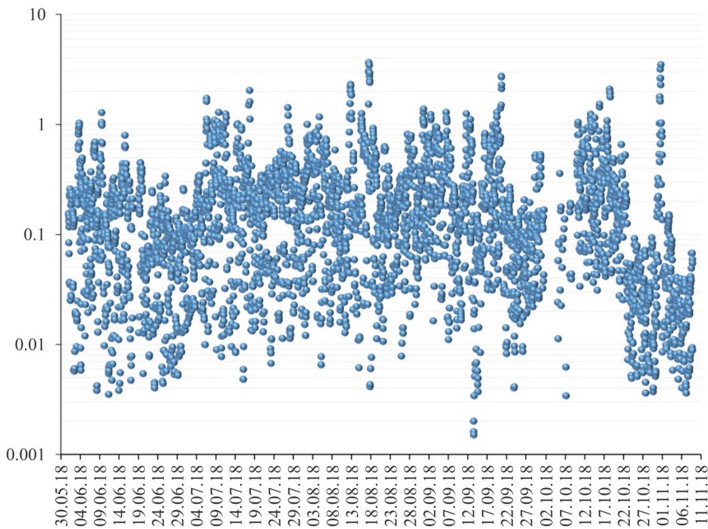


Fig. 2. Maximum effective dose (mSv) obtained in different simulations at distance 200 km from KhNPP.

Figure 3 presents locations of points in which maximum effective dose was reached at the territory of Belarus in different dispersion scenarios. Naturally the number of those points is significantly smaller than the number of points in Fig. 2 because not for all meteorological conditions plume reaches Belarus. Similar figures could be built also for other countries. The presented example of Belarus shows that not in all cases points of maximum dose are located near the border. Though the area density of points is largest near the border, but in some cases points of maximum dose are located more than in 200 km from the border. Therefore it is not possible to consider risk assessment for the specific country by analysis of the results only in a single point of this country, even if it is the nearest point to the release location (like points shown in Fig. 1).

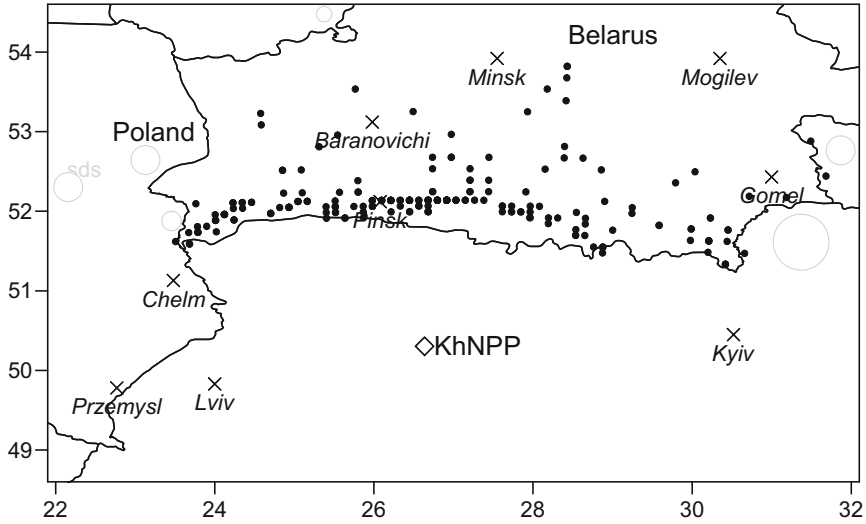


Fig. 3. Locations of maximum effective dose at the territory of Belarus reached in different dispersion scenarios following potential release from KhNPP.

Table 2 presents values of maximum effective doses reached at the territories of different countries in dispersion scenarios when plume was transported to the corresponding country. The probabilities of transport to countries are also presented in Table 2. This probability is greatest for Belarus (0.362) because Belarus is closest among all countries to KhNPP and also because the length of the joint border of Ukraine with Belarus is largest among all the considered countries. The maximum dose reached at Belarus (on average 0.118 mSv among all scenarios) is the largest as compared to other countries.

Table 2. Range of values of maximum effective doses reached at the territories of different countries in dispersion scenarios when plume was transported to the corresponding country.

Country	Distance from KhNPP to the nearest point at the border, km	Probability of transport to country	Min, mSv	Average, mSv	Max, mSv
Belarus	150	0.362	0.0001	0.1180	2.2711
Poland	180	0.176	0.0009	0.1204	0.7197
Slovakia	320	0.059	0.0002	0.0258	0.0754
Hungary	370	0.063	0.0001	0.0191	0.1030
Romania	230	0.300	0.0001	0.0404	0.3090
Moldova	210	0.126	0.0001	0.1418	0.7888

The probability of transport to Romania is also considerable (0.3). Though transport to Moldova is less probable (0.126) than transport to Romania, but the maximum doses obtained in Moldova (on average 0.14 mSv among all scenarios) are greater than maximum doses in Romania (on average 0.04 mSv) probably because of closer distance from Moldova to KhNPP. Despite considerable length and relatively close distance of joint border with Poland probability of transport to Poland (0.176) is more than by 2 times less than probability of transport to Belarus, because of prevailing transport of air masses from West. On average maximum doses reached in Poland are 0.12 mSv which is close to the corresponding value for Moldova. The smallest among all countries probabilities of transport and smallest maximum doses are obtained for Slovakia and Hungary.

3.2 Results of Statistical Processing

To determine the probability of effective dose exceeding 1 mSv as a result of a single release, the maximum dose series for each country presented in the previous sections were approximated by the generalized extreme value distribution function (GEV), which is often used for statistical analysis of extreme values [11]. The distributions of conditional probabilities of exceedance of the effective dose provided that plume was transported to corresponding country were obtained after fitting of the simulated series with GEV distribution using MATLAB `gevfit` function. The resulting distributions are presented in Fig. 4. The obtained conditional probabilities of exceeding the effective dose of 1 mSv for different countries are presented in Table 3.

As can be seen from Table 3, in all cases the probability of a dose exceeding 1 mSv is low, but differs quite different between different countries. The differences are due to both the different distances to these countries and the peculiarities of climatic conditions, which are naturally taken into account when calculating the consequences of identical emissions in different meteorological conditions. In Table 2 the maximum value of effective dose was obtained for Belarus (2.27 mSv), while according to Table 3 the probability of exceeding the dose of 1 mSv for Belarus is less than for Poland.

Table 3. Conditional and unconditional probabilities of exceedance of 1 mSv effective dose for different countries.

Country	Conditional probability of $D > 1$ mSv	Probability of $D > 1$ mSv
Belarus	0.011	3.1E-07
Poland	0.028	3.9E-07
Slovakia	0.003	1.4E-08
Hungary	0.005	2.5E-08
Romania	0.010	2.3E-07
Moldova	0.001	9.9E-09

This is explained both by the different length of the series and the fact that the approximating function cannot accurately reproduce the discrete distribution histogram due to smoothness requirements, and since the discrete histogram has an error related to the finiteness of a number of random values. Hence to clarify the effect of the error introduced by the finiteness of the series on the probabilities estimates, confidence intervals for estimating the probability of exceeding the dose of 1 mSv were also evaluated for Belarus and Poland by bootstrap method [11]. For Poland, the confidence interval of the studied probability turned out to be 0.8–5%, while for Belarus, this interval was 0.5–2%.

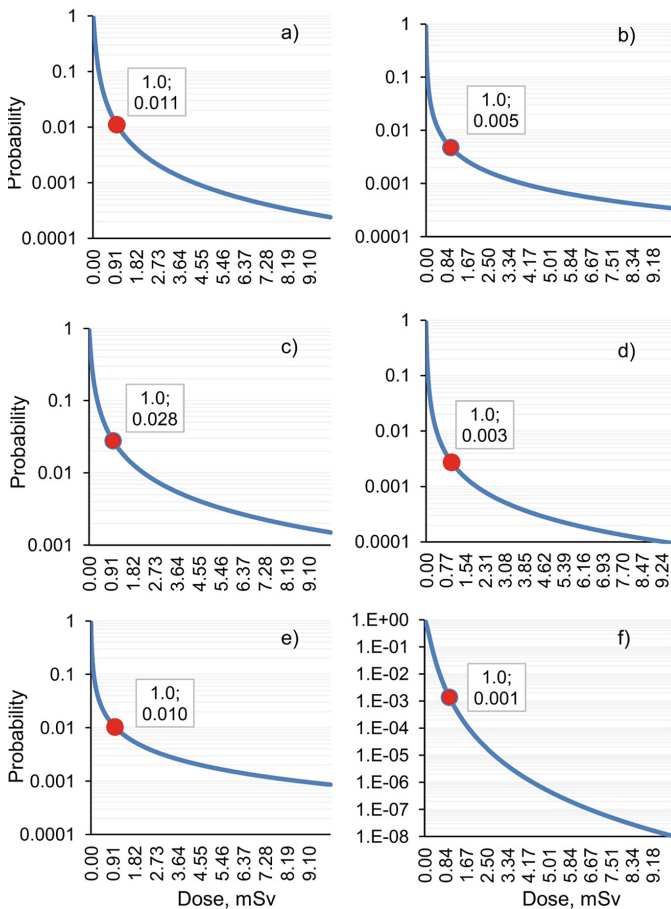


Fig. 4. Conditional probability of exceedance of given effective dose provided that plume was transported to corresponding country, obtained after fitting simulated series with GEV distribution for: a) Belarus; b) Hungary; c) Poland; d) Slovakia; e) Romania; f) Moldova. Circle shows probability of exceedance of 1 mSv.

The unconditional probability of effective dose exceeding 1 mSv could be obtained by multiplication of conditional probability presented in Table 3 on probability of transport to corresponding country (Table 2) and on probability of the accident. According to the document [12], the probability of an accident similar to the accident [10] considered in this paper is $7.82E - 05 \text{ yr}^{-1}$. The corresponding unconditional probabilities of effective dose exceeding 1 mSv for different countries are presented in Table 3. The largest of those probabilities is obtained for Poland ($3.9E - 07 \text{ yr}^{-1}$), which corresponds to the frequency of occurrence of such an accident 1 time in 2.6 million years. Finally the probability of effective dose exceeding 1 mSv in one of neighboring countries is $9.74E - 07 \text{ yr}^{-1}$, which corresponds to the frequency of occurrence of such event 1 time in 1 million years.

4 Conclusions

In this work we applied the JRODOS nuclear emergency response system for estimation of risks caused by transboundary transport of possible accidental release from the newly designed units 3&4 of Khmelnytsky NPP. The total effective potential dose for 1-year children received during 1 year after release from all pathways was calculated and used as measure of impact because total effective doses for this age group exceeded respective doses received by other age groups. About 4000 calculations of the dispersion scenarios were conducted using JRODOS atmospheric dispersion model and food dose model, with the start times of the releases falling within the time period of June–November, 2018. The numerical weather prediction data of WRF-Ukraine forecasting system operating in the Ukrainian Hydrometeorological Center were used to drive the JRODOS simulations. For each dispersion scenario maximum dose reached at the territory of every considered neighbor country (Belarus, Poland, Slovakia, Hungary, Romania, Moldova) was calculated. The decorrelation procedure was applied for the simulated series of effective doses to allow fitting of the simulated series with General-Extreme Value (GEV) distribution. Taking into account probability of the accident the total probability of effective dose exceeding 1 mSv in at least one of neighboring countries is $9.74E - 07 \text{ yr}^{-1}$, which corresponds to the frequency of occurrence of such event 1 time in 1 million years.


References

1. Landman, C., Päsler-Sauer, J., Raskob, W.: The decision support system RODOS. In: Kessler, G., Schlüter, F.-H., Raskob, W., Landman, C., Päsler-Sauer, G. (eds.) *The Risks of Nuclear Energy Technology*. Science Policy Reports, pp. 337–348. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-55116-1_21
2. Bradley, M.M.: NARAC: an emergency response resource for predicting the atmospheric dispersion and assessing the consequences of airborne radionuclides. *J. Environ. Radioact.* **96**, 116–121 (2007). <https://doi.org/10.1016/j.jenvrad.2007.01.020>

3. Hoe, S., McGinnity, P., Charnock, T., Gering, F., Schou Jacobsen, L.H., Havskov Sorensen, J., Andersson, K., Astrup, P.: ARGOS decision support system for emergency management. In: IAEA Proceedings Series STI/PUB/1460, pp. 1–10. International Atomic Energy Agency, Vienna (2010). http://www-pub.iaea.org/MTCD/Publications/PDF/P1460_Comp_CD/Start.pdf
4. Talerko, N.N.: Physical properties and limitations of radionuclides atmospheric transport models for different space and time scales. *Probl. Nucl. Power Plant. Saf. Chornobyl* **11**, 57–62 (2009). [in Russian]
5. Pirouzmand, A., Kowsar, Z., Dehghani, P.: Atmospheric dispersion assessment of radioactive materials during severe accident conditions for Bushehr nuclear power plant using HYSPLIT code. *Prog. Nuclear Energy* **108**, 169–178 (2018). <https://doi.org/10.1016/j.pnucene.2018.05.015>
6. Peltonen, T., Gering, F., Arnold, K., Duranova, T., Bujan, A., Duran, J., Bohun, L., Montero, M., Trueba, C.: Emergency preparedness for long lasting releases – assessment of radiological consequences. *Radioprotection* **51**, S79–S81 (2016). <https://doi.org/10.1051/radiopro/2016037>
7. Andronopoulos, S., Davakis, E., Bartzis, J.G., Kovalets, I.V.: RODOS meteorological pre-processor and atmospheric dispersion model DIPCOT: a model suite for radionuclides dispersion in complex terrain. *Radioprotection* **45**, S77–S84 (2010). <https://doi.org/10.1051/radiopro/2010017>
8. Ievdin, Y., Trybushny, D., Zheleznyak, M., Raskob, W.: RODOS re-engineering: aims and implementation details. *Radioprotection* **45**, S181–S190 (2010). <https://doi.org/10.1051/radiopro/2010024>
9. Becker, O., Mraz, G.: EIA Khmelnytsky 3&4 Procedure 2019. Final Expert Statement (Consultation Report). The Environmental Agency Austria Umweltbundesamt GmbH, REP-0699, Vienna (2019). <https://www.umweltbundesamt.at/fileadmin/site/publikationen/REP0692.pdf>
10. Power units №3,4 of Khmelnytsky. NPP construction. Vol. 13.14 Environmental Impact Assessment. Assessment of the effects of transboundary transport under normal and emergency conditions. (Строительство энергоблоков №3,4 Хмельницкой АЭС. Том 13.14. Оценка воздействия на окружающую среду. Оценка последствий трансграничного переноса при нормальных и аварийных режимах). Report № 43-814.203.004. ОЭ.13.14. NAEK «Energoatom, Kyiv, Ukraine (2016). (in Russian)
11. Wilks, D.S.: *Statistical methods in the atmospheric sciences*, 2nd edn. Elsevier, San Diego (2006)
12. Rivne NPP. Power unit No. 3. Security Analysis Report. Probabilistic Safety Assessment (PSA). Current Update Report on PSA. (Ривненская АЭС. Энергоблок №3. Отчет по анализу безопасности. Вероятностный анализ безопасности. Отчет по текущему обновлению ОВАБ). Report 22.3.133.ОБ.12.03.Ред.1, Rivne Nuclear Power Plant, NAEK «Energoatom», Varash, Ukraine (2019). (in Russian)



Evaluation of Spectral/Grid Nudging Methods for Weather Analysis and Forecasting in Kyiv Region with the Use of WRF Mesoscale Meteorological Model

Aleksander Khalchenkov^(✉)  and Ivan Kovalets 

Institute of Mathematical Machines and Systems Problems NAS of Ukraine,
Kiev, Ukraine

allexandro@ukr.net

Abstract. The possibility of using grid and spectral relaxation methods (nudging) and other options in the mesoscale model WRF for long-term continuous and forecast calculations on fine-resolution grid in the vicinity of Kyiv has been investigated. Results of comparison of selected meteorological parameters with surface measurements are presented. The basic recommendations for selecting the optimal combination of long-term calculation parameters are given. The use of the selected parameters allowed to obtain continuous meteorological fields over a long period (several months), which are well consistent with surface measurements. With grid and surface nudging and update of sea surface temperature for a temperature at a 2 m height in long-term continuous simulation a mean absolute error was $MAE = 1.63$ °C, and for wind speed at a 10 m height $MAE = 1.11$ m/s. Application of grid relaxation methods with grid surface nudging also allowed for improved results of calculations in forecast mode. The use of optimal combination allowed decrease of MAE for 2-m temperature from 2.06 °C to 1.85 °C for the first 48 h of forecast and from 2.35 °C to 2.14 °C for the next 49–96 h of forecast. The statistical measures of the quality of forecasted wind fields were also improved.

Keywords: WRF · Grid nudging · Spectral nudging · Meteorological fields · Data assimilation · Kyiv

1 Introduction

In many scientific and practical problems there occurs a need to estimate the state of the atmosphere in the past on a regular grid covering particular geographical region and prolonged time period – from several months to years. In several global products, such as final analysis archive of the Global Forecasting System (GSF) of the US National Center for Environmental Prediction (NCEP) [1] global fields of analyzed meteorological variables were calculated with using meteorological models and measurements combined by using data assimilation procedures and made available on a regular grids typically having spatial resolutions from 0.15 to 1 dec. deg (0.25 dec. deg in case of [1]). However in many applications, for instance, in urban air pollution

studies, occurs a need to calculate meteorological fields on even finer grids for certain historical time periods. If this is done in a single run of limited-area mesoscale meteorological model such as WRF [2] and using data of global reanalysis fields to specify boundary and initial conditions, the following known problem may occur. When historical time period is sufficiently long (e.g. a few months and more) the resulting synoptic-scale features obtained in the run of mesoscale model could unrealistically deviate from the input reanalysis data as a result of the inherent sensitivity of non-linear equations of atmospheric dynamics to small changes in initial and boundary conditions [3]. One way to overcome this difficulty is to split historical time period in a sequence of short-duration calculation periods having length from 48 to 96 h [3]. However this approach may lead to unphysical discontinuities in meteorological variables. Alternative approach is to use ‘nudging’ methods previously implemented in WRF model [4, 5]. This approach was tested in previous work (to be published) for the problem of reanalysis of meteorological conditions for the territory of Ukraine on a grid with resolution 0.15 deg. However when meteorological conditions are simulated on even finer grid new physical processes of smaller scale are resolved by the model, such as urban-scale circulations, convection. Therefore the goal of the present work was to test the nudging method for reanalysis of meteorological fields in fine-scale long-term simulations of weather conditions for the Kyiv region. Secondly we for the first time apply and test this method in forecasting mode for Kyiv region.

2 Model Setup and Numerical Experiment Description

The simulation was performed using the WRF model [2] in nested domains. A large external domain with dimensions 120×80 horizontal grids points with spacing of 0.15 degrees was created to cover the whole of Ukraine. An internal domain (Fig. 1) with 49×37 horizontal grid points with spacing of 0.025 degrees was created for detailed modeling in the vicinity of Kyiv. In both domains 28 vertical levels were used. The Modis_15s global dataset was used to obtain land use categories and the topo_gmt-ed2010_30s dataset was used to obtain topographical data. The resolution of input geographical data (15s for land use and 30s for topography) is sufficient for meteorological simulations on computational grid having 0.025 dec. deg. (90 s) resolution around Kyiv. However if resolution of computational grid is further increased it is reasonable to use more detailed geographical data such as products of Sentinel-2 (<https://land.copernicus.eu/global/products/lc>). Some parameters are responsible for physical processes, which supposedly should not significantly affect the differences between long-term and short-term modeling, were fixed in all the calculations. WSM 6-class graupel scheme was used for microphysics. For the long-wave radiation, we used the RRTM scheme, for the short-wave radiation Goddard Shortwave scheme. For the surface layer model, the Monin-Obukhov (Janjic Eta) scheme based on the similarity theory and for the boundary layer model, the Mellor-Yamada-Janjic (Eta) TKE scheme parameterization was used. The Kain-Fritsch cumulus region model was turned on only for the outer region and turned off for the detailed region. Therefore, the feedback option was disabled in all performed calculations. All mentioned parameterizations are described in [2].

Modeling was performed for the period from 2019/01/01 to 2019/06/11. This period includes the cold and warm seasons of the year, which are characterized by significantly different characteristics of meteorological conditions and processes (the presence or absence of snow cover, the presence or absence of ice on the surface of water bodies, soil heating, rainfall of various types, etc.). The end of the calculation on 2019.06.11 is due to the fact that at this moment the number of vertical levels in the global GFS model changed (in the WRF model, it is not possible to use boundary data with a different number of vertical levels within one continuous calculation). For initial and boundary conditions in long term calculations we used the $0.25^\circ \times 0.25^\circ$ NCEP Global Final Analysis (FNL) data [1] with 6 h' time step. In forecast mode we used $0.5^\circ \times 0.5^\circ$ NCEP Global Forecast System (GFS) data with 3 h' time step.

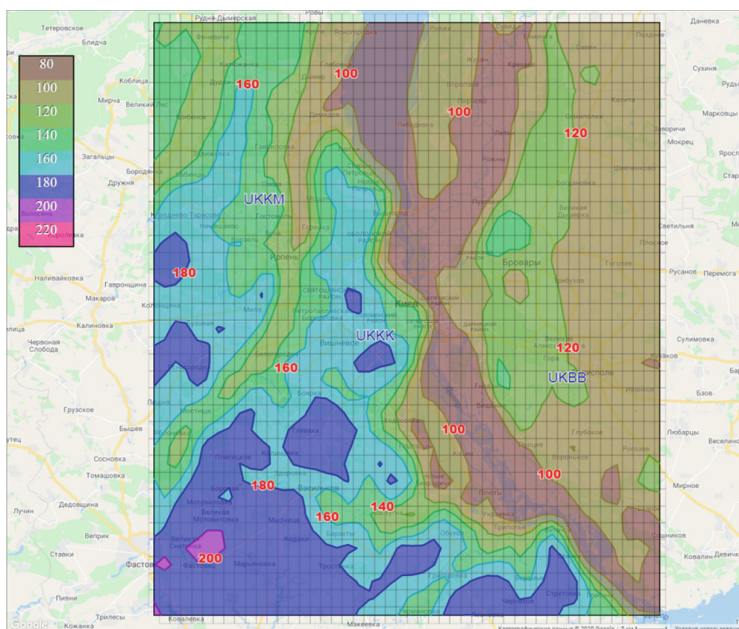


Fig. 1. Internal computational domain (d02) with fine-scale grid, topography, and surface weather stations

In this study, we first examined the effect of common parameters that are likely to affect the quality of long-term modeling, such as `sst_update` (update of sea surface temperature), `grid_fdda` (grid nudging), `grid_sfdda` (surface nudging). All these options were turned on only for the external domain. The simulation results in the detailed internal region changed only due to changes in the boundary conditions due to changes in the fields in the external domain. After that, for the chosen optimal combination, the influence of the different surface physics options was investigated. At the final stage, the optimal combination was checked when operating in the prognostic mode.

To apply grid surface nudging is necessary to create the wrfsfdda_d01 file. The Obsgrid program was used to create these files. This program was created by NCEP to assimilate measurements and fields of the global model in the calculation process. To create a wrfsfdda_d01 file, we did not use raw measurements, but only assimilated the data already contained in met_em files created by the real.exe program, which already contain information from measurements assimilated during analysis procedure of GFS model.

Comparison was performed for the results obtained in the inner domain with measurements of three weather stations (Fig. 1) located at the airports: UKKK (Kyiv, Zhulyany), UKBB (Boryspil), UKKM (Antonov, Gostomel). Basic statistical comparisons were made for temperature at 2 meters, wind speed and wind direction at 10 meters. For the statistical evaluation of the simulation results, we used the following measures: mean error (ME), mean absolute error (MAE) and correlation coefficient between the simulated and measured data. Precipitation reproduction quality assessment was performed only for UKKK station. All the comparisons in this Sects. 3 and 4 (analysis calculations) were performed against data of WMO stations that were involved in the creation of final analysis data which were used as input. Only in Sect. 5 (the forecast mode) the comparisons are performed against independent measurement data.

3 Results with Varying the General Parameters

Table 1 shows all the calculations performed during testing of the selected general parameters, which can affect the quality of long-term continuous modeling. Noah Land-Surface Model was used in all these calculations. To additionally assess the quality of long-term continuous calculations, short-term calculations (48h_restart in tables) were performed that fully covered the selected comparison period (the duration of each individual calculation was 54 h, of which the first 6 h were spent on initialization and were not taken into account in the comparison).

Table 1. Calculations with different the general parameters

ID	Case	WRF options		
		sst_update	grid_fdda	grid_sfdda
1	48h_restart	0	0	0
2	continuous	0	0	0
3	sst_update	1	0	0
4	gfdda	0	1	0
5	gfdda&sst_update	1	1	0
6	gfdda2&sst_update	1	2	0
7	gfdda&sfdda&sst_update	1	1	1
8	gfdda&sfdda2&sst_update	1	1	2

Tables 2 and 3 show the systematic error, the mean absolute error, and the correlation coefficient Corr for temperature at 2 m, precipitation accumulated for 24 h, wind speed and wind direction at 10 m. Statistical measures for precipitation were calculated only for one surface weather station (UKKK), which reported 283 mm of rainfall during comparison period from 2019/01/01 to 2019/06/11. Automatic updating of the water surface temperature positively affected the modeling of temperature at 2 m level, and is less statistically noticeable when modeling wind fields. Enabling this option reduced MAE for temperature from 2.11 to 2.04 °C.

Table 2. Statistical variables for temperature at 2 m level and precipitations.

ID	Case	Temperature, °C			Precipitation, mm/24 h		
		ME	MAE	Correlation	ME	MAE	Correlation
1	48h_restart	-0.47	1.73	0.97	0.26	1.88	0.45
2	continuous	-0.47	2.11	0.97	0.84	2.03	0.50
3	sst_update	-0.81	2.04	0.97	0.38	1.96	0.30
4	gfdda	-0.19	2.10	0.97	0.46	2.08	0.38
5	gfdda&sst_update	-0.47	2.03	0.97	0.62	2.23	0.43
6	gfdda2&sst_update	-0.54	2.05	0.97	1.36	2.72	0.37
7	gfdda&sfdda&sst_update	-0.43	1.63	0.98	0.39	1.99	0.41
8	gfdda&sfdda2&sst_update	-0.51	2.11	0.96	0.93	2.20	0.41

Table 3. Statistical variables for wind speed and wind direction at 10 m level

ID	Case	Winds Speed (m/sec)			Wind Direction (deg)		
		ME	MAE	Correlation	ME	MAE	Correlation
1	48h_restart	0.97	1.32	0.76	3.55	23.37	0.64
2	continuous	1.33	1.46	0.71	5.88	27.17	0.62
3	sst_update	1.34	1.39	0.70	4.61	27.71	0.62
4	gfdda	0.54	1.26	0.75	4.31	24.30	0.65
5	gfdda&sst_update	0.93	1.37	0.75	3.72	24.50	0.69
6	gfdda2&sst_update	0.96	1.43	0.75	2.88	25.34	0.64
7	gfdda&sfdda&sst_update	0.61	1.11	0.79	2.24	22.27	0.68
8	gfdda&sfdda2&sst_update	0.17	1.33	0.71	-10.11	30.11	0.63

The use of grid relaxation (option `grid_fdda = 1`) had a noticeable positive effect in modeling wind fields. Enabling this option allowed reducing MAE for wind speed from 1.46 m/s to 1.29 m/s and increasing the correlation coefficient from 0.71 to 0.75. For wind direction, MAE decreased from 27.17° to 24.30 with an increase in the correlation coefficient from 0.62 to 0.65. In addition, similarly to other works (for example [6]), it was possible to maintain the correct synoptic-scale structure of the pressure field in the long-term calculation. Figure 2 presents the example of simulated temperature at height

2 m, wind speed and wind direction at height 10 m and observations for surface weather station in Zhulyany (UKKK).

The use of spectral relaxation (option `grid_fdda = 2`, with additional parameters `xwavenum = 3` and `ywavenum = 3`) allowed to obtain close statistical results to the calculations with grid relaxation. Both relaxation options implemented in the WRF model improve the results and are recommended for use in continuous long-term modeling of meteorological conditions.

Using the option of grid surface nudging (`grid_sfdda = 1`) has significantly improved the modeling of surface meteorological fields. Simulation with the turning on of this option allowed to obtain the best results for all the studied variables. This allowed us to reduce the average error for temperature to $1.63\text{ }^{\circ}\text{C}$ and increase the coefficient of correlation to 0.98. For wind speed MAE reduced to 1.11 m/s and coefficient of correlation increased to 0.79. For wind direction MAE reduced to 22.27 deg and coefficient of correlation increased to 0.68. The FASDAS method of surface grid nudging (`grid_sfdda = 2`, flux-adjusted surface data assimilation system) turned out to be statistically worse than a number of other calculations.

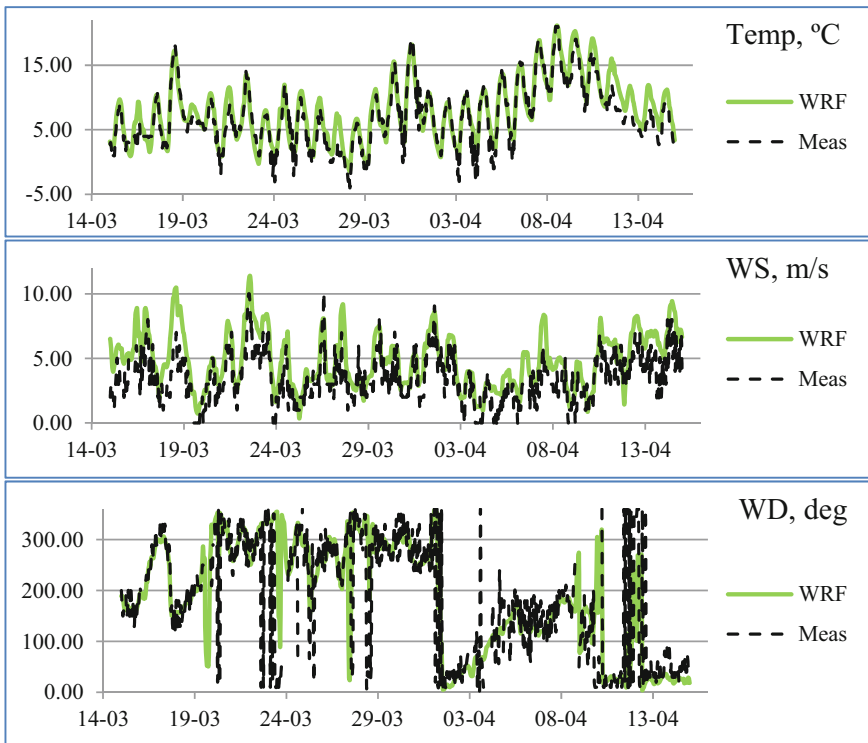


Fig. 2. Temperature at height 2 m, wind speed and wind direction at height 10 m for surface weather station in Zhulyany (UKKK). Solid line – WRF simulations (`gfdda&sst_update` case), dashed line – measurements.

4 Results with Varying the Surface Physics

The land surface model (`sf_surface_physics`) is one of the most important options for long-term modeling, as it is responsible for the properties of the underlying surface (temperature, humidity, albedo, snow depth, etc.). We tested four different land surface models (LSM) and evaluated their impact on the quality of continuous long-term modeling. The following models took part in testing: NOAH LSM (four layers of soil), RUC LSM (nine layers of soil), CLM4 (ten layers of soil), PLEIM-XIU (two layers of soil).

During testing, we found that the RUC earth surface model in WRFV3.9 (and newer versions) contains a critical error, which leads to an incorrect estimate (over-estimation) of the thickness of the snow cover. Therefore, for the RUC model, we used older WRFV 3.8.1, all other calculations were performed using WRFV 3.9. All calculations were performed with the turning on of updating the sea surface temperature and using grid nudging. The statistical results of comparing calculations with measurements of surface stations are given in Tables 4 and 5.

Table 4. Statistical variables for temperature at 2 m level for different land surface model

ID	Case	Temperature 2 m, °C			Precipitations, mm/24 h		
		ME	MAE	Correlation	ME	MAE	Correlation
1	NOAH	-0.47	2.03	0.97	0.62	2.23	0.43
2	RUC9	-1.33	2.38	0.96	1.31	2.38	0.54
3	CLM	0.54	1.93	0.97	0.21	1.70	0.42
4	PLXU	-2.68	3.02	0.96	-0.17	1.55	0.64

The best results for temperature modeling were obtained with the ten-layer model CLM4 (MAE = 1.93 °C). The two-layer PLXU model resulted in the worst results for temperature (MAE = 3.02 °C). When modeling wind speed at a height of 10 m, the best results were obtained using models PLXU (MAE = 1.03 m/sec) and RUC (MAE = 1.09 m/sec). At the same time, for the wind direction, the minimum error was obtained using the model NOAH (MAE = 24.50 deg).

Table 5. Statistical variables for wind speed and wind direction at 10 m level for different land surface model

ID	Case	Winds speed (m/sec)			Wind direction (deg)		
		ME	MAE	Corr	ME	MAE	Corr
1	NOAH	0.93	1.37	0.75	3.72	24.50	0.69
2	RUC9	0.44	1.09	0.76	-1.83	26.08	0.66
3	CLM	0.77	1.28	0.72	6.17	27.12	0.65
4	PLXU	0.16	1.03	0.75	-5.26	26.42	0.67

The best results for precipitation were found for the PLXU model (see Fig. 3). The NOAH and RUC models showed good agreement with snow depth obtained from the final analysis data (see Fig. 4). Moreover, the NOAH LSM reproduced better than other models the time of complete disappearance of the snow cover. Models CLM4 and PLXU significantly underestimated snow depth.

With continuous long-term calculation, none of the models was able to reproduce well the temperature of the lower soil layer given in the final analysis. This is probably due to difference in the depth of the lowest soil layer where a fixed boundary condition for the temperature is set.

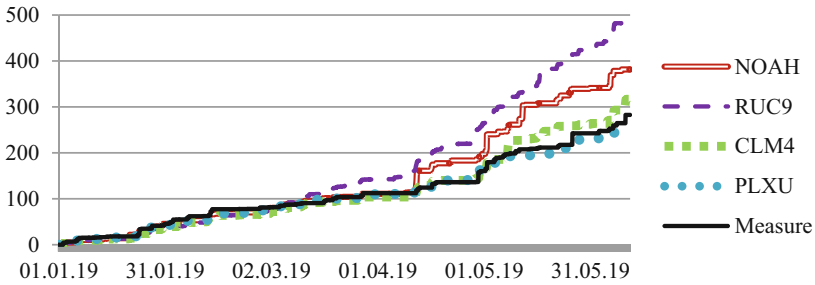


Fig. 3. Accumulated total precipitations, mm (UKKK)

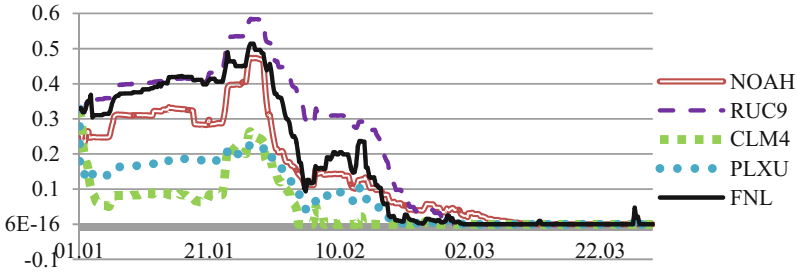


Fig. 4. Dynamics of snow depth in the center of the detailed domain, m

5 The Forecast Mode

In the previous sections we found that the use of grid nudging and option `sst_update` could significantly improve the results of long-term modeling of meteorological conditions in the vicinity of Kyiv. In this section, we will check the effect of these options on the quality of short-term modeling in the forecast mode. For testing, we performed four sets of predictive calculations: 1) without grid nudging and sea surface temperature updating (normal); 2) with turned on grid nudging and sea surface temperature (`wrfda`); 3) with turned on grid nudging and sea surface temperature only for the first 48 h of forecast (`wrfda48`); 4) with turned on grid nudging and sea surface temperature and turned on surface grid nudging (`wrfda&sfdda`);

Each set of calculations includes 80 separate prognostic calculations lasting 96 h. The period between individual calculations is 48 h. The performed series of calculations fully covered the same period over which the comparison in the previous sections was carried out (from 01/01/2011 to 2019/06/11). The comparison was carried out separately for the first 48 h of each forecast, and for the period from the 49th to the 96th hour of each forecast. Statistical error indicators of the calculated results are presented in Tables 6, 7 and 8.

The use of grid nudging allowed us to slightly improve all the simulation results in the prognostic mode. MAE for temperature decreased from 2.06 °C to 1.99 °C for the first 48 h and from 2.35 °C to 2.32 °C for the next 49–96 h. For wind speed MAE decreased from 1.51 m/sec to 1.42 m/sec for the first 48 h of forecast and from 1.84 m/sec to 1.71 m/s for the next 49–96 h.

The turning on of surface nudging had noticeable positive effect on the quality of forecasted temperature. The MAE additionally decreased from 1.99 °C to 1.85 °C for first 48 h of forecast and from 2.32 °C to 2.14 °C for the last 49–96 h. There is no noticeable error change for the wind direction, but using this option allowed to increase the correlation coefficient from 0.63 to 0.66 for the first 48 h and from 0.55 to 0.57 for period covered 48–96 h of forecast.

Table 6. Statistical variables for temperature at 2 m level for calculations in forecast mode

ID	Case	0–48 h			49–96 h		
		ME	MAE	Corr	ME	MAE	Corr
1	normal	−0.42	2.06	0.97	−0.17	2.35	0.96
2	wrfda	−0.39	1.99	0.97	−0.11	2.32	0.96
3	wrfda48	−0.39	1.99	0.97	−0.11	2.36	0.96
4	wrfda&sfdda	−0.25	1.85	0.97	−0.01	2.14	0.96

Table 7. Statistical variables for wind speed at 10 m level for calculations in forecast mode

ID	Case	0–48 h			49–96 h		
		ME	MAE	Corr	ME	MAE	Corr
1	normal	1.18	1.51	0.73	1.45	1.84	0.62
2	wrfda	1.07	1.42	0.74	1.31	1.71	0.64
3	wrfda48	1.07	1.42	0.74	1.40	1.78	0.63
4	wrfda&sfdda	1.07	1.43	0.73	1.31	1.70	0.63

Table 8. Statistical variables for wind direction at 10 m level for calculations in forecast mode

ID	Case	0–48 h			49–96 h		
		ME	MAE	Corr	ME	MAE	Corr
1	normal	1.56	28.06	0.64	2.27	35.44	0.55
2	wrfda	1.57	28.25	0.63	0.68	34.84	0.55
3	wrfda48	1.57	28.25	0.63	1.73	35.20	0.55
4	wrfda&sfdda	0.27	28.45	0.66	−1.12	34.90	0.57

6 Conclusions

The effectiveness of using grid and spectral relaxation methods and other options in the mesoscale model WRF for long-term historical calculations and in forecast mode on the detailed grid covering Kyiv region has been investigated for the period from January 1, 2019 to June 6, 2019. The use of grid nudging combined with the sea surface temperature updating option in WRF model allows long-term multi-month calculations with improved simulation results on detailed grid. Errors obtained for the internal high-resolution domain during long-term calculations using the grid nudging and grid surface nudging are smaller than the errors of the run splitted on sequence of short-term calculations. In long-term historical run with grid and surface nudging and update of sea surface temperature for a temperature at a 2 m height a mean absolute error was $MAE = 1.63$ °C, and for wind speed at a 10-m height $MAE = 1.11$ m/s. The obtained statistical indicators of the model performance are consistent with the results presented in similar works such as [6, 7].

The disadvantage of the CLM4 model is its complexity and significant requirements for computing resources (the calculation time using this surface model turned out to be approximately 8 times longer than the calculation time of other models). Models CLM4 and PLXU significantly underestimated the height of the snow cover, but better than other models reproduced precipitation in the warm season. In future studies it may also be useful to use for land surface model validation the Sentinel-3 Land Surface Temperature product and Snow and Ice Monitoring product.

Grid relaxation methods (especially grid surface nudging) also allowed for improved results of calculations in the prognostic mode. For temperature at 2 m height MAE decreased from 2.06 °C to 1.85 °C for the first 48 h of forecast and MAE decreased from 2.35 °C to 2.14 °C for next 49–96 h of forecast. The quality of forecasted wind fields using relaxation techniques was also improved.





References

1. NCEP GDAS/FNL 0.25 Degree Global Tropospheric Analyses and Forecast Grids. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce (2015). <https://doi.org/10.5065/D65Q4T4Z>. Accessed 20 May 2020
2. Skamarock, W.C., Klemp, J.B., Dudhia, J., et al.: A description of the advanced research WRF version 3 (No. NCAR/TN-475 + STR). University Corporation for Atmospheric Research, Boulder, Colorado (2008). <https://doi.org/10.5065/d68s4mvh>
3. Lo, J.C.F., Yang, Z.L., Pielke, R.A.: Assessment of three dynamical climate downscaling methods using the weather research and forecasting (WRF) model. *J. Geophys. Res.* **113**, D09112 (2008). <https://doi.org/10.1029/2007JD009216>
4. Staufer, D.R., Seaman, N.L.: Use of four-dimensional data assimilation in a limited area mesoscale model. part I: experiments with synoptic-scale data. *Mon. Weather Rev.* **118**, 1250–1277 (1990)

5. Míguez-Macho, G., Stenchikov, G.L., Robock, A.: Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations. *J. Geophys. Res.* **109**, D13104 (2004). <https://doi.org/10.1029/2003JD004495>
6. Omrani, H., Drobinski, P., Dubos, T.: Using nudging to improve global-regional dynamic consistency in limited-area climate modeling: what should we nudge? *Clim. Dyn.* **44**, 1627–1644 (2015). <https://doi.org/10.1007/s00382-014-2453-5>
7. Astrup, P., Mikkelsen, T.: Comparison of NWP prognostic and local monitoring data from NPPs. *Radioprotection* **45**, S97–S111 (2010). <https://doi.org/10.1051/radiopro/2010019>
8. Bhati, S., Mohan, M.: WRF model evaluation for the urban heat island assessment under varying land use/land cover and reference site conditions. *Theor. Appl. Climatol.* **126**, 385–400 (2015). <https://doi.org/10.1007/s00704-015-1589-5>



Simulation Modeling for Predicting the Formation of Municipal Waste

Viktoriia Khrutba¹ , Tetiana Morozova¹ , Iryna Kotsiuba² ,
and Volodymyr Shamrai² 

¹ National Transport University, Omelianovycha-Pavlenko Ave. 1,
Kiev 01010, Ukraine

viktoriia.khrutba@gmail.com

² Zhytomyr Polytechnic State University, Chudnivska Str., 103,
Zhytomyr 10005, Ukraine

Abstract. The study was carried out to evaluate the amount and composition of SMW and to determine the potential for waste treatment. The simulation model that underlies the information subsystem of management support for solid municipal waste management has been developed. The system dynamics model was used to estimate the average annual amount of SMW. Simulation was performed using the AnyLogic 7 environment. To analyze the sensitivity of the model, it is necessary to consider the main factors of influence: population, waste components.

For developing a simulation model, regression equations were used: the dynamics of population change, the dynamics of changes in the generated waste, the dynamics of changes in the volume of housing stock, the dynamics of changes indicators of retail trade and public catering, the dynamics of changes industrial production, the dynamics of changes household income.

The used amount depends on the produced and recycled wastes, while the compacted density is constant. This means that if the amount of collected wastes increases as a result of population decline, the volume of use will increase. The model allows to carry out the forecast of volumes of the generated waste for decision-making in system of regional management.

Keywords: Mathematical modeling · Solid waste management · Mathematical model · Morphological composition of waste

1 Prerequisites for Mathematical and Experimental Simulation for the Prediction of Municipal Waste Generation

1.1 Problem Statement and the Purpose of Research

According to the National Strategy of waste management in Ukraine until 2030 and the National Waste Management Plan until 2030 in the areas of Ukraine needs to develop a regional waste management plan. Much attention is paid to municipal waste management (municipal and municipal waste), details of their generation, and current state

of the system and the scenario of system development [1, 2]. The data on volumes of waste, their morphological structure, the number of processing and disposal can be obtained from official sources, but they are fragmented. That is why we have developed our own approach to determining the generation of municipal waste in settlements of Ukraine, which is described in this article.

The aim of this research is development of simulation prognostic model of waste generation.

1.2 Analysis of Recent Research and Publications

Municipal waste generation models are required for adequate planning and management of environmental impacts. The issue of predicting the generation of municipal waste has been addressed by various authors. Thus, V.M. Adamović et al. [3] have developed a general regression network model (GRNM) based on available social, economic, demographic and sustainability indicators. C. Ghinea et al. [4] used regression and time series analysis in their predictive models, using the following variables as baseline variables: population, age, urban life expectancy, total SMW.

1.3 Materials and Methods of Researches

The methodological basis for evaluating and selecting ways to improve the environmental safety of the region under consideration by improving the efficiency of the waste management system is a systematic approach that is increasingly used in research both at the stage of analysis of an existing problem and in the process of making management decisions. System analysis is the most constructive area that is used to put systems theory into practice in management tasks. Constructiveness of system analysis is connected with development of the method of carrying out works, which allows to take into account all essential factors that allow to build effective control systems in specific conditions [5].

As defined in [6], the introduction of waste management technologies includes measures aimed at preventing the generation of waste, its collection, transportation, storage, treatment, disposal, burial, solid waste disposal, as well as control over these operations and control sites for disposal.

If we designate by function $G(x, \xi, t, \tau)$ the waste system of the region, through $u(x, t)$ – the waste management system, which is formed under the influence of the following factors: volume of waste generation; basic professional ideas; $\bar{\omega}_u(x, t)$ – the distribution of management influence takes into account existing technological concepts; $\bar{\omega}_{b,z}(x, t)$ – external influences take into account the attitude of the society and its financial capabilities; $\bar{\omega}(\xi, \tau)$ – the standardization function of mitigation is based on the legal framework and the attitude of the state; then a generalized scheme of the waste management system can be constructed (Fig. 1).

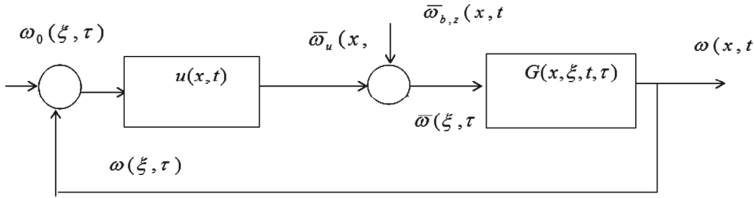


Fig. 1. Generalized scheme of waste management system: $\omega_0(\xi, \tau)$ – the initial state of waste generation in the region; $\omega(x, t)$ – the final state of waste generation in the region

We will apply the method of morphological analysis for construction a standardizing control function [7]. The essence of the method is that the system distinguishes several characteristic for its main functional elements of morphological features, each of which make the most complete list of different specific options (alternatives) expression of these features. Each characteristic characterizes the indicator, function, mode of operation (or state) of the system, the form of interaction of indicators (elements), etc., on which the solution of the problem depends and the main purpose of functioning of the system of distribution of project resources is achieved. The signs with their alternatives are arranged in the form of a morphological matrix, according to which it is possible to form a morphological model, which allows to systematically analyze the different structures of the waste management system, taking into account the peculiarities of the region.

In balance modifications, the appearance of waste is estimated according to information in the field of product use, trade, consumption of goods. Factor modifications are based on the consideration of conditions that show the formation of waste. Statistical modifications are regularities of change of formation of MSW. For today, the best and most accurate method, which covers all parameters, is a comprehensive method (factorial with elements of statistics). The method is based on retrospective analysis, which allows to identify the significance of the conditions. The model must take into account all the conditions that affect the studied indicator.

The proposed model of system dynamics identifies key elements that should be quantified as variables, and their effects are formulated mathematically. The model is finally determined when the parameters and initial values for state variables (stocks) are defined. For forecasting, you need to choose parameters that can be predicted with high accuracy for a long forecasting horizon.

The software product consists of four forms. The main window of the program presents eight controls (eight main blocks shown in the diagram), which allow to make settings for further forecasting the volume of municipal solid waste. The functionality of the program is presented in detail schematically in the Fig. 2.

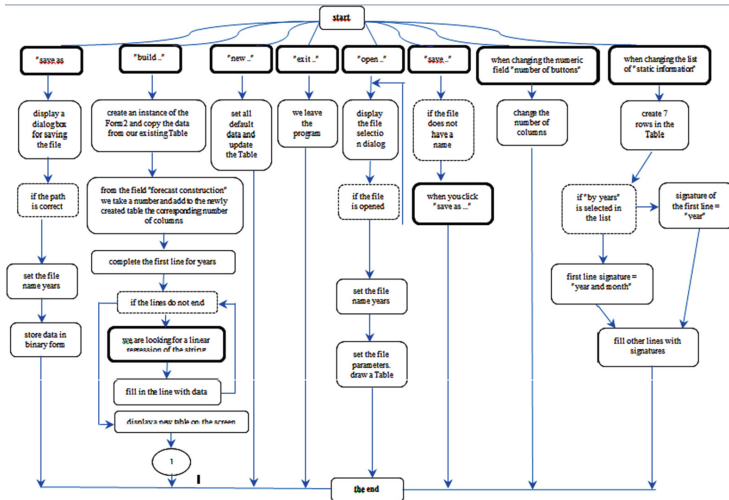


Fig. 2. Schematic model of the program «Municipal waste counter (MWC)»

The study uses the methodology of system dynamics Forrester. The software product consists of four forms. The main window of the program presents eight controls, which allow you to make settings for further forecasting the volume of solid waste.

The software product «Municipal waste counter (MWC)» was developed using the Visual C 7.0 programming language in the Microsoft Visual Studio 2017 programming environment using the .NET Framework 4.5.1.

The project is designed as a Windows Forms application. To run, you must have Windows 7/8/8.1 installed on your computer with the .NET Framework version at least 4.5.1 or Windows 10 installed. The project contains four forms: Form1, Form2 and Form3 and Form4. The main form of the Form1 project gives access to the main functions of working with files (loading and saving data), selection of a territorial feature, input of the initial information which is necessary for forecasting, etc.

2 Construction of Mathematical Simulation Models for the Prediction of Solid Municipal Waste (SMW) Production

2.1 Forecasting of the Volume of SMW Production

It is known that the quantitative and qualitative characteristics of municipal waste depend on the sources of their generation. In general, they include: paper, cardboard, glass, metals, plastic, bio-waste, wood, textiles, packaging, waste electrical and electronic equipment, waste batteries, batteries and accumulators, as well as bulky waste. Exact information on the composition and quantitative indicators of municipal waste management in Zhytomyr was not available at the time of the study. Therefore, we

used the “assumptions” approach to forecast. The data of 2016 year provided a “null hypothesis”. Population of Zhytomyr: in 2016 - 267.6 thousand people [8].

For every 1000 births and deaths were 10.4 and 16.2, respectively. The difference between births and deaths generates a net population (C_n), which is used to calculate waste generation (Q). The population growth rate was used to calculate waste generation. A production volume of SMW depends on the volume of housing stock (G_t), indicators of retail trade and public catering (P_n), industrial production (R_n), and household income (D). The dynamics of change according to [8] can be represented in the form of regression models (1)–(6):

✓ the dynamics of population change:

$$C_n = 0,3299c^4 - 3,9136c^3 + 15,963c^2 - 30,925c + 282,66, \quad (1)$$

where c - the quantity of population in the region, thousands of people;

✓ the dynamics of changes in the generated waste:

$$Q = -12500q^4 + 141667q^3 - 537500q^2 + 825054q + 161896, \quad (2)$$

where q - the quantity of waste generated, m^3 ;

✓ the dynamics of changes in the volume of housing stock:

$$G_t = -3,7916g^4 + 45,584g^3 - 189,72g^2 + 318,22g + 5215, \quad (3)$$

where g - housing, million square meters;

✓ the dynamics of changes indicators of retail trade and public catering:

$$P_n = 143,09p - 282660, \quad (4)$$

where p - indicators of retail trade and public catering, million UAN;

✓ the dynamics of changes industrial production:

$$R_n = 298,13r - 589634, \quad (5)$$

where r - industrial production, mil. UAN;

✓ the dynamics of changes household income:

$$D = 108,07d^2 - 433477d, \quad (6)$$

where d - household income, mil. UAN.

Using the obtained approximate dependencies, we can determine the predictive value and volume factors specific production of waste in the near future.

The dependence of actual, estimated and projected volumes of SMW in the period from 2009 to 2021 years for Zhytomyr as the main link of the waste-producing object indicates their growth. This function is most influenced by industrial production and retail trade. This is explained by the size of the population and the dependence of living standards on the development of certain industries of the city. The increase in the volume of SMW is due to the increase in the volume of housing stock, the growth of industrial production, retail trade and catering, and the increase in income.

The analysis of the data shows that the growth of each resource corresponds to the growth of SMW (Fig. 3, 4). There is no data on the change of each isolated resource, but it is known that the standard of living of the population remains low. This suggests that the function will increase as each argument grows. The scarcity of resources inhibits the positive population growth in the forecast period, so it can be assumed that the growth efficiency of this resource decreases due to its further increase. In addition, there is no data that contradicts the function homogeneity hypothesis. Therefore, there is reason to believe that the function satisfies neoclassical criteria.

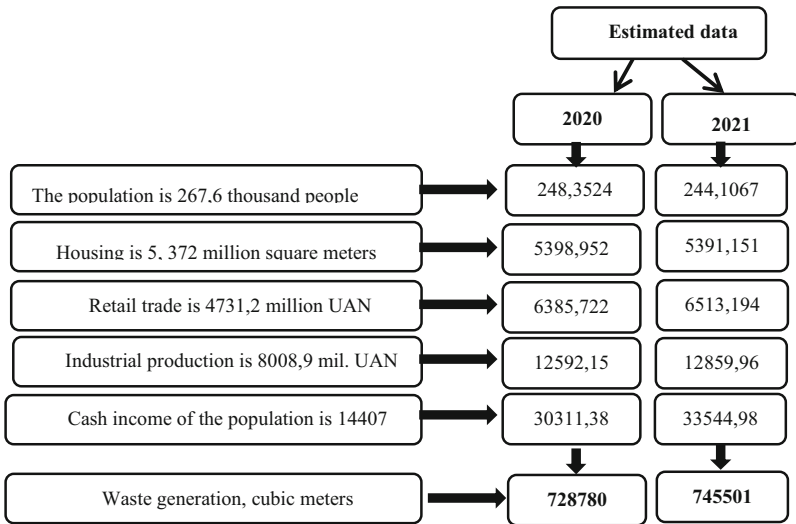


Fig. 3. Forecasting of the volume of SMW production in Zhytomyr

Regarding the possibility of replacing arguments of the waste function, a relatively stable level of technological exchange between factors is expected in the forecast period. The steady state condition roughly corresponds to the proportionality of the marginal and average productivity of the factors. For its part, this means that the elasticity of factor replacement is unity. Since the volume of SMW formation is influenced by heterogeneous factors, the least-squares function is chosen as the basis of the mathematical model.

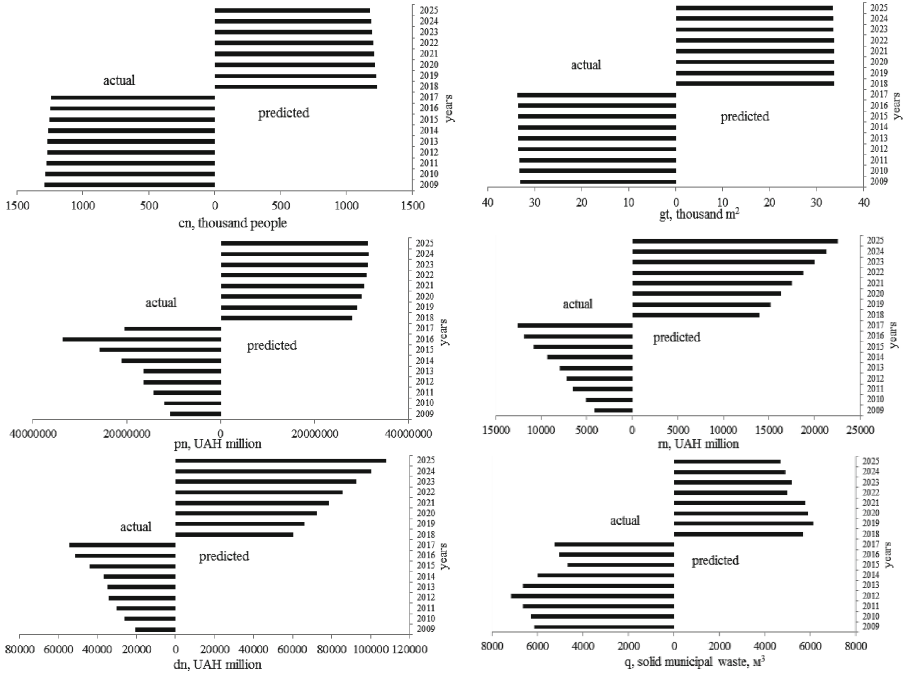


Fig. 4. Data of waste generation forecasting

The main disadvantages of an existing waste management organization are volume of waste which has to remove is determined on the basis of theoretical norms of formation, which lead to overestimation of funds for export; as payment for SMW export services is carried out by volume. Most cars are filled to half that reduces the quality of sanitation and increases cost. The lack of centralized control over the TKW removal process contributes to an increase in unauthorized landfills, which in turn leads to a rise in the cost of maintaining sanitary and environmental standards.

Data of actual, estimated and projected volumes of SMW from 2006 to 2025 years as the main link of the waste facility are indicate an increase in SMW volumes in Zhytomyr region compared to 2015 year. The situation might change with the use of modern methods of waste management. The projected amount of waste can be reduced by 2025.

Systemic dynamics methods were used to estimate the volume of urban solid municipal waste. Parameters for construction of a simulation model of waste production presented in Table 1. We will carry out simulation modeling by means of AnyLogic 7 environment. For developing a simulation model were used regression Eqs. (1)–(6). Figure 6 shows the projected population and garbage collection by 2025 year. Simulation modeling of waste generation was presented in Fig. 5. These two parameters are quite close in their values, so in 2014 the population amounted to 270.9 thousand people, while the projected one – 270.38 thousand people. This indicates that the model is 95% accurate. A decrease in population was noted - by the end of the simulation up to 244.106 thousand people.

Table 1. Parameters for construction of a simulation model of waste production

Parameter name	Value	Parameter name	Value
<i>natality</i>	0.002/year	<i>density_garbage</i>	0.08
<i>mortality</i>	0.004/year	<i>volume_garbage</i>	26*365
<i>density_glass</i>	1	<i>density_plastic</i>	0.07
<i>volume_glass</i>	720	<i>volume_plastic</i>	254*365
<i>density_metal</i>	0.36	<i>density_organics</i>	0.43
<i>volume_metal</i>	14*365	<i>volume_organics</i>	122*365
<i>density_cardboard</i>	0.08	<i>density_other</i>	0.27
<i>volume_cardboard</i>	74*365	<i>volume_other</i>	52*365
<i>Production_waste_per_person</i>			0.275 t/ year * people

Equation to describe population:

$$\begin{aligned} \text{population}(t) &= \text{population}(t - \Delta t) + \text{flow_population} * \Delta t, \text{ where} \\ \text{flow_population} &= \text{natality}(t - \Delta t) - \text{mortality}(t - \Delta t) \end{aligned} \quad (7)$$

To run the application, you must specify the initial conditions. Choose the population (0) = 271,000 people.

Equation to describe total waste:

$$\begin{aligned} \text{total waste}(t) &= \text{total waste}(t - \Delta t) + \text{flow_waste} * \Delta t, \text{ where} \\ \text{flow_waste} &= \text{population}(t - \Delta t) * \text{waste generation per person} \end{aligned} \quad (8)$$

Equation to describe the dynamics of constituent wastes, namely glass:

$$\begin{aligned} \text{glass}(t) &= \text{glass}(t - \Delta t) + \text{flow_glass} * \Delta t, \text{ where} \\ \text{flow_glass} &= \text{thick glass}(t - \Delta t) * \text{glass volume} \end{aligned} \quad (9)$$

Equation to describe the dynamics of metal waste:

$$\begin{aligned} \text{metal}(t) &= \text{metal}(t - \Delta t) + \text{flow_metal} * \Delta t, \text{ where} \\ \text{flow_metal} &= \text{density_metal} * \text{volume_metal} \end{aligned} \quad (10)$$

Equation to describe the dynamics of cardboard waste:

$$\begin{aligned} \text{cardboard}(t) &= \text{cardboard}(t - \Delta t) + \text{flow_cardboard} * \Delta t, \text{ where} \\ \text{flow_cardboard} &= \text{density_cardboard} * \text{volume_cardboard} \end{aligned} \quad (11)$$

Equation to describe waste paper dynamics:

$$\begin{aligned} \text{waste paper}(t) &= \text{waste paper}(t - \Delta t) + \text{flow_waste paper} * \Delta t, \text{ where} \\ \text{flow_waste paper} &= \text{density_waste} * \text{volume_waste} \end{aligned} \quad (12)$$

Equation to describe the dynamics of plastic waste:

$$\begin{aligned} \text{plastic}(t) &= \text{plastic}(t - \Delta t) + \text{flow_plastic} * \Delta t, \text{ where} \\ \text{flow_plastic} &= \text{thick_plastic} * \text{volume_plastic} \end{aligned} \quad (13)$$

Equations to describe the dynamics of organic waste :
 $organics(t) = organics(t - \Delta t) + flow_organics * \Delta t$, where (14)
 $flow_organics = organics_taste * organics_volume$

Equation to describe waste dynamics of the category "other waste" :
 $other_waste(t) = other_waste(t - \Delta t) + flow_other * \Delta t$, where (15)
 $flow_other = taste_other * volume_other$

A dynamic variable is introduced into the model :
 $General_waste = glass + metal + other\ waste + plastic + organic + waste\ paper + cardboard$ (16)
 $+ cardboard$

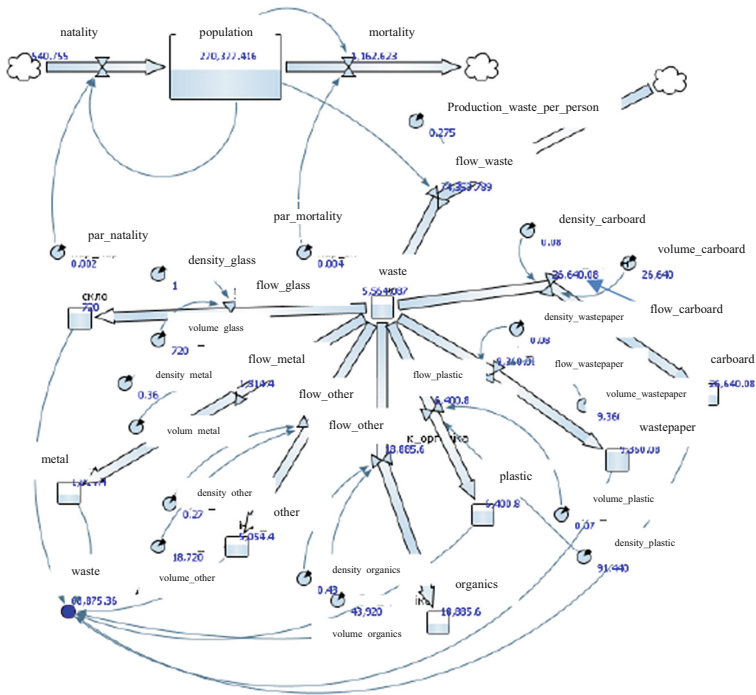


Fig. 5. Simulation modeling of waste generation

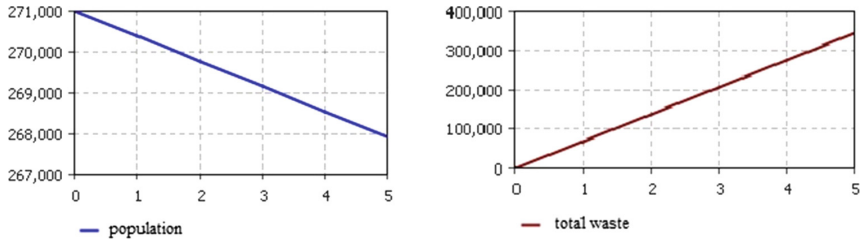


Fig. 6. Population dynamics and waste production over five years

Dependency graphs are shown in Fig. 6.

The total amount of waste generated is calculated by multiplying the rate of waste generation (kg/ day per capita) and the total population. The model demonstrates the total amount of waste generated by the population. The amount of waste generated is the amount accumulated, that is, the amount of waste generated in a given year is added to the amount of waste generated in the previous year.

Table 1 and Fig. 8 show the population at the annual baseline and representative indicative values of all waste fractions considered that may be generated in the year of assessment. It is expected that tonnage will continue to grow with decreasing population and changing socio-economic conditions. In the simulated study, the population was 270377, producing 18886 kg of food waste, 720 kg of glass waste (as of 2014 year). Therefore, glass waste in 2021 year will be 360 kg, and food waste - 94428 kg. The model takes into account different types of waste (cardboard, inert glass (less than 10 mm in diameter), metal, organic, paper, plastic and others). In addition, the model also shows the amount of each type of waste generated. The model considers the possibility of accepting a particular type of processing, depending on the economic region, where there are opportunities for different types of processing. By using the model from the point of view of waste treatment, considerable savings can be achieved (Fig. 8).



Fig. 7. Morphological composition of waste, %

It can be concluded that the morphological composition of SMW has significantly changed in the following components: increased organic component, plastic; the amount of waste paper (paper, cardboard) decreased (Fig. 7).

The system dynamics model allows to visualize the picture of solid waste generation for cities such as Zhytomyr, in comparison with traditional approaches. The model can be used as a decision support tool to predict the future situation and the amount of different types of waste generated in the Zhytomyr region.

The regional approach rather than a local approach is rational when organizing and managing SMW circulation processes. The city can be considered as a regional system, which has a multilevel, hierarchical structure, the elementary unit of which is a settlement, where mainly all processes of life, including waste, are concentrated.

The primary task is to predict the volume of solid municipal waste accumulation, mathematical modeling and study of the morphological composition of waste due to changes in environmental and social factors that affect on the morphological composition of waste. This problem is solved by the proposed system dynamics model.

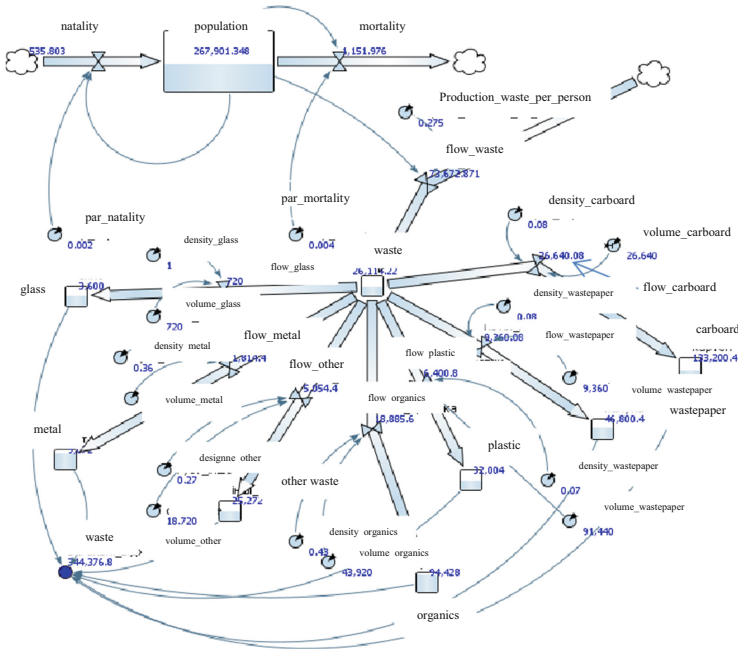


Fig. 8. Simulation modeling and forecasting for a five-year waste production period

3 Conclusions

The study was carried out to evaluate the amount and composition of SMW and to determine the potential for waste treatment. The simulation model that underlies the information subsystem of management support for solid municipal waste management

has been developed. The system dynamics model was used to estimate the average annual amount of SMW. Simulation was performed using the AnyLogic 7 environment. To analyze the sensitivity of the model, it is necessary to consider the main factors of influence: population, waste components.

For developing a simulation model were used regression equations: the dynamics of population change, the dynamics of changes in the generated waste, the dynamics of changes in the volume of housing stock, the dynamics of changes indicators of retail trade and public catering, the dynamics of changes industrial production, the dynamics of changes household income.



The used amount depends on the produced and recycled wastes, while the compacted density is constant. This means that if the amount of collected wastes increases as a result of population decline, the volume of use will increase. The model allows to carry out the forecast of volumes of the generated waste for decision-making in system of regional management.

References

1. Laznenko, D.: Vyznachennya parametriv utvorenniya komunal'nykh vidkhodiv u naselenykh punktakh Ukrainy dlya tsiley rehional'noho planuvannya. Kyiv (2019)
2. Morozova, T., Lukianova, V., Anpilova, Y.: Conceptualization of latent ecosystem services. *Ecol. Saf. Nat. Resour.* **1**(29), 54–64 (2019)
3. Adamović, V.M., Antanasijević, D.Z., Ristić, M.Đ., Perić-Grujić, A.A., Pocajt, V.V.: Prediction of municipal solid waste generation using artificial neural network approach enhanced by structural break analysis. *Environ. Sci. Pollut. Res.* **24**(1), 299–311 (2016). <https://doi.org/10.1007/s11356-016-7767-x>
4. Ghinea, C., Niculina, D.E., Comăniță, E., Gavrilescu, M., Câmpean, T.: Hybrid model for the prediction of municipal solid waste generation in Hangzhou. *China Waste Manag. Res.* **37**(8), 781–792 (2019). <https://doi.org/10.1177/0734242X19855434>
5. Anfilatov, V.S., Emelyanov, A.A., Kukushkin, A.L.: System analysis in management of Moscow “Finance and Statistics”, p. 368 (2002)
6. Khrutba, V.O.: Fundamentals of project management and waste management in the road complex. Monograph, Kiev, p. 192 (2013)
7. Odrin, V.M., Kartavov, S.S.: Morphological analysis of systems. Building a morphological matrix, Kiev, p. 183 (1977)
8. Main Department of Statistics in Zhytomyr Oblast. <http://www.zt.ukrstat.gov.ua/>
9. Curteanu, S., Gavrilescu, M.: Forecasting municipal solid waste generation using prognostic tools and regression analysis. *J. Environ. Manag.* **1**(182), 80–93 (2016). <https://doi.org/10.1016/j.jenvman.2016.07.026>



Modelling Radionuclide Scavenging in the Ocean by a Particle Tracking in Multicomponent Medium with First-Order Reaction Kinetics

Igor Brovchenko^(✉)  and Vladimir Maderich 

Institute of Mathematical Machine and System Problems, Glushkova Av., 42,
03187 Kiev, Ukraine

ibrovchenko@gmail.com, vladmad@gmail.com

Abstract. A new particle tracking method for modelling of pollutant propagation in multicomponent medium with first-order reaction kinetics and spatially varying model parameters was developed using method of moments. The proposed particle tracking method has weaker time step limit in comparison with standard method. This method was applied to simulate vertical transport by sinking particulate matter (scavenging) of the artificial radionuclides introduced to the ocean surface in result of the fallout from the nuclear weapon testing. The method was applied to simulate scavenging of plutonium $^{239,240}\text{Pu}$ in the northern Pacific Ocean. The results of simulation of plutonium $^{239,240}\text{Pu}$ scavenging in the northern Pacific Ocean agreed with measurements in period 1972–1997. We found that dispersion of reactive contamination caused by two-way phase transition can be much greater than vertical diffusion in the ocean. It was showed importance in the scavenging process of large fast sinking particles leaching contaminated layer of the ocean and process of dissolution with depth of the mainly organic sinking particulate matter.

Keywords: Particle tracking · Reaction kinetics · Method of moments · Radionuclide scavenging · Plutonium $^{239,240}\text{Pu}$

1 Introduction

A particle tracking methods originated from statistical physics [1] are widely used in the modelling of pollutant transport in environment including ocean (e.g. [2–6]) along with many other applications. In these methods the released pollutant amount is divided between a large number of particles tracked along the path. Unlike Eulerian numerical methods, particle tracking Lagrangian methods are conservative, they are free from numerical dispersion, artificial smoothing and oscillations. The particle tracking is especially computationally effective when pollution occupies a small part of the computational domain. First order reactions can be introduced in particle tracking algorithm assuming that at given time step, the particle changes phase state with some probability depending on the kinetic coefficients [2, 7]. However, this time step should be sufficiently small to guarantee that the probability of having more than one phase

change is negligible [7]. Therefore, a numerically efficient particle tracking method is needed to investigate the behavior of pollutants in multicomponent medium with first-order reaction kinetics.

In the paper, the particle tracking method for modelling of pollutant propagation in multicomponent medium with first-order reaction kinetics and spatially varying model parameters was developed using method of moments [8–10]. This method was applied to simulate vertical transport of the artificial radionuclides introduced to the ocean surface due the fallout from nuclear weapon testing. The profiles of several radionuclides show presence of maximum of concentration moving downward with time [11]. The reversible sorption process on falling particulate matter is effective removal mechanism (scavenging) for reactive radionuclides. An idealized one-dimensional problem was solved using particle tracking in multicomponent medium with first-order reaction kinetics.

The paper is organized as follows. Method of moments for particle tracking in the medium with first-order kinetics is briefly described in Sect. 2. The scavenging model is built in Sect. 3. In Sect. 4 the scavenging model and method of moments were applied for both idealized case of instantaneous release of radionuclide and realistic case of downward propagation of $^{239,240}\text{Pu}$ from fallout on the surface of northern Pacific. The results of method development and application are summarized in Conclusions.

2 Method of Moments for Particle Tracking in Medium with First-Order Kinetics

Consider for simplicity one-dimensional advection-diffusion equation for tracer with first-order reaction kinetics:

$$\frac{\partial C_i}{\partial t} = -\frac{\partial u_i C_i}{\partial x} + \frac{\partial}{\partial x} D_i \frac{\partial C_i}{\partial x} + \sum_{j=1}^n \alpha_{i,j} C_j, \quad i = \overline{1, n}, \quad (1)$$

where t is time, x is spatial coordinate, C_i is concentration of tracer in i -th state, n is number of states, u_i advection velocity in the state i , D_i is diffusivity in state i , $\alpha_{i,j}$ are kinetic coefficients of 1st order reaction terms, and n is a number of states. Representing tracer as an ensemble of equal particles, the Eq. (1) can be written in form of the Fokker-Planck equation for particle location probability density function $p_i(x, t)$ in state i :

$$\frac{\partial p_i}{\partial t} = -\frac{\partial}{\partial x} (u_i^{drift} p_i) + \frac{\partial^2 D_i p_i}{\partial x^2} + \sum_{j=1}^n \alpha_{i,j} p_j, \quad (2)$$

where $u_i^{drift} = u_i + \partial D_i / \partial x$. The particle located at position $x_i = 0$ is considered as infinitesimal plume of probability. Further evolution of $p_i(x, t)$ can be described using several first moments of order k :

$$M_i^{(k)}(t) = \int_{-\infty}^{\infty} x^k p_i(x, t) dx, \quad (3)$$

assuming that p_i and its derivatives tend to 0 at $x \rightarrow \pm\infty$. For the coefficients of Eq. (2) smoothly varying in space, we restrict ourselves to the first terms of the Taylor series expansions in the neighborhood of the local point

$$D_i = D_i^{(0)} + xD_i'; \quad \alpha_{ij} = \alpha_{ij}^{(0)} + x\alpha_{ij}'. \quad (4)$$

Multiplying on x^k and integrating (2) by parts after some simplifications yields system of equations for first three moments:

$$\frac{dM_i^{(0)}(t)}{dt} = \sum_{j=1}^n \alpha_{i,j}(\bar{M}_j^{(1)}) M_i^{(0)}, \quad (5)$$

$$\frac{d\bar{M}_i}{dt} = u_i^{drift} + \frac{1}{M_i^{(0)}} \sum_{j=1}^n \alpha_{i,j}(\bar{M}_j) M_i^{(0)} (\bar{M}_j - \bar{M}_i), \quad (6)$$

$$\frac{d\sigma_i^2}{dt} = 2D_i(\bar{M}_i) + \frac{1}{M_i^{(0)}} \sum_{j=1}^n \alpha_{i,j}(\bar{M}_j) M_i^{(0)} [\sigma_j^2 - \sigma_i^2 + (\bar{M}_j - \bar{M}_i)^2], \quad (7)$$

where \bar{M}_i is a normalized 1st moment, and σ_i^2 is a dispersion defined as

$$\bar{M}_i = \frac{M_i^{(1)}}{M_i^{(0)}}, \quad \alpha_i^2 = \frac{M_i^{(2)}}{M_i^{(0)}} - (\bar{M}_i)^2. \quad (8)$$

The Eq. (5) for zero-moment $M_i^{(0)}(t) \equiv P_i(t)$ is master equation for probability of particle $P_i(t)$ to be in i -th state. This equation describes the Markov processes in systems which jump from one state to another in continuous time [1]. The Eqs. (6) and (7) describe displacement of centre of mass and dispersion of plume. Change of particle position in state i by deterministic displacement of centre of mass and stochastic displacement by uncorrelated random walk from t to $t + dt$ is governed by equation

$$x_i(t + dt) = x_i(t) + \bar{M}_i(dt) + R\sqrt{\sigma_i^2(dt)}, \quad (9)$$

where R is a normally distributed random variable with zero mean value and unit variance. An important feature of Eqs. (6)–(7), (9) is that displacement of particles depend also on reactions kinetics (last terms in (6)–(7)). In turn, change of centre of mass affect probability evolution in (5) when kinetic coefficients vary in the space. The system of ordinary differential Eqs. (5)–(7) can be solved numerically for each particle. Then new position of particle is determined using (9).

Using first order Euler time stepping for (5)–(7) it is easy to calculate probability to remain in state i and change state to j and find first moment, dispersion and particle displacement after one time step Δt :

$$P_i(\Delta t) = 1 - \alpha_{ij}\Delta t \text{ at } i = j, \quad P_i(\Delta t) = \alpha_{ij}\Delta t, \quad \text{at } i \neq j; \quad (10)$$

$$\bar{M}_i(\Delta t) = (u_i + \partial D_i / \partial x)\Delta t \text{ at } i = j, \quad M_i(\Delta t) = 0 \text{ at } i \neq j, \quad (11)$$

$$\sigma_i^2 = 2D_i\Delta t \text{ at } i = j, \text{ at } i \neq j \quad \sigma_j^2 = 0, \quad (12)$$

$$x_i(t + \Delta t) = x_i(t) + (u_i + \partial D_i / \partial x)\Delta t + R\sqrt{2D_i\Delta t}, \quad (13)$$

The particle state, displacement of centre of mass, dispersion and new position are calculated sequentially. Notice that (13) is a standard numerical scheme of Euler-Maruyama of order $O(\Delta t^{1/2})$ in strong sense and $O(\Delta t)$ in weak sense [12]. The disadvantage of this method is limitation on time step $\alpha_{ij}\Delta t \ll 1$ and limitation on corresponding small values of the transition probabilities, which guarantees only a single phase change during each time step [7].

3 Scavenging Model

Consider an idealized one-dimensional problem for multicomponent medium with first-order reaction kinetics representing processes of sorption-desorption of radionuclide in the ocean water column. Following [13], the general Eq. (1) can be written for concentration $C_1 \equiv C_d$ (Bq m⁻³) in dissolved state ($i = 1$) and concentrations $C_i \equiv C_{p,i}$ (Bq m⁻³) in particulate state for fractions $i = 2, \dots, n$ as:

$$\frac{\partial C_d}{\partial t} = -\frac{\partial u_d C_d}{\partial x} + \frac{\partial}{\partial x} D \frac{\partial C_d}{\partial x} - \lambda C_d - a_{ds} \left(C_d^w \sum_{i=1}^n S_{p,i} K_{d,i} - C_p \right), \quad (14)$$

$$\frac{\partial C_{p,i}}{\partial t} = -\frac{\partial (u_d + u_{p,i}) C_{p,i}}{\partial x} + a_{ds} (C_d S_{p,i} K_{d,i} - C_{p,i}) - \lambda C_{p,i} + \frac{\partial}{\partial x} D \frac{\partial C_{p,i}}{\partial x}, \quad (15)$$

where x (m) is vertical coordinate directed downward from ocean surface $x = 0$, vertical advection velocities $u_1 \equiv u_d$ (m s⁻¹) in state $i = 1$ and $u_i \equiv u_{p,i}$ (m s⁻¹) in state $i = 2, \dots, n$, diffusivity $D \equiv D_i$ (m² s⁻¹), S_i (kg m⁻³) is concentration of particulate matter in state $i = 2, \dots, n$, the distribution coefficient $K_{d,i}$ (m³ kg⁻¹) in state $i = 2, \dots, n$ is related with equilibrium values of C_d and $C_{p,i}$ by the ratio

$$S_i K_{d,i} = \frac{C_{p,i}}{C_d}. \quad (16)$$

The total concentration of radionuclide fractions in particulate phase C_p is

$$C_p = \sum_{i=2}^n C_{p,i}. \quad (17)$$

The equations similar to (14)–(15) for single particulate fraction ($i = 1, 2$) were considered by [3, 14]. To obtain single equation for C_p assume that distribution on fractions of particulate matter is close to the equilibrium distribution, i.e.

$$C_{p,i} \approx \frac{S_i K_{d,i}}{\sum_{i=2}^n S_i K_{d,i}} C_p. \quad (18)$$

Then an equation for the total concentration of radionuclide in the particulate phase C_p^w is written as

$$\begin{aligned} \frac{\partial C_p}{\partial t} = & -u_d \frac{\partial C_p}{\partial x} - \sum_{i=2}^n u_{p,i} \frac{\partial}{\partial x} \left(\frac{C_p S_i K_{d,i}}{\sum_{i=2}^n S_{p,i} K_{d,i}} \right) - \lambda C_p + \\ & a_{ds} \left(C_d \sum_{i=1}^n S_i K_{d,i} - C_p \right) + \frac{\partial}{\partial x} v_T \frac{\partial C_p}{\partial x}. \end{aligned} \quad (19)$$

The total concentration C_e for low concentration of particulate matter is approximately

$$C_e(z, t) = C_d + C_p. \quad (20)$$

Then using (5) and (7) yields

$$C_d = \frac{1}{1 + \sum_{i=2}^n S_i K_{d,i}} C_e, \quad C_p = \frac{\sum_{i=1}^n S_i K_{d,i}}{1 + \sum_{i=2}^n S_i K_{d,i}} C_e. \quad (21)$$

A single equation for C_e was written using (1), (6) and (8) as

$$\frac{\partial C_e}{\partial t} = -u_d \frac{\partial C_e}{\partial x} - \sum_{i=2}^n u_{p,i} \frac{\partial}{\partial x} \left(\frac{S_i K_{d,i} C_e}{1 + \sum_{i=2}^n S_i K_{d,i}} \right) - \lambda C_e + \frac{\partial}{\partial x} D \frac{\partial C_e}{\partial x}. \quad (22)$$

This quasi-equilibrium equation was used in several scavenging models (e.g. [15]) for one-fraction particulate component.

4 Results

4.1 Idealized Case of Scavenging Due to Instantaneous Release of Radionuclide

In this subsection an idealized case of instantaneous release of radionuclide on the ocean surface is considered using particle tracking to highlight important transport and dispersion mechanisms. Only two state of radionuclide are considered: dissolved state

($i = 1$) and sorbed on single fraction of particulate matter state ($i = 2$). Initial instantaneous release of radioactivity at the ocean surface was $I_0 = 1 \text{ Bq m}^{-2}$. The ocean depth was 6000 m, $u_d = 0$. The concentration of particulate matter is $S \equiv S_2 = 0.25 \cdot 10^{-3} \text{ kg m}^{-3}$ and $u_{p2} = u_p = 5 \cdot 10^{-5} \text{ m s}^{-1}$. According [16], the distribution coefficient $K_d = K_{d,2}$ for $^{239,240}\text{Pu}$ is $K_d = 100 \text{ m}^3 \text{ kg}^{-1}$. The decay constant $\lambda = 3.34 \cdot 10^{-12} \text{ s}^{-1}$ is small making possible decay to be ignored at scale of several decades. Two scenarios were considered: (A) The diffusivity is $D = 10^{-5} \text{ m}^2 \text{ s}^{-1}$, and (B) $D = 0$.

A self-similar solution of Eq. (22) with abovementioned assumptions is

$$C_e(x, t) = \frac{I_0}{\sqrt{4\pi Dt}} \exp \left[-\frac{(x - u_e t)^2}{4Dt} \right], \quad (23)$$

where $u_e = u_p SK_d / (1 + SK_d)$. This solution was compared with numerical solution of particle tracking model Eqs. (5)–(7), (9) at $a_{ds} = 1.16 \cdot 10^{-5} \text{ s}^{-1}$ [17]. The Eqs. (5)–(7) were solved by 4th order Runge-Kutta method with time step $\Delta t = 10^5 \text{ s}$.

The results of comparison of concentration in dissolved state between scenario A and analytical solution (23) are given in Fig. 1a for different times. In both cases an initial injection of activity was transformed in symmetric bulge moving downward with velocity u_e . The scavenging by settling particulate matter resulted in formation of similar bulge of concentration on particles which transported down and released activity into water due to the desorption. However, in the non-equilibrium case spreading of concentration bulge is greater than in the quasi-equilibrium case (23). The comparison of results of simulation for scenarios A and B is shown in Fig. 1b. As seen in the figure, spreading of bell-shaped concentration profile occurs for both scenarios demonstrating importance of phase change in dispersion of concentration profile.

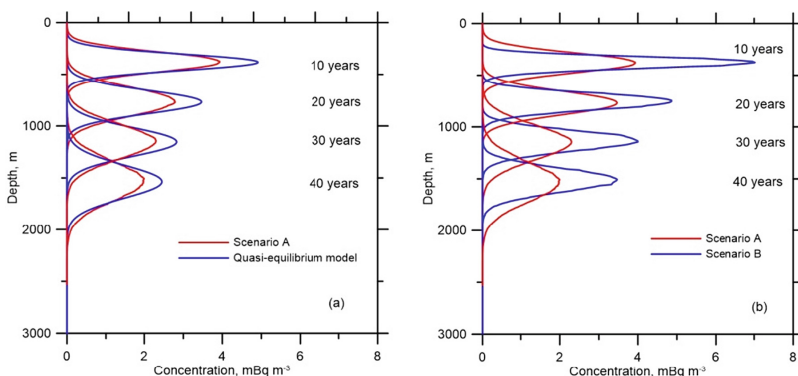


Fig. 1. Comparison of analytical solution (23) and simulation according scenario A (a) and comparison of scenarios A and B (b) in the idealized case of instantaneous release of radionuclides.

Consider this mechanism in more detail using particle tracking model (3), (5)–(7), (9). In the two state model the solution of master Eq. (3) for probability of particle to be in states $i = 1; 2$ is

$$\begin{aligned} P_1(t) &= \frac{k_2}{k_1 + k_2} \left(\frac{k_1 P_1(0) - k_2 P_2(0)}{k_2} e^{-(k_1 + k_2)t} + 1 \right), \\ P_2(t) &= \frac{k_1}{k_1 + k_2} \left(\frac{k_2 P_2(0) - k_1 P_1(0)}{k_1} e^{-(k_1 + k_2)t} + 1 \right), \end{aligned} \quad (24)$$

where $k_1 = a_{ds} S K_d$, $k_2 = a_{ds}$. At large times these probabilities tend to the quasi-equilibrium state:

$$P_1(t) = \frac{k_2}{k_1 + k_2}, \quad P_2(t) = \frac{k_1}{k_1 + k_2}. \quad (25)$$

The corresponding solutions of (6) are

$$\begin{aligned} \bar{M}_1(t) &= \frac{u_p k_1}{k_1 + k_2} t + \frac{u_p k_1}{(k_1 + k_2)^2} \left(e^{-(k_1 + k_2)t} - 1 \right), \\ \bar{M}_2(t) &= \frac{u_p k_1}{k_1 + k_2} t + \frac{u_p k_2}{(k_1 + k_2)^2} \left(1 - e^{-(k_1 + k_2)t} \right). \end{aligned} \quad (26)$$

The shift in position of centroids in different states at large time is

$$M_1(t) - M_2(t) = \frac{u_p k_1}{k_1 + k_2}. \quad (27)$$

The dispersions in states 1 and 2 are

$$\sigma_1^e(t) = \sigma_2^e(t) = 2 \left(D + \frac{u_p^2 k_1 k_2}{(k_1 + k_2)^3} \right) t. \quad (28)$$

The derived expressions show that contamination in dissolved and particulate states propagate with same velocity. As follows from (27) position of centre of mass of particulate component propagates at constant distance ahead of dissolved state position of centre of mass. The dispersion increases linearly with time. Dispersion (28) includes diffusivity and second term $u_p^2 S_p K_d / a_{ds} (S_p K_d + 1)^3$ caused by phase transfer. This term is responsible for dispersion of concentration profile for scenario B in Fig. 1b. With increase of u_p , S_p , K_d this second term dominates whereas increase of desorption rate a_{ds} results decrease of dispersion caused phase transfer.

The comparison of simulation results for Scenario B using method of moments and the Euler method is shown in Fig. 2. As seen in figure, the Euler method can reproduce method of moments results for profile of contamination and dispersion only at time step in 10 time less than time step in the method of moments. This conclusion confirm time step limit for Euler method. At $\Delta t = 0.5$ d parameter $a_{ds}\Delta t \approx 0.48$, whereas $a_{ds}\Delta t \approx 0.048$ at $\Delta t = 0.05$ d.

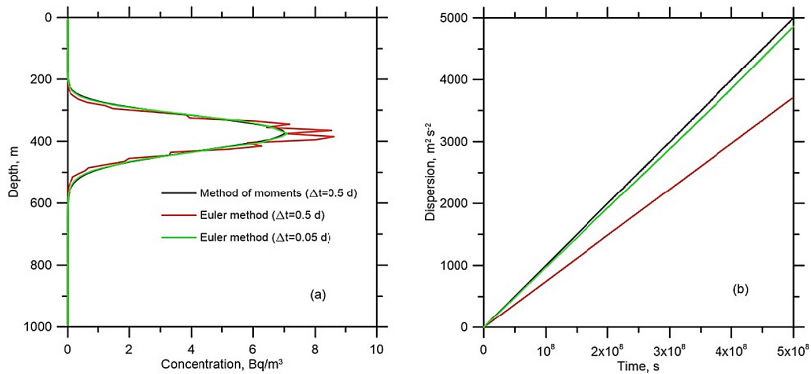


Fig. 2. Comparison of profiles of $^{239,240}\text{Pu}$ (a) and dispersions (b) computed by method of moments and first order numerical models in the idealized case of instantaneous release.

4.2 $^{239,240}\text{Pu}$ Scavenging in the North Pacific

The idealized case considered in previous section showed that instantaneous release of reactive radionuclides in the ocean with single class of suspended sediment result in formation of bulges of concentration of contamination in dissolved and particulate states moving downward with constant velocity. However, observations of $^{239,240}\text{Pu}$ scavenging in the ocean [11, 18] demonstrated more complicated behavior. This is caused by several reasons, among which: (i) Deposition of $^{239,240}\text{Pu}$ is a time varying process following to series of nuclear weapon testing (Fig. 3a); (ii) The particulate matter in the ocean proper is mostly organic origin with wide range of sizes, settling velocities and decay rates of particulate matter with depth due to dissolution [11]; (iii) The chemical forms of the plutonium change in water column implying the distribution coefficient variation with depth; (iv) Diffusivity in the oceanic thermocline and abyssal waters depend on the stratification; (v) The vertical velocity in the ocean varies with depth and can changes sign in anticyclonic gyres. Therefore, it is necessary to include some of abovementioned factors into consideration for scavenging simulation in the northwestern Pacific.

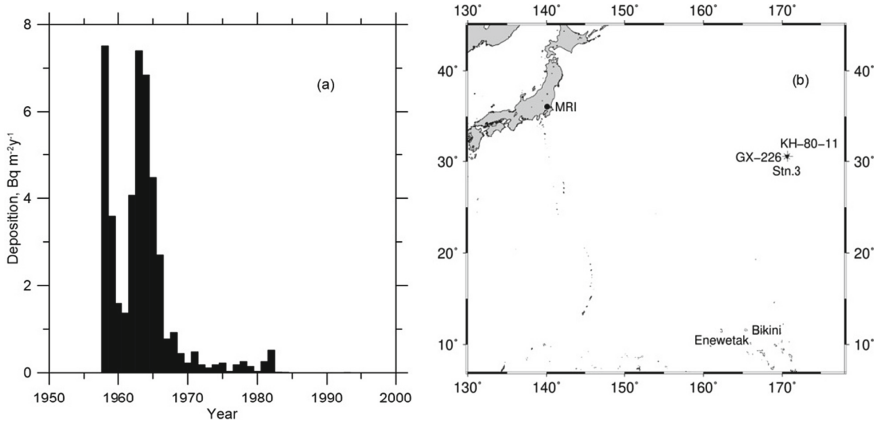


Fig. 3. The annual deposition of $^{239,240}\text{Pu}$ observed at the Meteorological Research Institute (MRI) of Japan [15] (a); Sampling stations in the northwestern Pacific Ocean which were used to compare with simulation results (b).

Following [3, 20], whole range of falling particles was divided on two classes: fine slow sinking particulate matter ($u_{p2} = 5 \cdot 10^{-5} \text{ m s}^{-1}$) distributed with depth following [15] as

$$S_2 = S_{20} \cdot 10^{-(z/z_0)}, \quad (29)$$

where $S_{20} = 0.25 \cdot 10^{-3} \text{ kg m}^{-3}$, $z_0 = 4 \cdot 10^3 \text{ m}$ and large fast sinking particles ($S_3 = 0.25 \cdot 10^{-5} \text{ kg m}^{-3}$, $u_{p2} = 5 \cdot 10^{-3} \text{ m s}^{-1}$). Diffusivity and distribution coefficient K_d were assumed constant ($D = 10^{-5} \text{ m}^2 \text{ s}^{-1}$, $K_d = 100 \text{ m}^3 \text{ kg}^{-1}$). The annual deposition of $^{239,240}\text{Pu}$ observed MRI [15] was used as source term in the modelling (see Fig. 3a). As seen in the figure, first pulse of activity was delivered by troposphere fallout due to the weapon testing at Marshall Islands with maximum in 1958. Second pulse was caused by global stratospheric fallout from atmospheric testing of nuclear weapon with maximum in 1963. The results of simulations carried out in location 31°N , 148°E (Fig. 3b) were compared with measurements performed in 1972 (Station GX-226), 1980 (Station KH-80-11) and 1997 (Station 3) [19]. The results of comparison of profiles of $^{239,240}\text{Pu}$ concentration in the dissolved and particulate states are given in Fig. 4. In general, the simulation results are in good agreement with the measurements, keeping in mind that the parameters of the model were estimated by few observations. Two deposition maxima in Fig. 3a merged on the vertical profile into one. The concentration of $^{239,240}\text{Pu}$ in the particulate matter does not exceed 2% of concentration in water, however sinking particulate matter play crucial role in vertical transport and dispersion of reactive radionuclide. Despite the relatively low concentration of large particles, the fall of these particles leads to leaching of activity from the upper layers of the ocean and the formation of an asymmetric concentration profile, while in the case of one class of particles the vertical distribution of activity remains symmetric (Fig. 1).

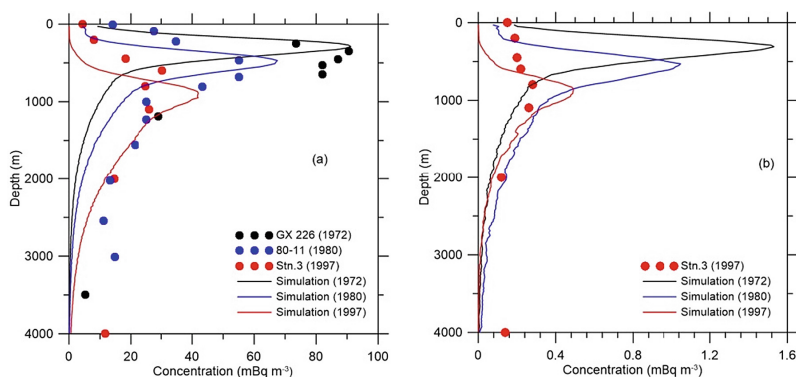


Fig. 4. The comparison between the vertical distribution of the observed concentration in dissolved (a) and particulate state (b) of $^{239,240}\text{Pu}$ in location 31°N , 148°E and simulated concentrations in the years of 1972, 1980 and 1997.

5 Conclusions

The new particle tracking method for modelling of pollutant propagation in multi-component medium with first-order reaction kinetics and spatially varying model parameters was developed using method of moments. The developed particle tracking method is free on time step limitation necessary for standard method. This method was applied to simulate scavenging of plutonium $^{239,240}\text{Pu}$ introduced to the ocean surface due to the fallout from the nuclear weapon testing. It was found that dispersion of reactive contamination caused by two-way phase transition that can be much greater than diffusion. The results of simulation for northwestern Pacific agreed with observation data. It was showed importance in scavenging process of large fast sinking particles leaching upper layer of the ocean and of dissolution with depth of the mainly organic sinking particulate matter.

Acknowledgments. This work was partially supported by IAEA CRP K41017 “Behavior and Effects of Natural and Anthropogenic Radionuclides in the Marine Environment and their use as Tracers for Oceanography Studies”.

References

1. Parzen, E.: Stochastic Processes. Holden-Day, San Francisco (1962)
2. Perri  ez, R., Elliot, A.J.: A particle-tracking method for simulating the dispersion of non-conservative radionuclides in coastal waters. *J. Environ. Radioact.* **58**, 13–33 (2002)
3. Nakano, M., Povinec, P.P.: Modelling the distribution of plutonium in the pacific ocean. *J. Environ. Radioact.* **69**, 85–106 (2003)
4. Maderich, V., Brovchenko, I., Jung, K.T.: Oil spreading in instantaneous and continuous spills on rotating Earth. *Environ. Fluid Mech.* **12**, 361–378 (2012)

5. Lynch, D.R., Greenberg, D.A., Bilgili, A., Mcgillicuddy, D.J., Manning, J.P., Aretxabaleta, A.L.: *Particles in the Coastal Ocean Theory and Applications*. Cambridge University Press, Cambridge (2015)
6. Periañez, R., Bezhenar, R., Brovchenko, I., Duffa, C., Iosjpe, M., Jung, K.T., Kim, K.O., Kobayashi, T., Liptak, L., Little, A., Maderich, V., McGinnity, P., Min, B.I., Nies, H., Osvath, I., Suh, K.S., de With, G.: Marine radionuclide transport modelling: recent developments, problems and challenges. *Environ. Model. Softw.* **122**, 104523 (2019)
7. Kinzelbach, W.: The random-walk-method in pollutant transport simulation. In: Custodio, E., et al. (ed.) *Advances in Analytical and Numerical Groundwater Flow and Quality Modelling*, pp. 227–246. D. Reidel, Norwell (1987)
8. Aris, R.: On the dispersion of a solute in a fluid flowing through a tube. *Proc. Roy. Soc. Lond. Ser. A* **235**, 67–78 (1956)
9. Hunter, J., Craig, P., Phillips, H.: On the use of random walk models with spatially variable diffusivity. *J. Comput. Phys.* **106**, 366–376 (1993)
10. Henri, C.V., Fernandez-Garcia, D.: Toward efficiency in heterogeneous multispecies reactive transport modeling: a particle-tracking solution for first-order network reactions. *Water Resour. Res.* **50**, 7206–7230 (2014)
11. Livingston, H.D., Povinec, P.P., Ito, T., Togawa, O.: The behavior of plutonium in the Pacific Ocean. In: Kudo, A. (ed.) *Plutonium in the Environment, Radioactivity in the Environment*, pp. 267–292. Elsevier, Amsterdam (2001)
12. Kloeden, P.E., Platen, E.: *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin (1992)
13. Maderich, V., Jung, K.T., Brovchenko, I., Kim, K.O.: Migration of radioactivity in multi-fraction sediments. *Environ. Fluid Mech.* **17**(6), 1207–1231 (2017)
14. Periañez, R.: Modelling the distribution of radionuclides in deep ocean water columns. Application to H-3, Cs-137 and Pu-239, Pu-240. *J. Environ. Radioact.* **38**, 173–194 (1998)
15. Tsumune, D., Aoyama, M., Hirose, K.: Numerical simulation of ^{137}Cs and $^{239,240}\text{Pu}$ concentrations by an ocean general circulation model. *J. Environ. Radioact.* **69**, 61–84 (2003)
16. IAEA (International Atomic Energy Agency): *Sediment distribution coefficients and concentration factors for biota in the marine environment*. Vienna, IAEA (2004)
17. Periañez, R., Brovchenko, I., Jung, K.T., Kim, K.O., Maderich, V.: The marine k_d and water/sediment interaction problem. *J. Environ. Radioact.* **129**, 635–647 (2018)
18. Hirose, K., Aoyama, M., Povinec, P.P.: Concentration of particulate plutonium in surface and deep-water samples collected during the IAEA1997 expedition. *Deep-Sea Res. II* **50**, 2639–2647 (2003)
19. MARiS (Marine Information System): *Radioactivity and stable isotope data in the marine environment*. <http://maris.iaea.org/>. Accessed 21 Mar 2020
20. Livingston, H.D., Anderson, R.F.: Large particle transport of plutonium and other fallout radionuclides to the deep ocean. *Nature* **303**, 228–231 (1983)



Information System of Ecological Monitoring “Small Mammals as Bioindicator”

Mariia Talakh¹ and Serhii Holub²

¹ Department of Computer Science, Yuriy Fedkovych Chernivtsi
National University, Chernivtsi, Ukraine
m.talah@chnu.edu.ua

² Department of Automated Systems Software,
Cherkasy State Technological University, Cherkasy, Ukraine

Abstract. The article presents the results of research on population dynamics *Martes martes* L. and *Lepus europaeus* Pall. depending on weather factors in the national park “Vyzhnytskyi” during the period 2002–2019 years. The effectiveness of the proposed method of multilevel monitoring for the studied bioindicator’s species was shown and the factors having a reliable influence on their population size were highlighted. For assessing the complex state of ecosystems often several bioindicator species are used. In this case, the trophic criterion was used as an approach for the formation of a complex bioindicator - the selected species belong to small mammals, while at the same time they are at different trophic levels and can be connected by food bonds. In this case, a complex bioindicator was obtained by the synthesis of models created for each of the bioindicators species. Modeling was carried out using a multi-row algorithm for group accounting of arguments, and a polynomial was used as a reference model. The proposed method allows to prioritize the use of individual bioindicators, depending on the modeling results. And proposed in this work set of indicators can be expanded by both bioindicators and climate characteristics.

Keywords: Monitoring · GMDH · Population · Modeling · *Martes martes* L. · Bioindicator · *Lepus europaeus* Pall

1 Introduction

Monitoring is a technology of providing information to decision-making processes. In the case of ecological monitoring, the requirements for this information technology are determined by what information is needed by the system to support the process of generating control impacts in managing the ecosystem [1].

In particular, the identification of population dynamics trends is necessary to predict possible undesirable phenomena and correction of their consequences. Study of population dynamics is important for environmental protection, as it allows predict the state of the ecological system, assessing the effects of environment factors on them, and is the basis for creating scientific bases for the rational use of natural resources. In particular, the use of the concept of environmental services of ecosystems is now relevant. These are the many and varied benefits to humans gifted by the natural

environment and from healthy ecosystems. Ecosystem services are grouped into four broad categories: provisioning, such as the production of food and water; regulating, such as the control of climate and disease; supporting, such as nutrient cycles and oxygen production; and cultural, such as spiritual and recreational benefits. Thus, one of the tasks for information technology of ecological monitoring is the search for organisms that can serve as indicators of the state of the environment and the factors that influence them [2]. This study focused on the use of bioindicators to assess the second group of environmental services, - climate parameter indicators. Small mammals provide important ecosystem services, and knowledge of their dispersal preferences is useful for population management and landscape planning. Small mammals can regulate insect populations, act as pollinators and seed dispersers, support forest regeneration, aerate soil and provide food for carnivores and humans. In addition, the transport of minerals from deeper layers of the soil to the surface can act as a source of nutrients for plants, and the burrowing activity of some rodent species influences surface water runoff. Small mammals are also disease reservoirs that can be detrimental to human health and they can also act as crop pests [3]. However, small mammals are a fairly diverse group of species, so selecting bioindicator species within it is not an easy task. When determining the state of ecosystems, it is logical to use several bioindicator species for complex assessment of the dynamics of environmental factors. In the biological sciences, the using of complex bioindicators is carried out through the use of various integral indices, mainly indices of biodiversity based on entropy coefficients or based on the study of bioindicators at different trophic levels or at different levels of life organization. In this case, we have also been limited to the list of species whose number will be monitored in the area National Park “Vizhnitsky”. As a solution to this problem, we suggest using of two species: *Lepus europaeus* Pall. and *Martes martes* L. The first of these species belongs to rodents, and the second to predatory species, moreover, these species are associated with food relations, although *Lepus europaeus* Pall. is not a staple food base for *Martes martes* L. [4, 5]. Although, small mammals have potential as indicators of sustainable forest management, which is particularly important for protected areas [6]. In this case, we have selected two species that can be conditionally assigned to one group – small mammals, which simultaneously represent different food groups and are associated with trophic links.

The aim of research is the construction of monitoring information systems (MIS) is an attempt to solve the problem of indicating the states of complex systems using available technical and algorithmic means. **The main purpose** of this work is to investigate the processes of ecological monitoring information technology to identify an interaction between weather factors and the populations’ dynamics of small mammal with different food strategies.

2 Analysis of Recent Research and Publications

The information obtained as a result of the monitoring it’s the data about the properties of the monitoring objects in the form of analytical models.

Statistical techniques are frequently used to perceive patterns of affiliation and forecasting. There are several works about the effect of exogenous elements of the

animal population, which include the instance of the investigated species. Based on [7] the evaluation of the space-time dynamics of the ecological niche of marten, the dependence of the nature of the placement of individuals on the weather conditions is shown. At the same time, the main function of such elements as snow cover height, temperature, and nature of biotopes are proven [5].

Other work additionally distinguishes a set of elements comparable to the preceding work, the definition of which is snow cover. Increasing it with elevating the temperature stage will increase marten’s connection with the vegetation structure. There is a tendency for it to transition to fairly enclosed forests at the lowest temperature and low snow cover, and vice versa [8].

One more study focused on the possible effects of climatic factors and the frequency of flooding on variation in the size of a population of brown hare (*Lepus europaeus*) living on a salt marsh in the Netherlands [9]. In the work on the basis of statistical data processing the role of climatic factors in the dynamics of hare was estimated.

Many forestry and forest management works have highlighted the important ecosystem role of small mammals and the ability to use them as bioindicators. In particular, it is mentioned the relationship of the species of this species with the level of recreational load, which is extremely relevant for the protected areas [10, 11]. Abundances of mammalian carnivores are also reduced by clear-cutting, with loss of preferred prey species, den sites, and other components of forest stand structure [12].

Some publications contain research of joint dynamics of studied species based on their statistical analysis. Analysis revealed that the 25-year period of fluctuations of marten and hare population with a phase shift of 12.5 (a high number of martens occurs 12.5 years after a high number of hare), as well as 8–10 years with a negative phase shift (a relatively high hare number follows 8–10 years after a high marten number) corresponded well to the “predator–prey” model [5].

3 Results and Discussion

According to the methodology of the development of automatic monitoring structures [1] in the first stage of MIS formation, the supersystem formulates necessities for the information contents that must be received from the monitoring results. In this case, monitoring data is required to define the response of the ecosystem to the dynamics of climate factors.

To find out about biota responses to environmental change is particularly essential for protected territories. But the challenge of entirely assessing the response of organic structures to local weather change can not be realized. Therefore, a reasonable choice of the most relevant and informative features is needed for sufficiently further research.

Small mammals, including predators, can also be the universal and appropriate objects for monitoring ecosystem responses to local weather change. This is decided through their close connection with several factors of the ecosystems and their active participation in the most important types of the biogenic cycle. They are well known for their sensitivity to small mammals for environmental change, their huge vary of

adaptations, and their capability to quickly enforce them in an ever-changing environment.

Therefore, the MIS requires a description of the dependence of the mammal's population on the territory of national park "Vizhnytskyi" on weather conditions. And it was the fifth stage of MIS construction. As climate characteristics of the studied territory, the observations of the park's weather station were used. In this research, we used two of them: air temperature and precipitation. These features were used as independent variables. Among the species' characteristics, the number of populations was selected. Because this is an integral feature of populations, that is convenient for the study and have close connections with many other parameters. Respectively, this metric was chosen as the dependent variable. In this case, the creation of a complex indicator was anticipated, so this was obtained by the synthesis of models created for each of the bioindicators species. This set of indicators was selected to test the suitability of using the proposed approach and can be expanded by both bioindicators and climate characteristics.

Numerical characteristics resulting from the implementation of the lower monitoring stratum includes: the amount of populations of the studied species, They were used as components of the initial description of the monitoring entities and as variables in the synthesis of models that are part of the higher stratum.

Table 1. List of features of the input array

№	The name of the feature	Number of variables
1	Amount of populations of the studied species (<i>Lepus europaeus</i> Pall. and <i>Martes martes</i> L.), individuals	2
2	Daily average air temperatures by months, degree	12
3	Average minimum air daily temperature by months, degree	12
4	Average maximum air daily temperature by months, degree	12
5	Minimum daily air temperature by month, degree	12
6	Maximum daily air temperature by month, degree	12
7	Average daily precipitation by months, mm	12
8	Maximum precipitation, by months, mm	12
9	Minimum precipitation, by months, mm	12
10	Average number of days with precipitation by months, mm	12
11	Average daily air temperature gradient, by months, degree	12
12	Average gradient of the maximum daily air temperature by months, degree	12
13	Average gradient of the minimum daily air temperature by months. degree	12

To enforce the second stage of MIS construction, the decomposition of the data transformation technique is carried out. The number of strata and the content of the duties for converting data for every stratum is determined.

In this case, it is essential to form three vertical strata. The lower stratum gives the transformation of ecosystem features into their numerical characteristics via making continuous observations. Means are measuring instruments for identifying the characteristics of the climate, an expert for determining the features of the marten population of the forest and the methodological assist of these processes.

The upper stratum provides the transformation of data from the shape of an array of numerical features to the form of models that reflect the exchange in the population of biological monitoring objects relying on the features of the weather conditions.

Features 1 (Table 1) were used as dependent variables (simulated metrics) of the set Y

$$Y = \{y_1, y_2\}. \quad (1)$$

Features 2–13 (Table 1) are independent variables and form the set X :

$$X = \{x_1, x_2, \dots, x_{155}\}. \quad (2)$$

Within the model construction, the problems of identification of functional dependence were solved:

$$Y = f(X). \quad (3)$$

The solution of the tasks of data transformation, formulated for each of the strata, was performed on the third stage of MIS construction. In particular, establishing the character of the weather factors influence the characteristics of a complex bioindicator, that was created taking into account trophic dependencies between two species of small mammals, was implemented.

Thus, it is necessary to solve the identification problem:

$$y_1 = f(x_1, x_2, \dots, x_{155}). \quad (4)$$

For the synthesis used the multi-model algorithm GMDH [7]. To construct the model for the input of the MIS synthesizer, the outcomes of observations got between 2002 and 2013 were used. The results of observations from 2014–2019 were used to test the model. For the models' selection, the regularity criterion was used. It was calculated as the mean square deviation of simulation results from real values on the test sequence of observation points.

By examining the model for the sensitivity to the dynamics of variables of the set X , we can evaluate the influence of factors and describe the pattern of association. The influence of the features was estimated via the values of the weighting coefficients of the modeling variables, calculated in accordance to the expression [1]:

$$W_1 = \frac{y'_1}{\sum_{i=1}^m y'_i}, \quad (5)$$

where y'_1 – the value of the partial derivative of the variable model x_1 ; m – the number of variables that are included in the model structure.

The average modeling error differs significantly by year and ranged from 0.34% to 8.22%. These results allow to conclude that the obtained model is accurate and stable. This means that the structure of this model reflects information about trends that can be used by the super system as useful.

After confirming the usefulness of the received models, the complete MIS information conversion subsystem is tested by an expert assessment of the monitoring data at the output of the system.

Thus, the proposed method for the formation of a complex bioindicator based on the synthesis of models includes the following steps:

1. Determining the purpose of the study.
2. Determination of the list of bioindicators.
3. Determination of a set of their characteristics.
4. Determining the number of observations that reflect the properties of bioindicators in the array of results.
5. Builds a database with a primary description.
6. Initial processing of observations results.
7. Building the array of input data.
8. Synthesis of the model - complex bioindicator.
9. Model testing.
10. Model study and calculation of the characteristics influence coefficients or each bioindicator.

Approbation of monitoring system work was carried out with the use of data collected from the territory of the National Park “Vizhnitsky”. In this case, we conduct an specialist assessment of the patterns of the association, which revealed the estimates of the influence of weather conditions on the populations of studied species. Analysis of the literature suggests that the most essential weather factors that affect the population of the martens are: altitude of snow cover, air temperature and precipitation [5, 13]. The marten is characterized by the presence of the so-called “false mating season” (false rut) during January-February. Its intensity and timing rely on the temperature of the air, particularly – the highest intensity of the process is tied to the maximum winter temperatures [11, 14]. Also, studies of the biology and ecology of marten allow us to talk about its steady connection with forest habitat, so the abundance of this species can serve as an indication of the forest cover of the territory and its dynamics [5].

Analysis of literature data suggests that the most important weather and climatic factors that may affect the population size of hares are: snow cover height, air temperature, duration of frost-free period, precipitation. Hare numbers and the change of it were correlated with total precipitation and the number of months in which precipitation exceeded 100 mm. The change in hare numbers negatively correlated with the same two factors, as well as total precipitation in the reproductive period, the number of months in the reproductive period in which total precipitation exceeded 100 mm, and the number of days in a year that the sea level was higher than 200 cm above normal level for investigated territory [9]. Researches results in [15] testify to defined correlation relationships presence between brown hare number and macroclimate components in 1984–2015. So the mid-annual temperature and annual precipitation are in positive correlation relationships of average force ($r = 0.50$ and $r = 0.57$ accordingly),

and mid-annual air relative humidity and average snow cover height are in weak positive correlation relationships ($r = 0.21$ and $r = 0.22$ accordingly) with number of studied species.

The results of the model’s study are presented in Table 2. According to the results of the model study, among all features that were tested, only three have a substantial (more than 10%) impact on the number of martens: the average maximum air temperature in February, the maximum daily temperature gradient in January and the minimal precipitation in June. Moreover, for all the identified features there is a direct influence. It is additionally worth noting that the absolute values of the weights. For all the studied factors, they range approximately from 25 to 27%, which could indicate their considerable and equal effect on the output. The fifth factor has confirmation in literature sources. In [5] it was shown that the minimum duration of trace activity (which is the most common way to evaluate number of animals in winter) for marten recorded in February and it connected directly to air temperature. Minimum precipitation in June may be a determining factor, since, in the summer, the percentage of plant feed in the marten diet rises notably [14]. Also, in the summer there is a mating season [5].

Table 2. The impact of model variables.

<i>Lepus europaeus</i> Pall.		<i>Martes martes</i> L.	
Variable	Weight, %	Variable	Weight, %
Daily average air temperatures in September	26,80	Average maximum air daily temperature in February	24,96
Average maximum air daily temperature in July	46,55	Average gradient of the maximum daily air temperature in January	26,78
Average daily precipitation in April	2,87	Minimum precipitation in June	25,46
Average gradient of the maximum daily air temperature in August	5,77	Average gradient of the maximum daily air temperature in May	3,44

For the second species, a reliable influence on the amount of maximum air temperature and precipitation in September (the first one by the principle of feedback) was established. For both studied factors they are much more than 10%, which indicates their significant impact on the resulting indicator. At the same time, the weight for the precipitation factor is approximately 2 times higher than the values for the maximum temperature factor in September. The concentration of the resulting indicators in September may have several reasons - this very month in the *Lepus europaeus* Pall. the last brood of is born, and local migrations and molting may begin, may depend both on the length of daylight and air temperature [5, 13].

Worth to mention, well-known function of forests, as “ecological engineers” [16]. Specifically, function of the forest can be viewed in the equalization of temperature of

the territory: decreasing the temperature under the forest canopy in summer and rising in winter [17]. Thus, the sensitivity of the marten to maximum winter temperatures is associated in particular to the animal's susceptibility to forest biotopes [18].

So, the established trends are confirmed by the literature data, which testifies to the possibility and efficiency of using studied species as a bioindicators. However, it would be wrong to associate directly the fluctuations in the animal population to the studied factors. For example, the "lag" effect is an essential factor in studying the responses of open systems [15, 19, 20], so in the following studies, it would be logical to take it into account when forming array of input data.

4 Conclusions and Suggestions

The use of the multilevel monitoring information system on the example of analyzed data to discover trends of weather factors has an impact on the population of biomonitoring objects was demonstrated. The efficiency of the use of the populations' dynamics of small mammals as one of the bioindicators of monitoring local weather change has been proven. The proposed method allows to prioritize the use of individual bioindicators, depending on the modeling results. The opposite impact of maximum air temperature and direct precipitation in September on the gray hare population was established. For the marten, another set of significant factors was determined. The direct impact of the average maximum air temperature in February, the maximum daily temperature gradient in January and the minimal precipitation in June on the population of the forest martens were established.



References

1. Holub, S.: Multilevel modeling in environmental monitoring technologies, Bogdan Khmelnytsky Cherkasy National University, Cherkasy (2007)
2. Sandifera, P., Sutton-Grierb, A., Wardc, B.: Exploring connections among nature, biodiversity, ecosystem services, and human health and well-being: opportunities to enhance health and biodiversity conservation. *Ecosyst. Ser.* **12**, 1–15 (2015)
3. Russo, I., Sole, C., Barbato, M., von Bramann, U., Bruford, M.: Landscape determinants of fine-scale genetic structure of a small rodent in a heterogeneous landscape (Hluhluwe-iMfolozi Park, South Africa). *Sci. Rep.* **6**, 22–30 (2016)
4. Twining, J., Montgomery, I., Fitzpatrick, V., Marks, N., Scantlebury, D., Tosh, D.: Seasonal, geographical, and habitat effects on the diet of a recovering predator population: the European pine marten (*Martes martes*) in Ireland. *Eur. J. Wildl. Res.* **65**, 1–15 (2019)
5. Zheltukhin, A., Puzachenko, Yu., Kotlov, I., Shironiya, I.: Spatial-temporal dynamics of the pine marten (*Martes martes*), mountain hare (*Lepus timidus*) and red squirrel (*Sciurus vulgaris*) trail activity in the European southern taiga. *J. Gen. Biol. Methodol. Zool.* **7**(1), 37–55 (2017)
6. Pearce, J., Venier, L.: Small mammals as bioindicators of sustainable boreal forest management. *Forest Ecol. Manag.* **208**(1), 153–175 (2005)
7. Ivakhnenko, A.H., Myuller, J.A.K.: *Self-Organization of Forecasting Models*. Naukova dumka, Kyev (1985)

8. Zalewski, A., Jedrzejewski, W.: Spatial organisation and dynamics of the pine marten *Martes martes* population in Białowieża Forest (E Poland) compared with other European woodlands. *Ecography* **29**(1), 31–43 (2006)
9. van Wieren, S., Prins, H., Wiersma, M.: Climatic factors affecting a brown hare (*Lepus europaeus*) population. *Lutra* **49**(2), 103–110 (2006)
10. Gołos, P.T.: Selected aspects of the forest recreational function in view of its users. *Forest Res. Pap.* **74**(1), 57–70 (2013)
11. Sheehy, E., O'Meara, D., O'Reilly, C., Smart, A.: A non-invasive approach to determining pine marten abundance and predation. *Eur. J. Wildl. Res.* **60**(2), 23–40 (2013)
12. Sullivan, T.P., Sullivan, D.S., Lindgren, P.M.F., Ransome, Douglas B.: If we build habitat, will they come? Woody debris structures and conservation of forest mammals. *J. Mammal.* **93**(6), 1456–1468 (2012)
13. Mikheyev, A.V.: The seasonal transformation of ration structure of martens genus in forest ecosystems of Ukraine steppe zone. *Naukovyj visnyk Ukrainian State For. Univ.* **14**(6), 71–75 (2004)
14. Vladimirova, E.D.: Features of false rut of the pine marten in vicinities of Samara. *Povolzhsky Ecol. J.* **17**, 14–21 (2019)
15. Belsky, V.V., Lebedeva, N.I., Maslova, O.V.: Macroclimate components influence on brown hare (*Lepus europaeus* Pallas, 1778) number dynamics in Zaporizhzhya region. *Bull. Zaporizhzhia Natl. Univ. Biol. Sci. Zool. Anim. Ecol.* (2), 35–40 (2016)
16. Miyawaki, A., Golley, F.B.: Forest reconstruction as ecological engineering. *Ecol. Eng.* **2**(4), 333–345 (1993)
17. Renaud, V., Innes, J., Dobbertin, M., Rebetez, M.: Comparison between open-site and below-canopy climatic conditions in Switzerland for different types of forests over 10 years (1998–2007). *Theor. Appl. Climatol.* (105), 119–127 (2011)
18. Stelmakh, S.: Biotopes, refuges and food of pine marten (*Martes martes* L.) on Roztochya (Lviv region). *Visnyk Lviv Univ. Ser. Biol.* (63), 35–43 (2013)
19. Ryznichenko, G.Yu.: *Mathematical models in biophysics and ecology* IKI, Moscow-Izhevsk (2003)
20. Murray, J.D.: *Mathematical Biology. An Introduction.* Springer, Heidelberg (2002)



Mathematical Modeling of the Summarizing Index for the Biosystems Status as a Tool to Control the Functioning of the Environmental Management System at Business Entities

Olena Barabash¹  and Ganna Weigang² 

¹ National Transport University,
Omelianovicha-Pavlenko ave. 1, Kiev 01010, Ukraine
el_barabash@ukr.net

² Banking University, Andriivska Street, 1, Kiev 04070, Ukraine
weigang.ganna@gmail.com

Abstract. The main characteristics of the environmental management systems at business entities depend to a large extent on the correct selection of criteria and indicators for controlling their functioning and implementing organizational changes in the Entity-Environment system. In an economic crisis and as a result of increased competition, there is a need to reduce the time and cost of monitoring studies, which are mandatory to control the functioning of environmental management systems and replace them with computational experiment using developed mathematical models. The purpose of this work is to develop multifactorial models of biosystems status for establishing the ecological efficiency of environmental management systems of business entities, by conducting complex monitoring of the system Entity-Environment using bioindication and biotesting methods. The processing of the experiment results allowed to obtain mathematical models of dependence for the biosystems status and the environmental quality from organizational changes in the system Entity-Environment before and after the implementation of the environmental management system. Graphical interpretation of modeling results allows to predict not only biosystems status by environmental indicators, but also to predict the level of environmental safety in the cities and regions, depending on the level of efficiency and control of activities at the business entities operating in various sectors of the economy.

Keywords: Mathematical modeling · Environmental management system · Summarizing index of the biosystems status · Bioindication · Biotesting

1 Introduction

1.1 Problem Statement and Purpose of the Study

The development of business entities in Ukraine should take place in the direction taken by the enterprises of the European Union. The introduction of environmental

management systems is a tool for achieving the economic interests of enterprises and the development of environmentally aware management by monitoring environmental aspects of activities, products and services aimed at implementing the environmental protection measures improving the natural components of the environment. The complexity and complication of the problems, the dynamism of the situations occurring in the Entity-Environment system, lead to the need to study the technical, economic, environmental, social, psychological, managerial and other aspects inside and outside the system. The reduction of ecosystems natural capital [1], the strengthening of the human factor role in management, necessitate the application of a systematic approach to changes in the organization of business entities. In today's world, which strives for ecological, socio-economic development, there is a need to introduce environmental management systems (EMSs) aimed at improving the environmental security of business entities. The main task in the EMS development and implementation is to ensure its continuous monitoring to achieve effective functioning. Therefore, carrying out monitoring studies to determine the viability of the Entity-Environment system during the EMS development, implementation and operation will allow controlling the impact of the business entities on the natural components of the environment to achieve the DSTU ISO 14001: 2015 requirements. The selection of criteria and indicators, the application of methods and tools of laboratory research is a characteristic feature of changes in the philosophy of organization of activity within the Entity-Environment system and should take place in the first stage of EMS implementation in order to control its efficiency of functioning. It is important to monitor group indicators of environmental performance, management effectiveness and environmental efficiency to ensure the control for the EMS functioning. An assessment of the environmental performance of an entity's environmental safety management system contains in its structure separate factors collectively determining the value of summarizing indexes - the environmental quality and the biosystems status. Given the large amount of numerical data included in these indicators, it is best to use mathematical modeling methods to calculate and visualize them. Thus, the formation of a mathematical apparatus for the development of models predicting the biosystems status allowing to determine the level of EMS environmental performance, which is a significant component of the complex criterion for its functioning for business entities operating in various sectors of the economy is a relevant scientific problem.

The purpose of this work is to develop a multifactor model of the biosystems status for establishing the ecological effectiveness of environmental management systems at business entities, by carrying out complex monitoring of the Entity - Environment system using bioindication and biotesting methods.

1.2 Actual Scientific Researches and Issues Analysis

A number of current domestic and international studies are devoted to the issue of enhancing the environmental protection of business entities through the introduction of environmental management systems and monitoring methods used in controlling the status of its performance indicators.

Thus, the works of A.I. Borodin [2] and N.V. Pakhomova [3] present research on a number of environmental problems, which rise necessitates the creation of a more

efficient system of environmental measures at the micro-level and explains the importance of developing new approaches for business entities in matters of interaction with the environment applying innovations. Research on the creation of national complex ecological and biological accounting systems for practical strengthening of the system of effective and efficient ecological management and increasing the efficiency of data collection has been considered in the works of A.K. Murtazov [4], N.F. Ivanitskaya [5], and V. Pohrebennyk [6].

The works in which bioindication and biotesting methods are used to monitor the state of the biosphere components in various industries, including chemical industry, TDK, oil and gas processing industries are of particular interest [7–12].

The issue of corporate governance greening was considered by O.Yu. Boyko [13]; Prospects for implementation of international environmental standards in Ukraine were investigated by Yu.V. Bilyavs'ka [14] and V.O. Anishchenko [15]. In the work of V.V. Hassiy, I.M. Potravny [16], methods of economic regulation for environmentally important projects in the field of nature management and environmental protection were investigated. The effectiveness of the implementation of the international standard ISO 14001 and the need for certification of environmental management systems to improve the economic performance of enterprises have been proved [17, 18]. However, there is no the experience of monitoring within the sanitary protection zone of business entities before the beginning of the development and after the implementation of environmental safety management systems using bioindication and biotesting methods through mathematical modeling to predict a group indicator of EMS environmental performance comprehensive criterion for controlling its functioning.

2 Systematic Approach to Modeling the Biosystems Status

To develop a mathematical model allowing to predict the biosystems status in the structure of a group indicator of the environmental effectiveness for the environmental management system at business entities, depending on the reduction of emissions into the air, discharges into water objects and reducing the load on soil cover within the sanitary protection zone for the conservation and restoration of the state of natural components of the environment by the entity, the approaches of experimental and statistical modeling were used. The objective of such modeling is to determine the relationship between the input parameters of the biosystems of the Entity-Environment biosystems, which were determined before and after the implementation of the environmental management system, and the output parameters - indicators for assessing the level of organizational changes and environmental measures optimizing functioning of the environmental management systems. Building a deterministic statistical model of the biosystems status as a tool for assessing the environmental performance of an entity's environmental management system requires knowledge of its structure and information about the processes occurring within it to reduce the negative impact on the state of the natural components of the environment.

Therefore, combining the inputs of the system with the output parameters, their consistency, which is based on the statistical processing of experimental data with interval representation, is achieved.

2.1 System Model of the Biosystems Status

Synthesis and analysis of complex systems allows to determine the factors influencing the properties and condition of the research objects. In the process of studying a complex system, there is a need to apply systematic approaches allowing to decompose the whole into components (to perform decomposition), or to combine the components into a whole (to carry out aggregation). The impact of business entities on the biosystems status is characterized by the level of environmental pollution. Using system analysis methods and IDEF0 methodology for Enterprise Modeling, the functional model of influence of business entities activity on the biosystems status is developed and presented in Fig. 1.

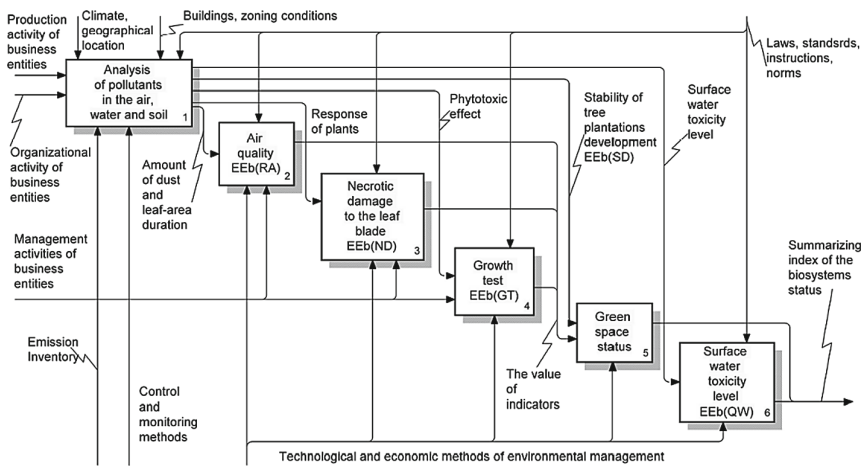


Fig. 1. Functional model of the business entities impact on the biosystems status

The analysis of the functional model of the biosystems status allowed to develop a conceptual system model, which is presented in the form of a tuple of functions determining the level of ecological instability of natural components. The summarizing index of the biosystems status (*EEb*) includes integrating indicators and factors determining the response of green spaces to the pollution degree of natural environmental components by the activities of business entities and can be presented as follows:

$$EEb = \langle EEb(RA), EEb(ND), EEb(GT), EEb(SD), EEb(QW) \rangle, \tag{1}$$

where *EEb(RA)* is a summarizing indicator characterizing the response leaf blade to the level of environmental pollution in the course of the business entities activity; *EEb(ND)* is a summarizing indicator determining the level of necrotic damage to the leaf blade; *EEb(GT)* is a summarizing indicator characterizing the response of the environment to the presence of toxic metal compounds; *EEb(SD)* is a factor determining the stability of the plantation; *EEb(QW)* is a factor determining the surface water toxicity level.

Summarizing indicators, in turn, are features including a set of values for individual targeting factors and characterizing the extent of damage to green spaces over a period of time:

$$\begin{cases} EEb(RA) = f(\min(EEb(RA)_1) \cup \max(EEb(RA)_2)) \\ EEb(ND) = f(\min(EEb(ND)_1) \cup \min(EEb(ND)_2) \cup \min(EEb(ND)_3)) \\ EEb(GT) = f(\min(EEb(GT)_1) \cup \min(EEb(GT)_2)) \end{cases} \quad (2)$$

where $\min(EEb(RA)_1)$ is minimization for the amount of dust per leaf area, mg/cm^2 ; $\max(EEb(RA)_2)$ is leaf blade area, cm^2 ; $\min(EEb(ND)_1)$ is minimization of necrotic lesions in angiosperms leaf blades, represented by the results of T.Cordata response, %; $\min(EEb(ND)_2)$ is minimization of necrotic lesions in gymnospermous leaf blades represented by the results of P.Sylvestris response (necrosis), %; $\min(EEb(ND)_3)$ is minimization of necrotic lesions in the gymnospermous leaf blades represented by the results of the P.Sylvestris response (chlorosis), %; $\min(EEb(GT)_1)$ is minimization of phytotoxic effect for precipitation, % $\min(EEb(GT)_2)$ is minimization of phytotoxic effect for soil, %.

2.2 Mathematical Aspects of Generalization for Biosystems Indicators

The significance of the factors forming the summarizing index of the biosystems status has been determined experimentally using of bioindication and biotesting methods. Based on the obtained data, the response of the sensitive indicator plants and the test function of the test objects on the state air, soil cover and surface water pollution within the sanitary protection zone of the business entities before and after the introduction of EMS were calculated for economic activities carried out by different types of entities within the administrative-territorial areas of Kyiv. Thus, the obtained data allowed us to determine not only the amount of dust and the presence of certain compounds that are by-products of an entity's activities in natural components the environment, but also establish their level of influence, which affected the physiological, morphological and biochemical changes occurred in the plant and animal organisms during all stages of ontogeny before and after EMS introduction.

A comprehensive analysis of the factors before and after the EMS introduction of the for the construction of a mathematical model was carried out taking into account the dynamics of changes in the data on the state of natural components over a certain period of time. In the process of complex assessment for the dynamics of the influence of various factors on the overall level of the biosystems status, normalization or scaling was conducted to determine the rate of change for the factors (Δ):

$$\begin{aligned} \Delta &= X_k - X_n, \text{ if } X_k \rightarrow \max \\ &OR \\ \Delta &= X_n - X_k, \text{ if } X_k \rightarrow \min, \end{aligned} \quad (3)$$

where X_n is the value obtained before the EMS introduction (initial value); X_k is the value obtained after the EMS introduction (final value).

The survey data were normalized according to the purpose of targeting the values of the sampling data for each indicator individually. The difference between the given analytical method and the common one is the presence of zero level, and the general limits of the obtained criterion are in the range [0; 1]. Zero or neutral levels were recorded when the system status did not change during the study period. This allowed us to scale values by the following rule:

$$\Delta_i = \left\{ \begin{array}{ll} 0,5 + \frac{0,5 * X_i}{\max_i(|\min_i|, |\max_i|)}, & \text{if } \Delta > 0 \\ 0,5 & \text{if } \Delta = 0 \\ 0,5 - \frac{0,5 * X_i}{\max_i(|\min_i|, |\max_i|)}, & \text{if } \Delta < 0 \end{array} \right\}, \tag{4}$$

where Δ_i is the dynamics of change for i -th indicator; $\max_i(|\min_i|, |\max_i|)$ – the maximum level between the absolute values of the maximum and minimum rates of increase/decrease and the i -th indicator.

Thus, it made it possible to determine the state of the system by the complex criterion for controlling the EMS functioning, which are presented in Table 1.

Table 1. Levels of indicators state and the biosystem as a whole

Level of the state	Value of the indicator
Critical	[0; 0.4)
Zero (neutral)	[0.4; 0.7]
Positive	(0.7; 1.0]

In the course of the experimental and statistical research, a mathematical module of a summarizing index of the biosystems status was formed, which is defined as follows:

$$EEb = \frac{1}{5} \left(\frac{\sum_{i=1}^{m_1} EEb(RA)_i}{m_1} + \frac{\sum_{i=1}^{m_2} EEb(ND)_i}{m_2} + \frac{\sum_{i=4}^{m_3} EEb(GT)_i}{m_3} + EEb(SD) + EEb(QW) \right), \tag{5}$$

where m_1, m_2, m_3 are the number of individual factors forming the separate groups ($m_1 = 2, m_2 = 3, m_3 = 2$). Depending on the location and nature of the business entities, not all individual factors can be included in the summarizing indicators.

3 Results and Discussion

3.1 Experimental Studies of the Summarizing Index for the Biosystems Status

The main objective of the experimental study is to conduct comprehensive monitoring to determine the state of chemical pollution of the atmospheric air, water bodies and soil cover and identify the response and test function of living organisms to the degree of toxicity, conservation or restoration of natural components of the environment through the of bioindication and biotesting methods, which allowed to obtain the values of summarizing indicator and individual factors determining the biosystems status within the sanitary protection zone of business entities before and after the EMS introduction. The obtained results allow us to analyze the level of functioning of the environmental management systems of business entities and are a tool for controlling the level of environmental pollution.

Experimental studies were carried out within the sanitary-protective zone of ten business entities carrying out different economic activities and registered in ten administrative and territorial districts of Kyiv. It should be noted that the experiment was conducted in two stages - before the development and after the implementation of the environmental management system at business entities. This step-by-step experiment made it possible to determine the level of organizational changes and environmental measures undertaken by business entities to implement the EMS environmental program.

3.2 Experimental Determination of Components for the Indicator of the Biosystems Status

The level of summarizing indicator and factors was determined for ten research entities operating in different administrative and territorial areas of Kyiv and related to different areas of economic activity.

For each summarizing indicator and factor, a table of normalized values calculation was constructed. Table 2 shows an example of the calculation for the values of the EEB (RA) air quality based on the experimental results of the study.

Table 2. Results of the air atmospheric air quality assessment (EEb) study

Business entities No.	Amount of dust per leaf blade area, (mg/cm ²), <i>EEb(RA)_i</i>			Normalized factor value	Leaf blade area (cm ²), <i>EEb(RA)₂</i>			Normalized factor value	Normalized value of the summarizing air quality indicator «EEb (RA)»
	Before EMS implementation	After EMS implementation	Δ		Before EMS implementation	After EMS implementation	Δ		
1	0.088	0.041	0.047	1.00	30.53	36.25	5.72	1.00	1.00
2	0.052	0.044	0.008	0.59	28.79	32.23	3.44	0.80	0.69
3	0.051	0.032	0.019	0.70	26.6	31.79	5.19	0.95	0.83
4	0.038	0.031	0.007	0.57	30.87	35.17	4.30	0.88	0.73
5	0.064	0.052	0.012	0.63	31.44	34.58	3.14	0.77	0.70
6	0.053	0.039	0.014	0.65	28.41	32.31	3.90	0.84	0.74
7	0.054	0.046	0.008	0.59	28.41	32.31	3.90	0.84	0.71
8	0.042	0.041	0.001	0.51	35.29	37.1	1.81	0.66	0.58
9	0.042	0.035	0.007	0.57	30.52	33.66	3.14	0.77	0.67
10	0.044	0.038	0.006	0.56	28.83	30.06	1.23	0.61	0.59
Maximum value			0.047				5.72		
Minimum value			0.001				1.23		

As a result of the experimental data processing, the values of all components included in the structure of determining the summarizing index of the biosystems status were obtained (Fig. 2).

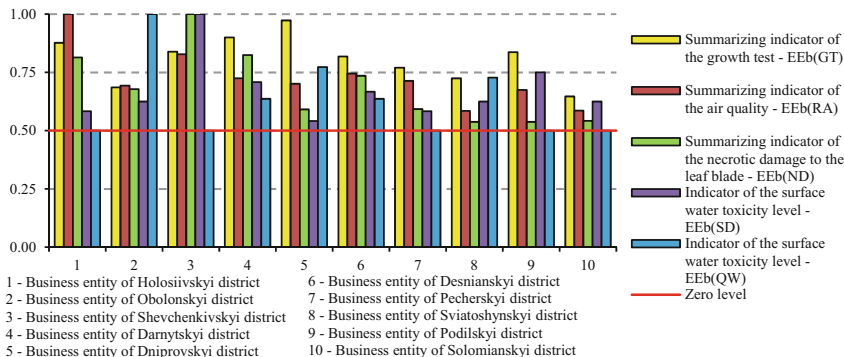


Fig. 2. Comparative analysis of the calculation results for the components of the summarizing index of the biosystems status

The graphical analysis showed that the entities with the highest proportion of indicators and factors, which are positive, operate in the Darniyskiy and Dniprovskiy districts. Low levels of management environmental activity are observed in the business entities of Solomianskiy and Sviatoshynskiy districts, most of the normalized values of indicators and factors of which are low or close to zero.

3.3 Investigation of the Level of the Summarizing Index of the Biosystems Status Using a Mathematical Model to Control the Environmental Management System Functioning

The use of the developed mathematical model allowed to search for the dependence of the summarizing index of the biosystems status (EEb) on the results of the application of bioindication and biotesting methods before and after the EMS introduction (Table 3).

The highest level of normalized value is for the enterprises of Shevchenkiy and Holosiivskiy districts, which are 0.85 and 0.81 respectively. This indicates that the activities of these enterprises are shaped with a view to minimizing the impact on the natural components of the environment. The greatest negative impact on the environment is caused by the activity of business entities in the Solomianskiy, Sviatoshynskiy and Pecherskiy districts, the normalized value of the summarizing index of the biosystems status is 0.58, 0.63 and 0.65 respectively.

Table 3. Normalized values of the summarizing index of the biosystems status (EEb) within a business entity’s sanitary zone

N	Summarizing index of the biosystems status - <i>EEb</i>	Administrative district of Kyiv	Type of economic activity (NACE code) of business entities
1	0.81	Holosiivskiy t	42.11. Construction of roads and motorways
2	0.72	Obolonskyi	10.11. Production of meat
3	0.85	Shevchenkivskiy	68.20 Renting and operating of own or leased real estate
4	0.78	Darnytskyi	20.42. Perfume and cosmetics production
5	0.73	Dniprovskiy	26.51. Manufacture of instruments and equipment for measurement, research and navigation
6	0.74	Desnianskyi	49.31. Land and suburban passenger land transport
7	0.65	Pecherskyi	56.10 Restaurants, mobile food service activities
8	0.63	Sviatoshynskiy	47.30 Retail trade of fuel (major)
9	0.67	Podilskyi	81.10. Integrated object maintenance
10	0.58	Solomianskyi	47.11. Retail sale of food, beverages and tobacco in non-specialized stores

The Fig. 3 shows the results of simulation of the summarizing index of the biosystems status in the form of a radar plot. In order to analyze the state of the environment and determine the level of functioning of the environmental management system at business entities, a comparison was made with zero or with the level of inactivity or passive environmental management.

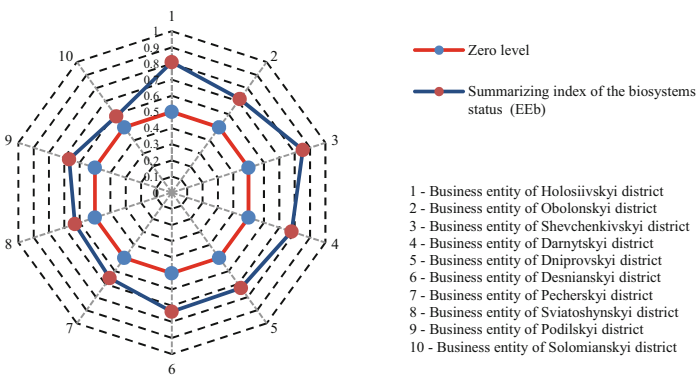


Fig. 3. Analysis of biosystems index (EEb) simulation results

Thus, the results of the research determine the dependence of the biosystems status on the level of organizational changes aimed at improving environmental management and activity in the Entity-Environment system. The constructed experimental and statistical model can be used for decision making and forecasting of the level of overall control of the environmental management systems functioning, taking into account the quality and effectiveness of the environmental measures implemented by business entities.

4 Conclusions

Systematic analysis and experimental studies of summarizing index of the biosystems status allowed to determine the factors influencing the state of organizational changes during the implementation of EMS by business entities carrying out different types of economic activity in ten administrative districts of Kyiv. These factors determine the insufficient reduction of pollutant emissions into the atmosphere, which leads to low levels of conservation or restoration of natural components of the environment and causes a negative response to their state of living organisms. The processing of the experimental results by the method of taking into account the interval changes of the state of the systems allowed us to develop a methodology for the formation of normalized factors and to build an experimental-statistical mathematical model, as the dependence of the biosystems status and the environmental quality on certain factors.

Graphical interpretation of the results made it possible to visualize the calculations results of both individual factors and a summarizing index of the biosystems status using an experimental-statistical model. The analysis of the research results showed the dependence of the biosystems status on the implemented nature conservation measures and the implementation of the environmental program of the environmental management system at business entities. The graphical interpretation of the calculations has shown that the EMS implementation with the implementation of all DSTU ISO 14001: 2015 requirements and the constant monitoring to control the EMS functioning, increases the level of its environmental performance, which is reflected in the improvement of the natural components of the environment.

The constructed mathematical model is a tool for effective environmental management decision making and forecasting of biosystems state. This will allow to predict the level of environmental safety in cities and regions, depending on the level of efficiency and control of the organizational activity at business entities operating in different sectors of the economy.

References





1. Barabash, O.V., Vozniuk, Ya.Yu.: Environmental footprint as an indicator of the natural capital of ecosystems. *Environ. Sci.: Sci. Pract. J.* **1**(20), 109–114 (2018)
2. Borodin, A.I.: Features of the formation of environmentally sound enterprise management. *Herald Kazan Technol. Univ.* **2**, 411–416 (2003)

3. Pakhomova, N.V., Rikhter, K.K., Malyshkov, G.B.: Sustainable development strategy and transition to a green economy: updating priorities and mechanisms. *Vestnik St. Petersburg Univ.* **4**(5), 35–54 (2013)
4. Murtazov, A.K.: *Ecological Monitoring. Methods and Means*. Ryazan State University named for S. Yesenin, Ryazan (2008)
5. Ivanitskaya, N.F., Grishchenko, S.V., Stepanova, M.G.: Environmental monitoring of heavy metals in the environment of the Donetsk region as an element of public health management. *Collected Pap. DonSU* **36**, 99–108 (2004)
6. Pohrebennyk, V., Melnyk, N., Boichuk, M.: Environmental monitoring: concepts, principles, systems. *Meas. Eng. Metrol.* **65**, 164–171 (2005)
7. Klemm, D.J., Lobring, L.B., Horning, W.H.: *Manual for the Evaluation of Laboratory Performing Aquatic Toxicity Tests*. U.S. Environmental Protection Agency, Cincinnati (1990)
8. Michaud, A., Chappelaz, C., Hinsinger, P.: Copper phytotoxicity affects root elongation and iron nutrition in durum wheat (*Triticum turgidum durum* L.). *Plant Soil* **310**(1–2), 151–165 (2008)
9. Aranda, E., García-Romera, I., Ocampo, J.A., Carbone, V., Mari, A., Malorni, A., Sannino, F., De Martino, A., Capasso, R.: Chemical characterization and effects on *Lepidium sativum* of the native and bioremediated components of dry olive mill residue. *Chemosphere* **69**(2), 229–239 (2007)
10. Blok, C., Persoone, G., Wever, G.: A practical and low cost microbiotest to assess the phytotoxic potential of growing media and soil. *Acta Hort.* **779**, 367–374 (2008)
11. Pinho, I.A., Lopes, D.V., Martins, R.C., Quin, M.J.: Phytotoxicity assessment of olive mill solid wastes and the influence of phenolic compounds. *Chemosphere* **185**, 258–267 (2017)
12. Kebrom, T.H., Woldesenbet, S., Bayabil, H.K., Garcia, M., Gao, M., Ampim, P., Awal, R., Fares, A.: Evaluation of phytotoxicity of three organic amendments to collard greens using the seed germination bioassay. *Environ. Sci. Pollut. Res.* **26**(6), 5454–5462 (2019). <https://doi.org/10.1007/s11356-018-3928-4>
13. Boiko, O.Y.: Greening of corporate governance at agribusiness enterprises. *Bull. Sumy State Agrarian Univ.* **2**, 308–310 (2015)
14. Bilyavs'ka, Yu.V.: Ecological management of an enterprise. *Econ. Ukraine* **4**(645), 104–111 (2016)
15. Anishchenko, V.O., Halkevych, T.L.: *Modern methodological principles of ecological management formation in Ukraine*. MAUP, K. (2008)
16. Hassiy, V.V., Potravnyi, I.M.: Ecological responsibility of business as an element of public-private partnership. *Marketing and management* **3**, 179–187 (2011)
17. Firuz, D.: Economic instruments of environmental management. *Proc. Int. Acad. Ecol. Environ. Sci.* **1**(2), 97–111 (2011)
18. Gwen, Ch., Fetsko, M., Hendrickson, Ch.: Environmental management systems and ISO 14001. Certification for construction firms. *J. Constr. Eng. Manag.* **130**(3), 330–336 (2004)

Mathematical Modeling and Simulation of Systems in Manufacturing



Theoretical Studies of the Rectilinear Motion of the Axis of the Locked Wheel After Braking the Vehicle on the Uphill

Serhii Shuklinov¹ , Dmytro Leontiev¹ ,
Volodymyr Makarov² , Victor Verbitskiy³ , and Anton Hubin¹ 

¹ Department of Automobiles named after A.B. Gredeskul,
Kharkiv National Automobile and Highway University,
Yaroslava Mudrogo Street, 25, Kharkiv 61002, Ukraine
dima.a3alij@gmail.com

² Vinnytsia National Technical University, 95 Khmelnytske Shose,
Vinnytsia 21021, Ukraine

³ V. N. Karazin Kharkiv National University,
4 Svobody Sq., Kharkiv 61022, Ukraine

Abstract. The article deals with the mathematical model of wheel axle displacement after the end of the vehicle braking process on the uphill in the absence of moving the pneumatic tire contact patch relative to the road surface. The proposed dependencies for determining the coordinate of the wheel axis position and the speed of its movement on the basis of the general dynamics equation allow to determine the time during which the wheel axis deviates by the maximum distance from the axis of symmetry of the contact patch of the pneumatic tire, which is stationary relative to the surface of the pavement. The obtained simulation results made it possible to establish a relationship between the position of the wheel axis, the mass of the vehicle and the value of the longitudinal uphill of the pavement surface, and made it possible to evaluate the possibility of using the energy of the elastic deformation of the pneumatic tire at the moment when the vehicle started to moving on the uphill. The proposed mathematical model of rectilinear motion of the axle of the locked wheel, the contact patch of which does not move relative to the surface of the pavement, may be a theoretical basis for the analysis and synthesis of the algorithm for controlling the system of motion of the vehicle on the uphill.

Keywords: Vehicle · Uphill · Longitudinal deformation of the tire · Equation of motion · Damping · Extremum

1 Introduction

The process of moving a vehicle [1, 2], especially on the uphill [3, 4], affects the traffic flow and on road safety, which is why modern vehicles are equipped with automatic systems that help the driver begin to move on the uphill. A lot of scientific works have been devoted to the study of vehicle moving processes, which indicates the relevance of this area of research. An analysis of these works showed that the process of starting a

vehicle moving on the uphill [3, 5, 6] is determined by the change in the elastically deformed state of the vehicle pneumatic tires [7, 8], the applied forces and the corresponding reactions from the side of the road surface. It is known that after a vehicle stops on the uphill [9, 10], the pneumatic tires of its braked wheels remain deformed by the longitudinal component of the gravity. In this case, elastic forces are formed in the pneumatic tires of the braked vehicle wheels, which compensate for a part of the longitudinal force acting on the vehicle. In this case, the elastic longitudinal deformation of the pneumatic tire determines the longitudinal displacement of the wheel axis and, accordingly, of the car body. It is obvious that the dynamic state of the axis of the blocked wheel after the vehicle stops on the uphill determines the nature of the operation of automatic systems that help the driver when the vehicle begins to move.

2 Analysis of Publications

An analysis of scientific papers [3–8, 11] on modeling the process of starting a vehicle's movement showed that they are based on a description of the dynamic state of the vehicle, and the deformation of pneumatic tires [12] mounted on wheels affects the control and dynamics of the vehicle's movement. To determine the displacement of the reaction of the surface of the road surface in [13], a study was carried out determining the nature of the vehicle's movement, and in [14] equations were given for determining the longitudinal shear of the vertical reaction during vehicle movement. An analysis of [15] showed that, based on theoretical and experimental studies, its author has developed a method for calculating the position of a normal reaction relative to the center of the contact patch of a pneumatic tire with the surface of the road. The studies of the angular deformation of the pneumatic tire of a wheel, carried out in [12, 16], made it possible to obtain equations for describing the angularly damped oscillations in a flexible pneumatic tire depending on the magnitude of the torque, wheel load, and tire characteristic.

3 Methods and Materials

Determining the stress state of the tires of all the braked wheels of a vehicle standing on the uphill is a difficult statically indefinable task. In this regard, at the first stage of the study of the dynamic state of the axis of the locked wheels under the action of the longitudinal component of the vehicle's gravity (Fig. 1), we accept the following assumptions and conditions:

- after stopping the car on the uphill, it is held in place by the wheels of one axis;
- wheels of the same axles have the equal loads, dimensions and characteristics;
- the wheels of one axis are replaced by one wheel located in the central longitudinal plane of the vehicle;
- there is no air movement around the vehicle;
- the suspension of all wheels is absolutely stiff;
- the surface is absolutely flat and hard, located at an angle α to the horizon;

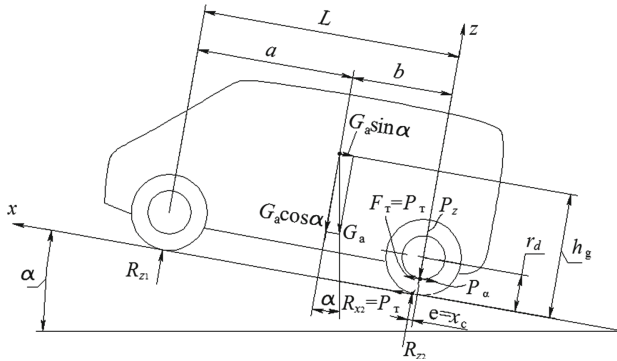


Fig. 1. Scheme of external influences acting on a stationary vehicle: xoz - motionless coordinate system; L - the distance between the axles of the wheels of the vehicle; a, b, h_g - coordinates of the center of gravity of the vehicle; r_d - dynamic radius of the wheel; G_a - vehicle gravity; R_z - wheel load; parameters having an index of “1” correspond to the vehicle front axis, and an index of “2” corresponds to the vehicle rear axis

- the force of adhesion of the wheels to the supporting surface is greater than the longitudinal forces acting on the wheels.

A stationary vehicle standing on the uphill in the longitudinal plane is affected by the component of gravity $P\alpha = G_a \cdot \sin\alpha$, which is balanced by the tangential reaction R_{x2} on the braked wheels. In the case of a stationary vehicle, the tangential reaction R_{x2} is equal to the braking force P_T . In this case, the tires of the braked wheels are subjected to longitudinal deformation. Elastic tire deformation force F_T with a stationary vehicle modulo equal to $P\alpha$ and P_T . As a result, in the static state of the vehicle, the normal reaction on the braked wheels shift “ $e = x_c$ ” relative to the axis of the wheel in the longitudinal plane. To compile a mathematical model of the rectilinear motion of the axis of the locked wheel after braking the vehicle on the uphill to a static state, we consider the diagram shown in Fig. 2. In the diagram below, in accordance with the accepted assumptions, the wheels of one axis are replaced by one wheel located in the central longitudinal plane of the vehicle. The scheme corresponds to the cases of braked wheels, both the front and rear axles, therefore the designation of reactions on the wheel is common. When braking a vehicle on the uphill, the braking moment M_T created by the brake mechanism acts on the wheel, and forces are applied to the wheel axis.

In this case, in the longitudinal plane of the vehicle are: part of the force of inertia of the vehicle P_{jx} ; the longitudinal component $P_{k\alpha}$ of the force of gravity G_a and the braking force P_T . In the vertical plane of the vehicle acts normal component P_z of gravity G_a .

The equilibrium equation of the wheel axis in the longitudinal direction is

$$P_x^j - P_{k\alpha} - P_T = 0. \tag{1}$$

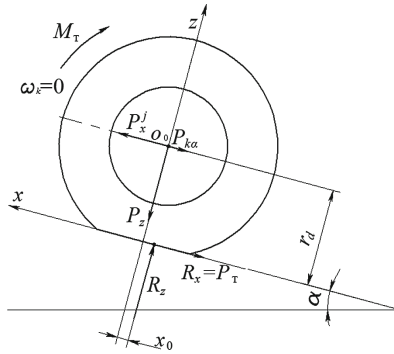


Fig. 2. The scheme of external influences acting on the locked wheel at the moment of stopping its contact spot with the road surface: xz - fixed coordinate system

A longitudinal tangent reaction R_x and a vertical reaction R_z act in the contact patch of the wheel with the road bearing surface. The vertical reaction R_z due to the tangential elasticity of the tire is shifted from the axis of the wheel o_0 by x_0 .

The scheme shown in Fig. 2 corresponds to the point in time when the linear velocity of the wheel axis and the angular velocity of the wheel ω_k is equal to zero, but the tire is in an elastically deformed state.

At the moment of stopping the wheel axis, the inertia force is zero, the realized braking force $P_{T\varphi}$ takes the value of the elastic longitudinal deformation force F_T of the tire. In this case, the wheel axis and, accordingly, the vehicle body begin to move in the opposite direction under the action of the sum of the forces $P_{k\alpha}$ and F_T . In this case, the vector of elastic force F_T decreases, vanishes and then, changing direction, increases (see Fig. 3), which affects the magnitude of the longitudinal deformation of the tire.

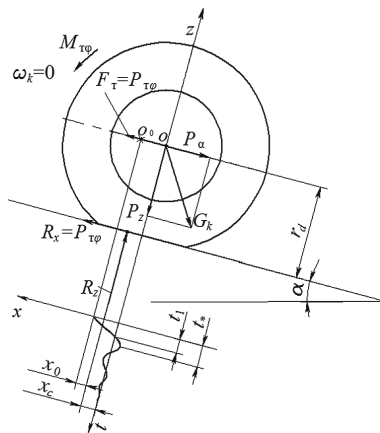


Fig. 3. The scheme of external influences acting on the locked wheel and the dependence of the displacement of its axis relative to the fixed contact patch: xz - fixed coordinate system;

The elastic force F_T of the tire increases with its deformation and at a stationary point x_c takes a steady-state value, which depends on the value of the uphill, braking force and adhesive properties of the tire. Since after stopping a vehicle on the uphill, it is held in place by the wheels of one axis, the longitudinal component of the vehicle's gravity is perceived by the wheels of this axis. This leads to an increase in tire deformation and the realized braking force $P_{T\varphi}$. The direction of the elastic force F_T , the deformed tires coincides with the direction of the coordinate axis x . The reactions R_x and R_z form the total moment, respectively, on the arms x_c and r_d , which causes a change in the direction of the braking moment M_T (see Fig. 3). In this case, the magnitude of the realized braking torque also changes.

The magnitude of the realized braking torque $M_{T\varphi}$ is determined by the coefficient of realized adhesion f_x and the coefficient of adhesion $f_x^{max} = \mu$ to the supporting surface, and if the condition $M_T \leq M_\varphi = R_z \cdot \mu \cdot r_d$ is fulfilled, then it acquires the value.

$$M_{T\varphi} = R_z \cdot f_x \cdot r_d. \tag{2}$$

The realized friction coefficient f_x depends on the uphill angle and the normal reaction of the surface to the axis of the braked wheels R_z .

$$f_x = \frac{P_\alpha}{R_z} = \frac{G_a \cdot \sin \alpha}{G_a \cdot m_{z\alpha}} = \frac{\sin \alpha}{m_{z\alpha}}, \tag{3}$$

where $G_a = m \cdot g$ vehicle gravity (m - vehicle mass, kg; $g = 9.81 \text{ m/s}^2$ - gravitational acceleration);

$m_{z\alpha}$ is the coefficient of distribution of the normal load on the braked axis of a vehicle standing on the uphill.

The distribution of the normal load on the front and rear axles of a stationary vehicle on the uphill can be determined from the corresponding expressions

$$m_{z\alpha 1} \cdot L = \cos \alpha \cdot (b - h_g \cdot \text{tg} \alpha); \quad m_{z\alpha 2} \cdot L = \cos \alpha \cdot (a + h_g \cdot \text{tg} \alpha). \tag{4}$$

Taking into account (4), Eq. (3) for the corresponding axis of the vehicle will take the form

$$f_{x1} = L \cdot \text{tg} \alpha \cdot (b - h_g \cdot \text{tg} \alpha)^{-1}; \quad f_{x2} = L \cdot \text{tg} \alpha \cdot (a + h_g \cdot \text{tg} \alpha)^{-1}. \tag{5}$$

If the realized friction coefficient f_x is less than the friction coefficient μ with the supporting surface, then the speed of the wheel axis at a stationary point with coordinate x_c becomes equal to zero.

Obviously, rolling the wheel along in the uphill is possible under conditions $f_x < \mu$ and $M_T < M_\varphi$. Under the conditions $M_T \geq P_\alpha \cdot r_d > M_\varphi$, the contact patch of the wheel and its axis, and accordingly the vehicle will slide down the uphill.

The last two cases correspond to a violation of the stability of the car on the uphill and are not considered in this paper. Based on the assumptions made and in accordance

with the diagram presented in Fig. 3, the equation of motion of the axis of the locked wheel caused by the longitudinal force has the form

$$m \cdot \ddot{x} + k \cdot \dot{x} + c \cdot x = P_{\alpha}, \quad (6)$$

where x, \dot{x}, \ddot{x} is the movement, speed and acceleration of the wheel axis relative to the center of the typo of the tire in the direction of the longitudinal axis; k – coefficient of inelastic resistance of tires of a braked axle; c – coefficient of tangential (longitudinal) stiffness of tires of the braked axle.

When solving Eq. (6), we assume that: the mass of the vehicle, the coefficients of dynamic viscosity and tangential stiffness are not changed: m -const; k -const; c -const; the uphill angle is constant and positive: α -const; $\alpha > 0$.

We transform Eq. (6) to the form

$$\ddot{x} + 2n \cdot \dot{x} + \omega^2 \cdot x = q, \quad (7)$$

where $n = 0.5 \cdot k \cdot m^{-1}$ – tire damping coefficient in the longitudinal direction; $\omega = \sqrt{c \cdot m^{-1}}$ – natural frequency of the vehicle axis in the longitudinal direction; $q = P_{\alpha} \cdot m^{-1}$ – specific longitudinal force applied to the axis wheel of the vehicle.

From Eq. (6) for a stationary point, the dependence of the coordinate of the wheel axis for a given longitudinal force has the form

$$x_c = P_{\alpha} \cdot c^{-1}. \quad (8)$$

To solve the second-order differential Eq. (7), we write the equation

$$\lambda^2 + 2n \cdot \lambda + \omega^2 = 0, \quad (9)$$

where λ – characteristic root of the equation.

The roots of the characteristic Eq. (9) determines the equality

$$\lambda_{1,2} = -n \pm \sqrt{n^2 - \omega^2}. \quad (10)$$

The solution of Eq. (7), and therefore the nature of the change in the coordinate of the wheel axis, depends on the ratio of the damping coefficient of the tire “ n ” and the frequency of own vibrations “ ω ”. For the case $n \geq \omega$, the aperiodic movement of the axis from the origin to the position of the stationary point is characteristic. Under the condition $n > \omega$, a particular solution of Eq. (7) has the form

$$\bar{x} = c_1 \cdot e^{-\lambda_1 \cdot t} + c_2 \cdot e^{-\lambda_2 \cdot t} \quad (11)$$

and the general solution of this equation, taking into account the dependence (8)

$$x = P_{\alpha} \cdot c^{-1} + c_1 \cdot e^{-\lambda_1 \cdot t} + c_2 \cdot e^{-\lambda_2 \cdot t}, \quad (12)$$

where c_1, c_2 – constant values; t – independent variable time.

The constant values c_1, c_2 of Eq. (12), which describes the dynamics of the movement of the wheel axis relative to the center of the tire contact patch on the road surface, are determined for the initial conditions. The initial conditions for the study of the rectilinear motion of the axis of the locked wheel after braking the vehicle on the uphill are determined at a point in time corresponding to the zero speed of the wheel axis (car body) moving forward on the uphill. In this case, the coordinate $x(0)$ of the wheel axis relative to the geometric center of the tire contact patch has a certain value x_0 , due to the longitudinal deformation of the tire. The longitudinal deformation of the tire at the initial moment is determined by the intensity of the previous braking, the amount of the uphill and tire parameters. We write down certain initial conditions in the form: the initial instant of time $t = 0$, the coordinate $x(0) = x_0$, and the initial velocity $\dot{x}(0) = 0$. For given initial conditions, c_1 and c_2 are determined from the expressions

$$\left. \begin{aligned} x(0) = x_0 &\Rightarrow x_0 = P_\alpha \cdot c^{-1} + c_1 + c_2 \\ \dot{x}(0) = 0 &\Rightarrow \lambda_1 \cdot c_1 + \lambda_2 \cdot c_2 = 0 \end{aligned} \right\} \Rightarrow \begin{aligned} c_1 &= (P_\alpha \cdot c^{-1} - x_0) \cdot \lambda_2 \cdot (\lambda_1 - \lambda_2)^{-1}; \\ c_2 &= -(P_\alpha \cdot c^{-1} - x_0) \cdot \lambda_1 \cdot (\lambda_1 - \lambda_2)^{-1}. \end{aligned}$$

Substituting in the Eq. (11) the values of c_1 and c_2 after the transformations, we have the dependence of the coordinate of the axis on time for this case

$$x = P_\alpha \cdot c^{-1} [1 + (1 - c \cdot x_0 \cdot P_\alpha^{-1}) \cdot (\lambda_2 \cdot e^{-\lambda_1 t} - \lambda_1 \cdot e^{-\lambda_2 t}) (\lambda_1 - \lambda_2)]. \quad (13)$$

When the condition $n = \omega$ is fulfilled, the stationary position of the point of the axis of the wheel of the vehicle occurs most quickly. In this case, the root of Eq. (9) takes a unique value $\lambda = -n$ and its particular solution has the form

$$\bar{x} = (c_3 + c_4 \cdot t) e^{-n \cdot t}, \quad (14)$$

and accordingly, taking into account dependence (8), the general solution of Eq. (6) can be written in the form

$$x = P_\alpha \cdot c^{-1} + (c_3 + c_4 \cdot t) \cdot e^{-n \cdot t}. \quad (15)$$

Given the initial conditions, the constant values of Eq. (15) are defined as

$$\left. \begin{aligned} x(0) = x_0 &\Rightarrow x_0 = P_\alpha \cdot c^{-1} + c_3 \\ \dot{x}(0) = 0 &\Rightarrow -n \cdot c_3 + c_4 = 0 \end{aligned} \right\} \Rightarrow \begin{aligned} c_3 &= (P_\alpha \cdot c^{-1} - x_0); \\ c_4 &= -n \cdot (P_\alpha \cdot c^{-1} - x_0). \end{aligned}$$

Substituting the values c_3 and c_4 into the Eq. (15) after the transformations, we obtain the dependence of the coordinate of the axis on time for this case

$$x = P_\alpha \cdot c^{-1} \cdot [1 + (1 - c \cdot x_0 \cdot P_\alpha^{-1}) \cdot (-1 - n \cdot t) \cdot e^{-n \cdot t}]. \quad (16)$$

For the case $n < \omega$, the general solution of the Eq. (9) has the form

$$\bar{x} = e^{-\lambda_3 t} \cdot [c_5 \cdot \cos(\lambda_4 \cdot t) + c_6 \cdot \sin(\lambda_4 \cdot t)], \quad (17)$$

where $\lambda_3 = n$, $\lambda_4 = \sqrt{\omega^2 - n^2}$; c_5 , c_6 – constant values.

The general solution of Eq. (7), taking into account dependence (8) in this case, has the form

$$x = P_\alpha \cdot c^{-1} + e^{-\lambda_3 t} \cdot [c_5 \cdot \cos(\lambda_4 \cdot t) + c_6 \cdot \sin(\lambda_4 \cdot t)]. \quad (18)$$

The values of the constant values c_5 and c_6 in the general solution of the equation are determined for the above initial conditions

$$\begin{aligned} x(0) = x_0 &\Rightarrow x_0 = P_\alpha \cdot c^{-1} + c_5 \Rightarrow c_5 = x_0 - P_\alpha \cdot c^{-1}; \\ \dot{x}(0) = 0 &\Rightarrow -\lambda_3 \cdot c_5 + \lambda_4 \cdot c_6 = 0 \Rightarrow c_6 = -\lambda_3 \cdot \lambda_4^{-1} \cdot (x_0 - P_\alpha \cdot c^{-1}). \end{aligned}$$

Substituting the values of c_5 and c_6 into Eq. (18) after transformations, the general solution of Eq. (3) for this case will take the form

$$x = P_\alpha \cdot c^{-1} \{1 - e^{-n \cdot t} \cdot (1 - x_0 \cdot c \cdot P_\alpha^{-1}) \cdot [\cos(\lambda_4 \cdot t) + n \cdot \lambda_4^{-1} \cdot \sin(\lambda_4 \cdot t)]\}. \quad (19)$$

From Eq. (19) it can be seen that the axis of the wheel will make damped oscillations around a stationary point defined by dependence (8). The position of the axis of the wheel will correspond to the position of the stationary point when the condition

$$\cos(\lambda_4 \cdot t) + n \cdot \lambda_4^{-1} \cdot \sin(\lambda_4 \cdot t) = 0. \quad (20)$$

Accordingly, in the first period of oscillations, the coordinate of the wheel axis will correspond to the dependences (8) at time

$$t_1 = \lambda_4^{-1} \cdot (\pi - \arctg(\lambda_3^{-1} \cdot \lambda_4)). \quad (21)$$

The maximum displacement of the wheel axis from the center of the tire patch corresponds to the first extremum of function (19). This extremum corresponds to zero for the first derivative of function (19).

$$\begin{aligned} \dot{x} = P_\alpha \cdot c^{-1} \{n \cdot e^{-n \cdot t} \cdot (1 - x_0 \cdot t \cdot P_\alpha^{-1}) \cdot [\cos(\lambda_4 \cdot t) + n \cdot \lambda_4^{-1} \cdot \sin(\lambda_4 \cdot t)] \\ - e^{-n \cdot t} \cdot (1 - x_0 \cdot c \cdot P_\alpha^{-1}) \cdot [-\lambda_4 \cdot \sin(\lambda_4 \cdot t) + n \cdot \cos(\lambda_4 \cdot t)]\} = 0. \end{aligned} \quad (22)$$

After the transformations, the condition of the extremum of function (19) takes the form

$$P_\alpha \cdot c^{-1} \cdot \omega^2 \cdot \lambda_4^{-1} \cdot e^{-n \cdot t} \cdot (1 - x_0 \cdot c \cdot P_\alpha^{-1}) \cdot \sin(\lambda_4 \cdot t) = 0. \quad (23)$$

The left-hand side of Eq. (23) is a harmonic function that will be zero provided

$$\lambda_4 \cdot t = \pi \cdot i. \quad (24)$$

Accordingly, the first extremum of maximum value will occur at time

$$t_* = \pi \cdot \lambda_4^{-1}, \quad (25)$$

and the value of the first extremum of function (19) determines the expression

$$x(t_*) = P_\alpha \cdot c^{-1} \cdot (1 + e^{-n \cdot t_*}). \quad (26)$$

Expressions (13), (16), (19) reflect the change in the coordinate of the linear motion of the wheel axis after braking the vehicle on the uphill with a fixed tire patch, taking into account damping in the tire. These expressions correspond to various combinations of tire characteristics and vehicle mass. In the event of damped oscillations of the vehicle axis relative to the fixed center of the tire patch, dependence (21) determines the time moments corresponding to the position of the stationary point of the wheel axis, and expression (25) determines the time points corresponding to the maximum deviations of the axis from the fixed center of the patch.

4 Analysis of the Parameters of the Rectilinear Motion of the Axis of the Locked Wheel After Braking the Vehicle on the Uphill

To analyze the influence of tire parameters on the dynamics of the wheel axis after braking the vehicle on the uphill with a fixed patch of tire contact and a non-rotating brake wheel, the Simulink environment was used. When modeling, it was accepted: the mass of the vehicle is 1500 kg, the tire 155R13 is installed on all four wheels. Tire parameters [12]: longitudinal stiffness coefficient $c = 162$ N/mm; the coefficient of inelastic resistance of the tire is $k = 5000$ N · s/m. Braking the vehicle before it stops on the uphill causes the tire to deform and determines the initial coordinate x_0 of the wheel axis relative to the center of the fixed tire patch. The value x_0 is determined from the equation of motion of the axis of the wheel (and body) of the vehicle at the time the tire patch stops. Longitudinal deformation of tires is determined by the total force P_x

$$P_x = m \cdot j_T - P_\alpha. \quad (27)$$

where P_x is the total rolling resistance of the braking wheels applied to the axles of the wheels of the vehicle.

Assuming a uniform distribution of the total force P_x between the wheels, having set the vehicle to decelerate on the uphill with an elevation angle α , we determine the initial coordinate x_0 of the wheel axis relative to the center of the fixed tire patch

$$x_0 = P_x \cdot c^{-1} \cdot n_w^{-1}, \tag{28}$$

where $n_w = 4$ - the number of brake wheels of the vehicle.

The results of modeling the rectilinear motion of the wheel axis relative to the center of the fixed tire patch are presented in Fig. 4. The curves 1 and 2 can be used to evaluate the effects of vehicle deceleration when stopping on the uphill on the dynamic behavior of its wheel axis. An analysis of Eq. (25) shows that the initial coordinate of the x_0 axis can have a positive and negative value depending on the ratio of the uphill of the road and the braking coefficient $z = j_T/g$.

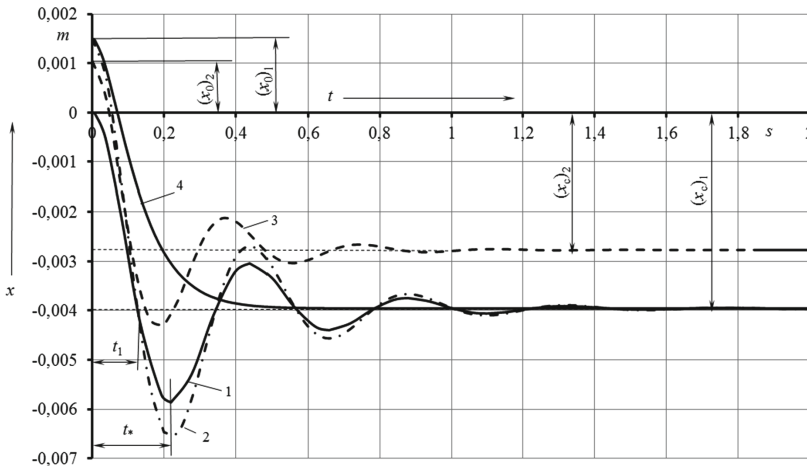


Fig. 4. The movement of the axis of the braked wheel relative to the center of the fixed tire patch on the longitudinal uphill of 5°

Curve 1 reflects the change in the coordinate of the wheel axis in the case of a vehicle stop with a slowdown $j_T = g \cdot \sin\alpha$, which is characterized by the absence of longitudinal force and longitudinal deformation of the tire x_0 at a given time. For the uphill of the supporting surface of the road of 5° , this case is observed when the vehicle slows down $j_T = 0.855 \text{ m/s}^2$. And curve 2 is obtained for the case of a vehicle stop with a deceleration of 1.5 m/s^2 . It should be noted that the results are obtained for braking with four wheels, and holding the vehicle on the uphill is performed by two wheels of a non-driving axle. Obviously, the time to reach the position of the stationary point and the extremum of the amplitude, as well as the coordinate of the stationary point, do not depend on the level of deceleration during braking. A decrease in the mass of the vehicle determines a decrease in the amplitude of the oscillations, but the oscillatory process is retained in view of the low damping ability of the tire.

The movement of the wheel axis relative to the fixed tire patch acquires an aperiodic character only when the tire inelastic drag coefficient $k = 22000 \text{ N} \cdot \text{s/m}$ (curve 4 of Fig. 4). When the vehicle stops with a deceleration of 1.5 m/s^2 on the rise with a different uphill, the value of the initial coordinate can take both positive and negative values (Fig. 5). This is explained by a change in the ratio of the constituent components of the right-hand side of dependence (27). With an increase in the uphill of the road, the value of the initial coordinate x_0 of the wheel axis decreases. It should be noted that the coordinate of the stationary point is uniquely determined only by the value of the uphill of the road surface. Moreover, the parameters characterizing the oscillatory process at any uphill have the same values.

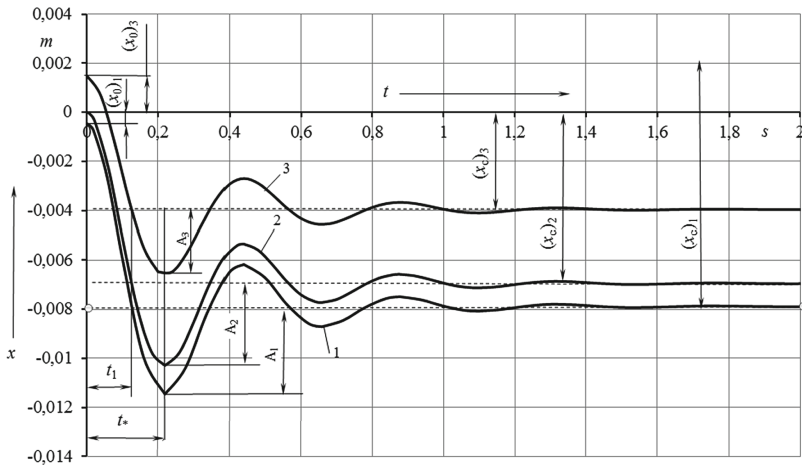


Fig. 5. The movement of the axis of the braked wheel relative to the center of the fixed of the tire patch when the mass of the vehicle is 1500 kg on the rise with different uphill: 1 - with uphill of 10° ; 2 - with uphill of 8.8° ; 3 - with uphill of 5° ; A - amplitude

5 Conclusion

The developed mathematical model of the rectilinear motion of the axis of the locked wheel after braking the vehicle on the uphill allows us to evaluate the change in the coordinate and speed of the axis taking into account damping in the tire. The results of a theoretical study characterize the initial conditions for starting a vehicle on the uphill. The direction of the elastic force of the deformed tires coincides with the direction of the coordinate x -axis, which makes it possible to use the elastic strain energy of the braked tires to improve the energy parameters of the start of movement on the uphill.

The presented mathematical model of the rectilinear motion of the axis of the locked wheel after braking the vehicle on the uphill can serve as a theoretical basis for the formation of the structure of automatic driver assistance systems at the beginning of movement, analysis and synthesis of control algorithms in this mode. Taking into account changes in the coordinate and speed of the axis, the braked wheels at various

dynamic conditions of the vehicle at the beginning of the movement on the rise allows to more correctly assess the dynamic and energy parameters of the beginning of movement, to clarify the requirements for the algorithm of automatic brake control systems of the vehicle and, as a result, to improve the quality of process control.







References

1. Nghia, L.: Feedback in automated clutch control circuit for truck start-up process. *Sci. Tech.* **17**(5), 421–431 (2018). <https://doi.org/10.21122/2227-1031-2018-17-5-421-431>
2. Yingchun, L., Lifu, L.: Fuzzy fractional order sliding mode control for automatic clutch of vehicle AMT. *Int. J. Smart Home* **9**, 53–68 (2015). <https://doi.org/10.14257/ijsh.2015.9.2.05>
3. Turenko, A.N., Shuklinov, S.N.: Adaptivnoe tormoznoe upravlenie kolesnyh mashin [Adaptive braking control of wheeled vehicles]. *Zhurnal avtomobilnyh inzhenerov* **5** (64), 18–21 (2010)
4. Tarasik, V.P.: Matematicheskoe modelirovanie pryamolinejnogo dvizheniya avtomobilya [Mathematical modeling of rectilinear motion of a car]. *Vestnik Belorussko-Rossijskogo universiteta* **2**(55), 119–130 (2017)
5. Zaharik, Yu.M., Zaharik, An.M.: Algoritm elektronnoj sistemy ZS, isklyuchayushej skatyvanie avtomobilya pri troganii na podeme. [Algorithm of the electronic system ZS, excluding rolling the car when starting off on the rise]. *Avtomobilnaya promyshlennost* **2**, 14–17 (2006)
6. Mikhalevich, M., Yarita, A., Leontiev, D., Gritsuk, I., et al.: Selection of rational parameters of automated system of robotic transmission clutch control on the basis of simulation modelling. *SAE Technical Paper 2019-01-0029* (2019). <https://doi.org/10.4271/2019-01-0029>
7. Mikhalevich, M., Yarita, A., Turenko, A., Leontiev, D., et al.: Assessment of operation speed and precision of electropneumatic actuator of mechanical transmission clutch control system. *SAE Technical Paper 2018-01-1295* (2018). <https://doi.org/10.4271/2018-01-1295>
8. Shuklinov, S.N., Gubin, A.V.: Issledovanie rezhimov tormozhenie-stop-nachalo dvizheniya avtomobilya na uklone [The study of the braking-stop-start of vehicle movement on the uphill]. *Avtomobilnyj transport* **42**, 61–69 (2018)
9. Bulgakov, M., et al.: IOP Conf. Ser.: Mater. Sci. Eng. **776**, 012022 (2020). <https://doi.org/10.1088/1757-899x/776/1/012022>
10. Jiang, D., Huang, Y., Zuo, Z., Li, H.: Drivability improvement control for vehicle start-up applied to an automated manual transmission. *Hindawi. Shock and Vibration*. Article ID 578527, 12 p. (2017). <https://doi.org/10.1155/2017/5783527>
11. Leontiev, D.N., Nikitchenko, I.N., Ryzhyh, L.A., Lomaka, S.I., Voronkov, O.I., Hritsuk, I.V., Pylshchyk, S.V., Kuripka, O.V.: About application the tyre-road adhesion determination of a vehicle equipped with an automated system of brake proportioning. *Sci. Tech.* **18**(5), 401–408 (2019). <https://doi.org/10.21122/2227-1031-2019-18-5-401-408>
12. Knoroz, V.I., Klennikov, E.V., Petrov, I.P., Sheluhin, A.S., Yurev, Yu.M.: Rabota avtomobilnoj shiny [The work of a car tire]. *Transport* (1976)
13. Balakina, E.V., Golubev, T.A., Kozlov, Yu.N.: O neobhodimosti modelirovaniya dinamiki elastichnogo koleasa mashiny s uchetom sostavlyayushih snosov reakcij opornoj poverhnosti [On the need to simulate the dynamics of the elastic wheel of a machine, taking into account the component drifts of the reactions of the supporting surface]. *Vestnik mashinostroeniya* **2**, 16–20 (2018)

14. Balakina, E.V., Zotov, N.M.: Raschet velichiny prodolnogo snosa normalnoj reakcii iz-za uprugih uglovyh deformatsij shiny [Calculation of the longitudinal drift of the normal reaction due to elastic angular deformations of the tire]. *Avtomobilnaya promyshlennost* **4**, 25–27 (2015)
15. Balakina, E.V., Zotov, N.M.: Determination of the mutual arrangement of forces, reactions, and friction zones in the contact zone of an elastic wheel with a solid surface. *J. Frict. Wear* **36**(1), 29–32 (2015)
16. Zotov, V.M., Havronin, V.P., Balakina, E.V., Zotov, N.M.: Uglovaya deformaciya avtomobilnogo koleasa pod dejstviem krutyashego momenta: fizika processa [Torque Angular Deformation of a Car Wheel: Process Physics]. *Izvestiya Nizhnevolzhskogo agrouniversitetskogo kompleksa: nauka i vysshee professionalnoe obrazovanie* **1**(45), 1–9 (2017)



Analytical Methods for Determining the Static and Dynamic Behavior of Thin-Walled Structures During Machining

Sergiy Plankovskyy¹ , Vitalii Myntiuk¹ ,
Yevgen Tsegelnyk¹  , Sergiy Zadorozhnyi² ,
and Volodymyr Kombarov¹ 

¹ National Aerospace University “Kharkiv Aviation Institute”,
17 Chkalova Street, Kharkiv 61070, Ukraine
y.tsegelnyk@khai.edu

² Kharkiv State Aircraft Manufacturing Company,
134 Sumska Street, Kharkiv 61023, Ukraine

Abstract. The flexibility of thin-walled component and presence of a low natural vibrations frequency significantly affects the efficiency and accuracy of the milling, so inevitably arise the problem of determining the details of influence of the deformation of the cutting tool and the natural vibration frequencies. In this paper we construct analytical solutions for these problems for a plate, two opposite edges of which are simply supported, and the other two may be supported by the beam and the plates. Action of the milling cutter modeled by the concentrated force. The results were compared with results obtained by finite element method. The influence of the stiffness of the reinforcing plates on the magnitude of deflection of the loaded plate is studied, and the variation of the deflection depending on the point of application of force is shown. The obtained analytical solutions can be easily used in the software of CNC machines for selecting milling regimes and adjusting of the cutting tool path.

Keywords: Thin-walled component milling · Bending and vibrations · Analytical solution

1 Introduction

Many elements of modern aircraft structures (stiffened panels, spars, frames, ribs, etc.) have thin-walled structures obtained by milling [1, 2]. This manufacturing method avoids the use of additional riveted and bolted joints, which leads to a better quality of airfoils and, more importantly, more smooth field stresses arising from external loads.

As a result of this manufacturing method, the material of the part can be 5–10% of the workpiece material. Removal of a large amount of excess material may take a long time. To reduce it using high-speed milling techniques with CNC [3], thus there are two major problems associated with the fact that the processed thin-walled parts have a small bending stiffness. The first problem (static) consists of deforming the workpiece by the forces of interaction with a cutter, whereby a processed surface shape is different

from the expected [2]. The second problem (dynamic) is associated with low natural frequencies of the workpieces vibrations, which gives rise to self-excited vibrations and consequently increases the roughness of the final part, tool wear and even failure of the equipment [4, 5]. The same problems arise in machining thin-walled parts which are produced or repaired by additive manufacturing [6, 7].

According to [2], approaches to solving these problems can be divided into two groups:

- industrial approaches – development of special clamp systems [8], dampers, online tool condition monitoring, etc. [9];
- computational solutions – modification of cutting tool trajectories [10], feed rate, cutting depth, etc. based on preliminary simulation [11].

Computational solutions shown in most studies [12–16], anyway include CAE simulations, which are based on the finite element method (FEM). At present, this method is the most common due to its versatility. However, its use involves carrying out laborious research on the accuracy and convergence of the numerical solution. Analytical methods are deprived of this drawback, but their use is limited to areas of a simple form with ideal boundary conditions [17, 18]. The main drawback of such solutions is that the idealization of the boundary conditions in the direction of tightening or loosening can significantly affect the final result. In this paper, we obtained analytical solutions for the plate bending and vibrations, taking into account its interaction with surrounding elements. Solutions are based on the “boundary conditions identification” method, according to which the first determined boundary stiffness, then the problem is solved for a selected element with inhomogeneous boundary conditions. This method was tested in [19, 20].

2 Problem Solution Strategy

As the workpiece is considered the divide wall (Fig. 1). At the bottom it connects to the skin, the free edge can be supported by a beam (a beam of rectangular cross-section $H \times B$ is shown on Fig. 1).

The problem is considered only normal to the surface cutting force component P (along axis z) both at up milling, as down milling. Two other components acting in the plane of the plate, have a small effect on its deformation. The concentrated force incorrectly simulates the impact of most types of cutters on the workpiece, but the deflection function obtained as a result of such a solution will be the Green’s function. Therefore, the transition to any other external load (distributed over the surface or along the curve) can be easily implemented, e.g. in [18].

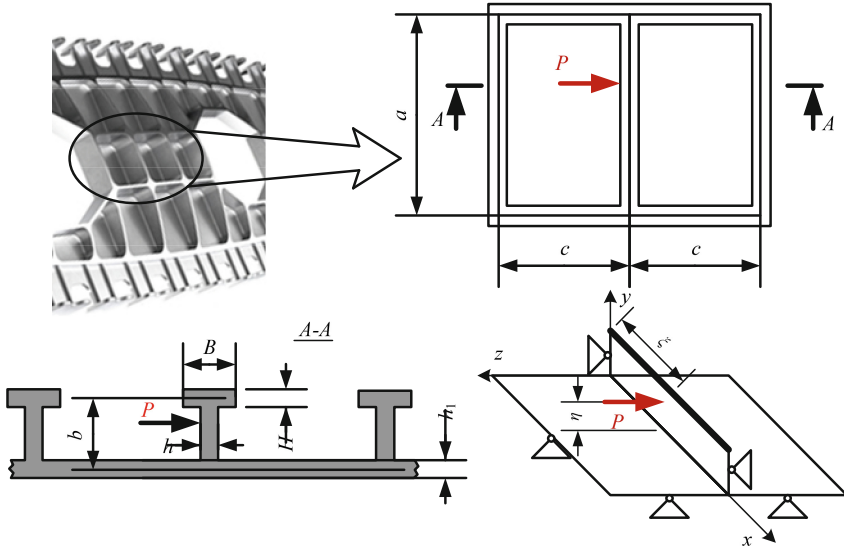


Fig. 1. Dimensions, position and calculated scheme of the workpiece (photo from <https://www.preziss.com/>).

All plates along the boundary are simply supported except boundaries $y = 0$ and $y = b$ (Fig. 2a). For these boundaries are set criteria of elastic support. Simply support for all plates at $x = 0$ and $x = a$ allows using Levy-type solution, i.e., deflection function is defined as

$$w(x, y) = \sum_{n=1}^{\infty} W_n(y) \sin \alpha_n x, \tag{1}$$

where $W_n(y)$ is an unknown function, $\alpha_n = \frac{n\pi}{a}$.

Functions $W_n(y)$ must satisfy to the boundary conditions:

- 1) $W_n|_{y=0} = 0,$
 - 2) $[-D(W_n'' - \mu\alpha_n^2 W_n) - 2K_n W_n']_{y=0} = 0,$
 - 3) $[-D(W_n''' - (2 - \mu)\alpha_n^2 W_n') - K_{wn} W_n]_{y=b} = 0,$
 - 4) $[-D(W_n'' - \mu\alpha_n^2 W_n) - K_{\phi n} W_n']_{y=b} = 0.$
- (2)

Here $D = \frac{Eh^3}{12(1-\mu^2)}$ is the flexural rigidity; E is a Young's modulus; μ is the Poisson's ratio; $K_n, K_{wn}, K_{\phi n}$ are stiffness coefficient.

Below is first determined coefficients $K_n, K_{wn}, K_{\phi n}$, and then construct a function $W_n(y)$.

2.1 Determination of Elastic Boundary

Stiffness coefficients K_{wn} , $K_{\varphi n}$ related to bending and torsional stiffness of the beam [17]. They are equal to the transverse force and torque that arise when a single deflection and a twist angle are applied to the reinforcing beam (Fig. 2b):

$$\begin{aligned} K_{wn} &= \alpha_n^4 EI, \\ K_{\varphi n} &= -\alpha_n^2 GI_k. \end{aligned} \quad (3)$$

Here EI and GI_k is bending and torsional rigidity of beam ($G = E/2(1 + \mu)$ is shear modulus). For a beam of rectangular cross-section $I = BH^3/12$,

$$I_k = B^3 \left(\frac{H}{3} - \frac{64B}{\pi^5} \sum_{m=0}^{\infty} \frac{\text{th}(2m+1)\pi H/2B}{(2m+1)^5} \right).$$

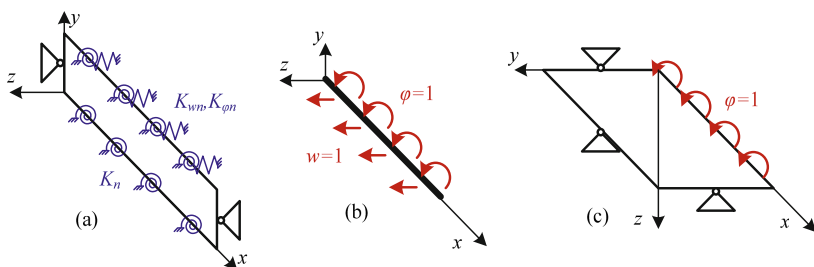


Fig. 2. Determination of elastic boundary.

The flexural rigidity of the plate K_n is defined by solving the boundary-value problem (BVP) for the reinforcement plate (Fig. 2c):

$$\begin{aligned} W_n^{IV} - 2\alpha_n^2 W_n'' + \alpha_n^4 W_n &= 0, \\ 1) W_n|_{y=0} &= 0, & 2) W_n'|_{y=0} &= 1, \\ 3) W_n|_{y=c} &= 0, & 4) [-D(W_n'' - \mu\alpha_n^2 W_n)]_{y=c} &= 0. \end{aligned} \quad (4)$$

Stiffness coefficient K_n will be equal to

$$K_n = M_{yn}|_{y=0} = [-D(W_n'' - \mu\alpha_n^2 W_n)]_{y=0} = -\frac{2D\alpha_n \text{sh}^2\alpha_n c}{\text{ch}\alpha_n c \text{sh}\alpha_n c - \alpha_n c}, \quad (5)$$

here $W_n = \frac{(2c-y)\text{sh}\alpha_n y - y\text{sh}\alpha_n(2c-y)}{2\alpha_n c - \text{sh}2\alpha_n c}$ is the BVP solution (4).

2.2 Bending Problem Solution

Plate deflection is defined as the integral of the differential equation

$$W_n^{IV} - 2\alpha_n^2 W_n'' + \alpha_n^4 W_n = \frac{q_n}{D}, \tag{6}$$

where the external load q_n is the Fourier series expansion coefficient of the Dirac delta function $-P\delta(x - \xi)\delta(y - \eta)$ along the coordinate x : $q_n = -P\delta(y - \eta) \frac{2 \sin \alpha_n \xi}{a}$.

After integrating we finally obtain

$$W_n(x, y, \xi, \eta) = (C_1 y + C_2) \operatorname{ch} \alpha_n y + (C_3 y + C_4) \operatorname{sh} \alpha_n y + \frac{P \sin(\alpha_n \xi)}{\alpha_n^3 Da} H(\eta - y) (\alpha_n (y - \eta) \operatorname{ch}(\alpha_n (y - \eta)) - \operatorname{sh}(\alpha_n (y - \eta))). \tag{7}$$

Here $H(y)$ is the Heaviside step function; C_i is a integrating constants which are determined from the boundary conditions (2) in which the stiffness coefficients $K_n, K_{wn}, K_{\varphi n}$ are computed from (3) and (5).

Note that (7) is a Green’s function, which allows you to easily switch to any other type of external load.

2.3 Natural Vibrations Problem Solution

The fuselage skin plates have much lower bending stiffness than the studied plate, therefore, in the problem of natural vibrations along the boundary $y = 0$, the conditions of simply support are accepted. According to another boundary $y = b$ plate is connected with the beam. To determine the stiffness coefficients, BVP for the beam is first solved

$$\left\{ \begin{array}{l} w^{IV} - \lambda_1^4 w = \frac{q \sin \alpha_n x}{EI}, \\ w|_{x=0} = w|_{x=a} = w''|_{x=0} = w''|_{x=a} = 0, \end{array} \right\} \Rightarrow w = \frac{q \sin \alpha_n x}{EI (\alpha_n^4 - \lambda_1^4)}, \tag{8}$$

$$\left\{ \begin{array}{l} \varphi'' - \lambda_2^2 \varphi = \frac{m \sin \alpha_n x}{GI_k}, \\ \varphi|_{x=0} = \varphi|_{x=a} = 0, \end{array} \right\} \Rightarrow \varphi = -\frac{m \sin \alpha_n x}{GI_k (\alpha_n^2 - \lambda_2^2)},$$

where $\lambda_1 = \sqrt{\omega^4 \frac{\rho F}{EI}}$, $\lambda_2 = \omega \sqrt{\frac{\rho}{G}}$, $\omega = \frac{2\pi}{T}$ is the circular frequency, ρ is the density,

$F = BH$ is the beam cross-sectional area, T is an oscillation period.

After that, the stiffness coefficients are defined as

$$\left\{ \begin{array}{l} \bar{K}_{wn} = EI (\alpha_n^4 - \lambda_1^4), \\ \bar{K}_{\varphi n} = -GI_k (\alpha_n^2 - \lambda_2^2). \end{array} \right. \tag{9}$$

Note that unlike (3) the coefficients (9) are dependent on the total system-wide parameter ω . This parameter is determined from the condition of the existence of a nontrivial solution of the BVP

$$\begin{cases} W_n^{IV} - 2\alpha_n^2 W_n'' + (\alpha_n^4 - \lambda^4) W_n = 0, & 0 < y < b, \\ W_n|_{y=0} = W_n''|_{y=0} = 0, \\ [-D(W_n'''' - (2 - \mu)\alpha_n^2 W_n') - \bar{K}_{wn} W_n]_{y=b} = 0, \\ [-D(W_n'' - \mu\alpha_n^2 W_n) - \bar{K}_{\varphi n} W_n']_{y=b} = 0, \end{cases} \quad (10)$$

which, after some transformations takes the form

$$\begin{aligned} & s_1 \operatorname{th}(s_2 b) \left(\overline{EI} \overline{GI}_k s_3 s_4 - ((1 - \mu)\alpha_n^2 + \lambda^2)^2 \right) \\ & - s_2 \operatorname{tg}(s_1 b) \left(\overline{EI} \overline{GI}_k s_3 s_4 - ((1 - \mu)\alpha_n^2 - \lambda^2)^2 \right) \\ & + 2\lambda^2 (s_3 \overline{EI} \operatorname{th}(s_2 b) \operatorname{tg}(s_1 b) + s_4 \overline{GI}_k s_1 s_2) = 0. \end{aligned} \quad (11)$$

Here $\lambda = \sqrt{\omega^4 \frac{\rho h}{D}}$ and we use the notation $\overline{EI} = EI/D$, $\overline{GI}_k = GI_k/D$, $s_1 = \sqrt{\lambda^2 - \alpha_n^2}$, $s_2 = \sqrt{\lambda^2 + \alpha_n^2}$, $s_3 = \lambda_1^4 - \alpha_n^4$, $s_4 = \lambda_2^2 - \alpha_n^4$.

Transcendental Eq. (11) can be easily solved by the method of bisection. Intervals boundaries, on which the radices are determined by the sequence

$$\left\{ 0, \frac{\pi^2}{a^2 b^2} \left(\frac{\rho h}{D} \right)^{-\frac{1}{2}} \left(\left(m - \frac{1}{2} \right)^2 a^2 + b^2 n^2 \right) \right\}, \quad m = 1, 2, \dots \quad (12)$$

3 Numerical Verification

To confirm the reliability of the results compare them with the results obtained by the finite element method (FEM), implemented in software NASTRAN. Here, a structure with dimensions (see Fig. 1) $a = 100$ mm, $b = 50$ mm, $c = 80$ mm, $h = 4$ mm, $h_1 = 2$ mm, $B = 12$ mm, $H = 5$ mm is used. The material of the workpiece is 2024 aluminum alloy. Its density, Poisson's ratio and elastic modulus are 2700 kg/m³, 0.3 and 72 GPa, respectively.

3.1 Static Deformation

In a numerical experiment, the magnitude of the maximum deflection w_{max} is compared. This deflection will occur at the point of application of force, if the normal component of the cutting force $P = 1$ is applied to the symmetry axis. In calculating such values taken as $x = \xi = 50$ mm and $y = \eta = 50$ mm.

The Fig. 3a shows how the error in the determination of maximum deflection decreases with increasing the additive components in Eq. (1) and the number of nodal unknowns in the FEM. As a reference value in evaluating the analytical solutions (1) taken $w_1 = 2.7124 \cdot 10^{-7}$ value defined at the upper limits of summation in Eq. (1) $N = 20$. For FEM, the reference solution is the $w_{FEM} = 2.7124 \cdot 10^{-7}$ solution obtained

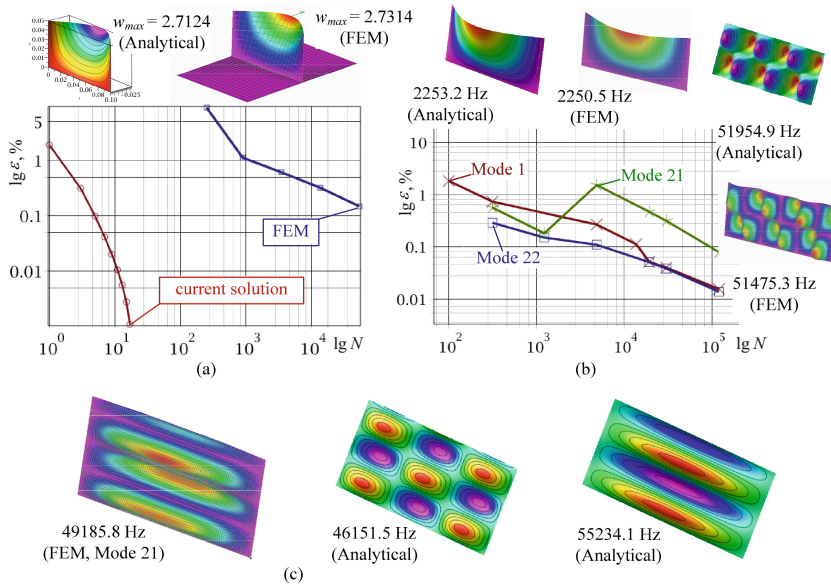


Fig. 3. Verification of the results (a) – static deformations, (b), (c) – modal analysis.

with a finite element size of 1×1 mm. The difference amount of deflection is $(w_{\text{FEM}} - w_1)/w_1 \cdot 100\% = 0.7\%$, which indicates a good matching results.

Note also that to obtain solutions with an accuracy of 1% is enough to confine the three additive components in (1). One first component gives results with an accuracy of 1.9%, and the magnitude of the deflection at the point of application of force can be approximately represented as

$$\begin{aligned}
 w = & 0.0002094 \sin^2(31.42 \xi) [(\eta + 0.1147)\eta \cosh^2(31.42 \eta) \\
 & + 0.7686 (\eta + 0.09801)(\eta + 0.06216) \sinh^2(31.42 \eta) \\
 & - 0.8250(\eta + 0.1437)(\eta + 0.01539) \sinh(62.83 \eta)]. \quad (13)
 \end{aligned}$$

3.2 Modal Analysis

Similar results can be seen in determining the frequencies of natural vibrations. For comparison, the lowest frequency was determined, and the two closest frequency to 50 kHz. Figure 3b shows the convergence of the determination of these frequencies in FEM and the waveforms corresponding to the lowest frequency and the 22nd FEM numbering. The difference in the determination of these frequencies does not exceed 1%. The situation is different with a frequency lower than the closest to 50 kHz. According to calculations, FEM is the frequency of 49185.8 Hz, analytic calculations show 46151.5 Hz and, as shown in Fig. 3c, they have different waveforms. The analytically determined frequency for a plate having a same waveform as in the FEM (Fig. 3c), is 55234.1 Hz.

4 Computational Results

The dependence of the maximum deflection from changing stiffness reinforcement plates: thickness h_1 , width c and length a was studied. The graphs of these curves are shown in Fig. 4a. Initial immutable data are taken from Sect. 3.

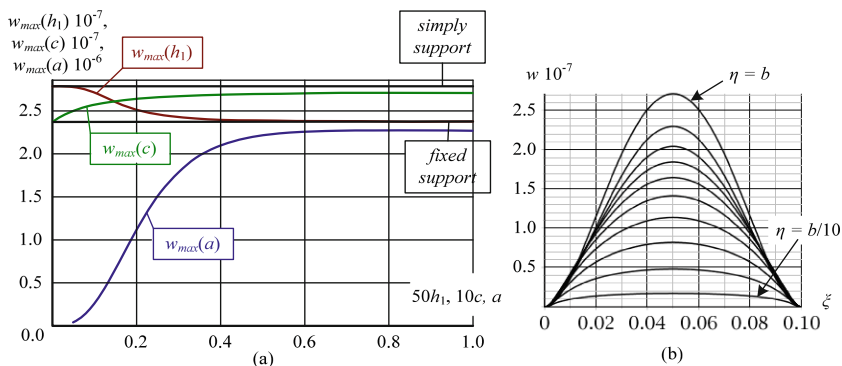


Fig. 4. The dependence of the maximum deflection of: (a) the thickness h_1 , the width c and the length a ; (b) from the point of force application.

As can be seen, in this example, the lower edge of the plate under investigation can accept the fixed support with $h_1 > 9$ mm. When $h_1 < 3$ mm can accept the simple support. In the range from 3 mm to 9 mm adoption of certain conditions can cause errors up to 18%.

At about the same maximum deflection range is changed when changing the width of the support plates c . In this example, the plate width $c > 60$ mm of its reinforcing effect is not changed. The plate can be considered infinitely wide, and for the coefficient of stiffness (5) to take a simpler form

$$K_n = - \lim_{c \rightarrow \infty} \frac{2D\alpha_n \operatorname{sh}^2 \alpha_n c}{\operatorname{ch} \alpha_n c \operatorname{sh} \alpha_n c - \alpha_n c} = 2D\alpha_n \tag{14}$$

Maximum deflection ceases substantially unchanged if the length of the structure reaches a certain value (in this example $a > 60$ mm).

The Fig. 4b shows the deflections calculated at the point of application of force ($x = \xi, y = \eta$), the position of which varies. These graphs show how to adjust the milling path to eliminate distortion caused by the flexibility of the workpiece.

5 Conclusion

The solutions of static deformation and natural vibrations problems of a flexible plate with elastic support are presented. The presented solutions as analytical expressions (see, e.g. Eq. (13)) retain high accuracy. Such solutions can be used when selecting milling regimes and adjusting the path of movement of the cutting tool. The results obtained in this paper can be easily implemented in the software of the CNC machine, which is an advantage over the results obtained by using FEM.






References

1. Herranz, S., Campa, F.J., De Lacalle, L.L., et al.: The milling of airframe components with low rigidity: a general approach to avoid static and dynamic problems. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **219**(11), 789–801 (2005). <https://doi.org/10.1243/095440505X32742>
2. Del Sol, I., Rivero, A., López de Lacalle, L.N., Gamez, A.J.: Thin-wall machining of light alloys: a review of models and industrial approaches. *Materials* **12**(12), 2012 (2019). <https://doi.org/10.3390/ma12122012>
3. Aksonov, Y., Kombarov, V., Fojtů, O., et al.: Investigation of processes in high-speed equipment using CNC capabilities. *MM Sci. J.* **2019**(04), 3271–3276 (2019). https://doi.org/10.17973/MMSJ.2019_11_2019081
4. Arnaud, L., Gonzalo, O., Seguy, S., et al.: Simulation of low rigidity part machining applied to thin-walled structures. *Int. J. Adv. Manuf. Technol.* **54**(5–8), 479–488 (2011). <https://doi.org/10.1007/s00170-010-2976-9>
5. Kononenko, S., Dobrotvorskiy, S., Basova, Y., et al.: Deflections and frequency analysis in the milling of thin-walled parts with variable low stiffness. *Acta Polytech.* **59**(3), 283–291 (2019). <https://doi.org/10.14311/AP.2019.59.0283>
6. Korzhyk, V., Khaskin, V., Voitenko, O., et al.: Welding technology in additive manufacturing processes of 3D objects. *Mater. Sci. Forum* **906**, 121–130 (2017). <https://doi.org/10.4028/www.scientific.net/MSF.906.121>
7. Grzesik, W.: Hybrid additive and subtractive manufacturing processes and systems: a review. *J. Mach. Eng.* **18**(4), 5–24 (2018). <https://doi.org/10.5604/01.3001.0012.7629>
8. Ivanov, V., Dehtiarov, I., Pavlenko, I., et al.: Parametric optimization of fixtures for multiaxis machining of parts. In: Hamrol, A., et al. (eds.) *Advances in Manufacturing II. MANUFACTURING 2019*. LNME, pp. 335–347. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-18789-7_28
9. Jáuregui, J.C., Reséndiz, J.R., Thenozhi, S., et al.: Frequency and time-frequency analysis of cutting force and vibration signals for tool condition monitoring. *IEEE Access* **6**, 6400–6410 (2018). <https://doi.org/10.1109/ACCESS.2018.2797003>
10. Kombarov, V., Sorokin, V., Fojtů, O., et al.: S-curve algorithm of acceleration/deceleration with smoothly-limited jerk in high-speed equipment control tasks. *MM Sci. J.* **2019**(04), 3264–3270 (2019). https://doi.org/10.17973/MMSJ.2019_11_2019080
11. Kolar, P., Sulitka, M., Matyska, V., Fojtu, P.: Optimization of five-axis finish milling using a virtualmachine tool. *MM Sci. J.* **2019**(05), 3534–3543 (2019). https://doi.org/10.17973/MMSJ.2019_12_2019037

12. Rai, J.K., Xirouchakis, P.: Finite element method based machining simulation environment for analyzing part errors induced during milling of thin-walled components. *Int. J. Mach. Tools Manuf.* **48**(6), 629–643 (2008). <https://doi.org/10.1016/j.ijmachtools.2007.11.004>
13. Wan, M., Zhang, W.H., Qin, G.H., Wang, Z.P.: Strategies for error prediction and error control in peripheral milling of thin-walled workpiece. *Int. J. Mach. Tools Manuf.* **48**(12–13), 1366–1374 (2008). <https://doi.org/10.1016/j.ijmachtools.2008.05.005>
14. Chen, W., Xue, J., Tang, D., et al.: Deformation prediction and error compensation in multilayer milling processes for thin-walled parts. *Int. J. Mach. Tools Manuf.* **49**(11), 859–864 (2009). <https://doi.org/10.1016/j.ijmachtools.2009.05.006>
15. Yan, Q., Luo, M., Tang, K.: Multi-axis variable depth-of-cut machining of thin-walled workpieces based on the workpiece deflection constraint. *Comput. Aided Des.* **100**, 14–29 (2018). <https://doi.org/10.1016/j.cad.2018.02.007>
16. Dobrotvorskiy, S., Basova, Y., Kononenko, S., et al.: Numerical deflections analysis of variable low stiffness of thin-walled parts during milling. In: Ivanov, V., et al. (eds.) *Advances in Design, Simulation and Manufacturing II. DSMIE 2019. LNME*, pp. 43–53. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-22365-6_5
17. Timoshenko, S., Woinowsky-Krieger, S.: *Theory of Plates and Shells*, 2nd edn. McGraw-Hill Education, New Delhi (2015)
18. Aijun, T., Zhanqiang, L.: Deformations of thin-walled plate due to static end milling force. *J. Mater. Process. Technol.* **206**(1–3), 345–351 (2008). <https://doi.org/10.1016/j.jmatprotec.2007.12.089>
19. Khalilov, S.A., Myntiuk, V.B.: Postbuckling analysis of flexible elastic frames. *J. Appl. Ind. Math.* **12**(1), 28–39 (2018). <https://doi.org/10.1134/S1990478918010040>
20. Myntiuk, V.B.: Postbuckling of a uniformly compressed simply supported plate with free in-plane translating edges. *J. Appl. Ind. Math.* **14**(1), 176–185 (2020). <https://doi.org/10.1134/S1990478920010160>



Modelling and Simulation of Metal Construction Stress-Strain Behaviour When Designing Road-Building Machines

Oleksandr Rieznikov¹ , Dmytro Klets¹ ,
Anton Kholodov¹ , Leonid Khmara², Mykhailo Kholodov¹ ,
and Natalia Didenko¹ 

¹ Kharkiv National Automobile and Highway University,
25 Yaroslava Mudrogo Street, Kharkiv, Ukraine

d.m.klets@gmail.com, antonkholodov23@gmail.com

² Prydniprovskaya State Academy of Civil Engineering and Architecture,
24a Chernyshevsky Street, Dnipro, Ukraine

Abstract. It is determined that the process of formation of stress-strain behaviour of the motor grader main frame depends on the influence of several factors: soil parameters, parameters and geometry of the machine, parameters of the metal structure itself. The length of the spinal beam makes the main contribution to the stress-strain behaviour formation of the metal structure of the motor grader main frame. The effect of the angle of spinal beam inclination and the distance from the spinal beam to the sub-engine frame have less effect on the stress-strain behaviour of the motor grader metal structure. The design of main frame box is more practical for motor graders in which planning of the roadbed, or cutting into the ground with the entire blade are the characteristic work operation, and the tubular main frame design is more suitable for motor graders with characteristic eccentric external load application. The method of finite elements determines the zones of action of maximum stresses with high degree of precision.

Keywords: Road-building machines · Stress-strain behaviour · CAD systems

1 Choosing Rational Parameters for Metal Structure of Motor Grader Main Frame

For modern earth-moving machinery the improved productivity due to the increase in operating speeds is characteristic. All this ultimately results in increase of the workloads acting on the machine. The experience of operating the middle-class motor graders indicates that an increase in external loads causes the appearance of fatigue cracks in metal structure of the motor grader. This is due to the fact that the dynamic component of the loading of the machine and the parameters of the main frame influencing the formation of its stress-strain behaviour (SSB) are not taken into consideration when designing the load bearing metal structure of the main frame. The

necessity arises for developing a more accurate method of selecting rational parameters for metal structure of a motor grader main frame.

1.1 Papers Review

The problem of calculating metal structure of the motor grader main frame is examined in the works of Kholodov A.M. [1], Sevrov K.P. [2], however, the proposed design schemes do not take into account the dynamic loads acting on the motor grader main frame during operation [3]. In the work [4] it is proposed a model of operation processes of a motor grader engine under unsteady load during technological processes of road construction works. The method which makes possible to take into account the dynamic loads acting on the main frame is considered in the work [5]. In the work [6] the author presents a method for evaluation of the moldboard structure resistance of the motor- graders. The paper [7] proposes a methodology for maintaining and restoration of operating capability of construction and road-building machines. Hydrostatic transmission system modelling and simulation is proposed by Ho and Ahn [8]. However, it should be noted that methods of selecting the rational parameters of motor graders metal construction details require further researches.

1.2 Purpose and Problem Description

The purpose of the work is to develop methods of selecting the rational parameters of metal structure of the motor grader main frame and to estimate loading operating in the motor grader centre girder. In accordance with the stated purpose of the work, the research tasks were determined: to make a quantitative assessment of the influence of geometric parameters of motor grader metal structure on the main frame stress-strain behaviour and to select the rational form of the motor grader metal structure based on the analysis. Scientific novelty: a spatial model of the loading of motor grader main frame metal construction and the formation of its stress-strain behaviour are obtained.

1.3 Problem Solution

Several factors influence the process of forming the stress-strain behaviour of the main frame: the parameters of the soil (the intensity of increase of digging resistance, the specific resistance to digging, the soil density), the parameters and geometry of the machine (the speed of the motor grader, the engine capacity, the geometry of external loads application, the elasticity of pneumatics), the parameters of the metal structure itself (the cross-section of the frame, the length of the spinal beam, the angles of transition of the spinal part of the main frame to the sub-engine one).

Since in the process of the motor grader operation the parameters of the soil and the machine vary over wide range, it is more expedient to select the rational parameters of the metal structure of the motor grader main frame depending on the parameters of the metal structure itself. To quantify the effect of the geometric parameters of the metal structure (Fig. 1) on the main frame stress-strain behaviour the dispersion analysis was performed. From the results of modelling the modes of the motor grader main blade penetration into the ground [5], a group forming the complete factor experiment was

singled out. Table 1 shows the results of calculations for each of the combinations of factors *A*, *B*, *C*, *D*. The following correspondences are established:

1. Factor *A* corresponds to the influence of the angle of inclination of spinal beam β and varies at two levels $A_1 = 0^\circ$, $A_2 = 45^\circ$;
2. Factor *B* corresponds to the influence of the angle of inclination of spinal beam α and varies at two levels $B_1 = 20^\circ$, $B_2 = 80^\circ$;
3. Factor *C* corresponds to distance *h* from the spinal beam to the sub-engine frame and varies at two levels $C_1 = 0.5$ m, $C_2 = 1$ m;
4. Factor *D* corresponds to the length of spinal beam *l* and varies at two levels $D_1 = 1.5$ m, $D_2 = 2.4$ m.

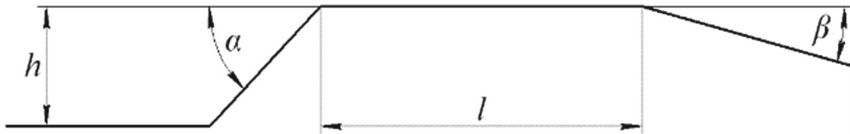


Fig. 1. Geometric parameters of the motor grader metal structure

Table 1. The main stresses acting in the metal structure of the motor grader main frame for each of the combinations of factors *A*, *B*, *C*, *D*

Combinations of factors		<i>D</i> ₁		<i>D</i> ₂	
		<i>B</i> ₁	<i>B</i> ₂	<i>B</i> ₁	<i>B</i> ₂
<i>C</i> ₁	<i>A</i> ₁	64	70	110	120
	<i>A</i> ₂	78	100	130	136
<i>C</i> ₂	<i>A</i> ₁	90	92	118	130
	<i>A</i> ₂	82	87	120	142

With the help of dispersion analysis we will be able to evaluate the effect of geometric parameters of the metal structure of the main frame, highlight the most significant ones and estimate the simultaneous interaction of different factors.

In order to implement the dispersion analysis, a standard method was used [9]. The analysis of the results of the dispersion analysis (see Table 2) indicates that the main contribution to the stress-strain behaviour formation of the metal structure of the motor grader main frame is the length of the spinal beam (factor *D*). It should also be noted that the combination of *ABC* and *ABCD* factors has a significant impact on the formation of stress-strain behaviour. The influence of the angles of the spinal beam inclination (factors *A*, *B*) and the distance from the spinal beam to the sub-engine frame (factor *C*) have a smaller effect on the stress-strain behaviour of the motor grader metal structure by one order on average. In accordance with the results of dispersion analysis, the change in factor *D* and the simultaneous interaction of *ABC* factors should be considered a particularly promising measure to reduce the stress-strain behaviour of the

Table 2. The results of the dispersion analysis

Source of variability	Sum of squares	Average square	Standard deviation
<i>A</i>	410	410	20,2
<i>B</i>	451	451	21,2
<i>AB</i>	39	39	6,2
<i>C</i>	175	175	13,2
<i>AC</i>	390	390	19,7
<i>BC</i>	1	1	1
<i>ABC</i>	2311	2311	48,1
<i>D</i>	7353	7353	85,7
<i>AD</i>	23	23	4,8
<i>BD</i>	15	15	3,9
<i>ABD</i>	10	10	3,2
<i>CD</i>	40	40	6,3
<i>ACD</i>	76	76	8,7
<i>BCD</i>	94	94	9,7
<i>ABCD</i>	2265	2265	47,6
Error	0,00	–	–
Total	9123	–	–

metal structure of the motor grader main frame. Based on the analysis of the obtained results, a step-by-step search for the rational form of the main frame metal structure was made. It was decided to consider 15 different layout forms of metal structures with a circular and rectangular cross-section of the spinal beam.

As a result, the rational shape of the main frame metal structure for the middle class motor graders with a wheel formula of $1 \times 2 \times 3$ was determined (see Fig. 2).

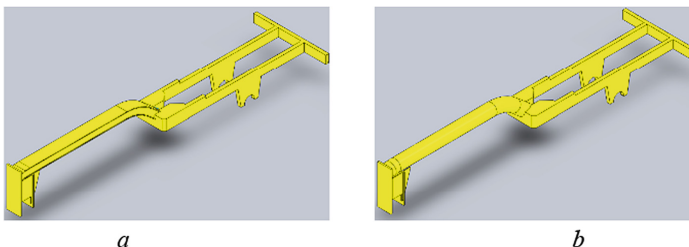


Fig. 2. Rational form of the metal structure of the motor grader main frame: a – for motor graders in which planning the roadbed or cutting into the ground with the whole blade is the characteristic work operation; b – for motor graders with a characteristic eccentric external load application, for example, cutting into the ground with the edge of the blade

So, the process of formation of stress-strain behaviour of the motor grader main frame depends on the influence of several factors: soil parameters, parameters and geometry of the machine, parameters of the metal structure itself.

The length of the spinal beam makes the main contribution to the stress-strain behaviour formation of the metal structure of the motor grader main frame. The effect of the angle of spinal beam inclination and the distance from the spinal beam to the sub-engine frame have less effect on the stress-strain behaviour of the motor grader metal structure.

The design of box main frame is more practical for motor graders in which planing of the roadbed, or cutting into the ground with the entire blade are the characteristic work operation, and the tubular main frame design is more suitable for motor graders with characteristic eccentric external load application.

2 Loading Operating in the Motor Grader Centre Girder

For modern earth-moving machines the growth of productivity due to increasing working speeds is considered typical. This ultimately results in the growth of working loads acting on the machine.

The experience of operating motor grader of the middle class points out that the existing methods of calculating the metal structure of the motor grader main frame is based on considering plain design diagram for the situations of the worst application of the maximum working loads. In calculations on durability these loads are regarded as static ones that do not allow calculating accurately the stresses that act in the frames of modern motor graders. Developing more accurate methods for calculating the loads that act in the motor grader main frame becomes necessary because it is the place where fatigue cracks can be observed.

The *objective* of this section is to analyse the data received experimentally on loading the main frame of the motor grader and to compare them with the calculated values and, in accordance with the experimental research, to propose a design diagram of the main frame that allows determining more precisely the stresses that occur during the performance.

To solve this problem in the process of durability calculations it was decided to consider stress-strain behaviour of the motor grader main frame with the help of the method of the finite elements. Simulation was carried out on the example of a frame made of steel 09G2S. Boundary conditions:

1. Metal structure (the motor grader main frame) was fixed in accordance with the operation conditions of the machine.
2. The load was applied: at the ball pin; at the point of hydraulic cylinders attachment and in the place of the balancing trolley fastening.

The three-dimensional model of the motor grader main frame was built to which the loads corresponding to the working operations of the motor grader dynamic loading were applied. In particular, in the course of calculations the following was realized:

- intensive penetrating into soil with the right edge of the main blade at the angle of gripping 45° ;
- moving at a transportation mode through irregularities of the basic surface.

In the process of calculating the motor grader main frame by the method of finite elements the results shown in Fig. 3 were received. The zones of acting maximum stresses received during calculations coincided with the zones of fatigue crack formation observed in the process of operation (see Fig. 4, Fig. 5), that in general proves the correctness of the calculation method. The experiments were carried out on motor grader DZk-251 in the field conditions on loam of I–III categories. During the experiment the operations of intensive penetration of the main blade into soil and grader moving at transportation speeds through irregularities were realized.

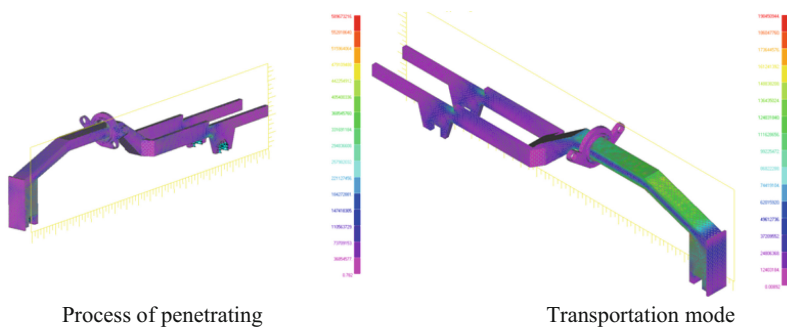


Fig. 3. Equivalent stresses (Pa) on the surface of the main frame model

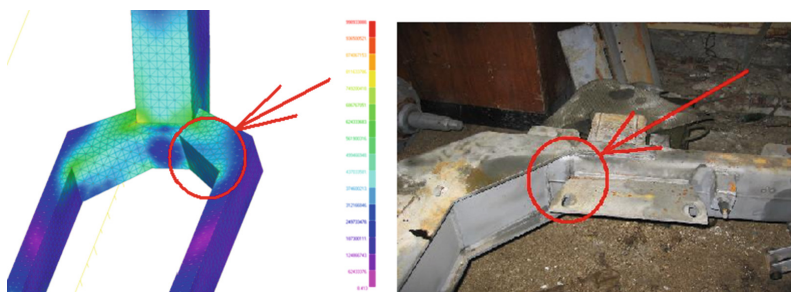


Fig. 4. Crack in the sub-frame

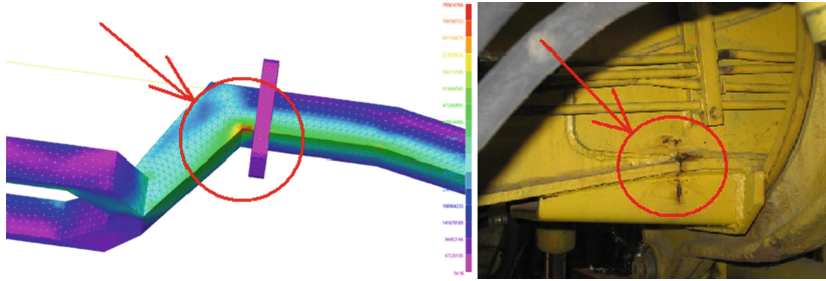


Fig. 5. Crack in the centre girder of the main frame

For transportation mode the artificially produced irregularities of semi-cylindrical shape with the height of 170 mm and the width along the foundation of 310 mm were used. The distance between irregularities made up 1.5 m that corresponds to the base of the balancing wagon of the motor grader.

Comparing maximum values of operating main stresses received during the experiment and during calculation proves their rather high convergence (see Table 3).

Table 3. Comparing maximum values of operating main stresses received during the experiment and during calculation

Operation to be carried out	σ_{\max} (calculated), MPa	σ_{\max} (experimental), MPa
Intensive penetration into soil with the right edge of the main blade at the angle of gripping 45°	120	122,5
Moving at a transportation mode through irregularities of the basic surface	115	99,2

At the same time typical oscillograms that record stresses in the main frame of a motor grader while performing working operations have clearly non-stationary character of changing in time (see Fig. 6).

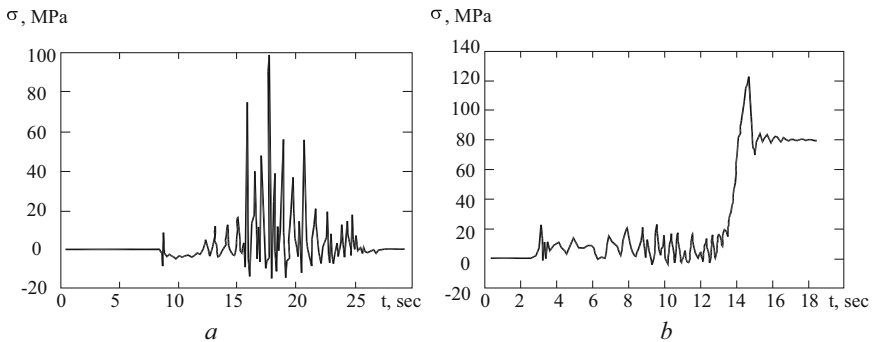


Fig. 6. Graphs of changing main stresses in time: a – Moving through irregularities of the right basic surface; b – Intensive penetrating of the edge of the main blade into soil

The received experimental data point out that the process of loading the main frame has the oscillating character. The existing methods of calculating stresses that act in motor grader metal structures do not take into consideration the oscillating processes that occur during operation. Therefore, developing methods for calculating maximum stresses in the motor grader frame becomes necessary.

3 Conclusions

Based on comparison of the given calculated and experimental data the following *conclusions* can be made:

- The method of finite elements determines the zones of action of maximum stresses with high degree of precision.
- Fatigue cracks appear in the calculated zones in real conditions of operation that is the result of acting variables during loading.
- The process of formation of stress-strain behaviour of the motor grader main frame depends on the influence of several factors: soil parameters, parameters and geometry of the machine, parameters of the metal structure itself.
- The length of the spinal beam makes the main contribution to the stress-strain behaviour formation of the metal structure of the motor grader main frame.
- The effect of the angle of spinal beam inclination and the distance from the spinal beam to the sub-engine frame have less effect on the stress-strain behaviour of the motor grader metal structure.
- The design of main frame box is more practical for motor graders in which planning of the roadbed, or cutting into the ground with the entire blade are the characteristic work operation, and the tubular main frame design is more suitable for motor graders with characteristic eccentric external load application.







References

1. Kholodov, A.: Earth Moving Machinery Designing. High school, Kharkiv (1986)
2. Sevrov, K.: Motor Graders. Design, Theory, Calculation. Mechanical Engineering, Moscow (1970)
3. Khmara, L., Kholodov, A.: Increasing the efficiency of the operation of the batch-type earth-moving machines at the expense of using the pneumatic storage system. In: Gašić, M., Conference Heavy Machinery 2014, HM 2014, Kraljevo, pp. 23–28 (2014)
4. Klets, D., Krasnokutsky, M., Hatsko, V., Barun, M.: Modeling of operation processes of a motor grader engine during work under unsteady load. East. Eur. J. Enterp. Technol. **4/7**(88), 45–50 (2017)
5. Shevchenko, V., Reznykov, A.: Dynamic model of formation of stress-strain behavior of motor grader main frame. Bull. KhNAHU **57**, 112–116 (2012)
6. Debeleac, C.: Evaluation of the moldboard structure resistance of the grader equipment. Eftimie Murgu **XX**(2), 171–176 (2013)

7. Maximenko, A., Zarovchatskaya, E., Maslovskaya, S.: Determination of main output parameters for hydroficated construction and road-building machines at operational stage of their life cycle. *Nauka i Tehnika* (5), 60–66 (2015)
8. Ho, T.H., Ahn, K.K.: Modeling and simulation of hydrostatic transmission system with energy regeneration using hydraulic accumulator. *J. Mech. Sci. Technol.* **24**(5), 1163–1175 (2010)
9. Vynarsky, M.S., Lur'e, M.V.: *Planning the Experiment in Engineering Research*. Technics, Moscow (1975)



Transport Category Airplane Fuselage Master Geometry Parametrical Modeling Method

Anton Chumak¹  , Liliia Buival¹ , Andrii Humennyi¹ ,
Oleksandr Grebenikov¹ , and Dmytro Konyshchuk² 

¹ National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine
a. chumak@khai.edu

² State Company “ANTONOV”, Kiev, Ukraine

Abstract. The analysis of transport category airplane fuselage design features and master geometry modeling methods showed that the methods should be updated to provide quicker automated editing of the model as well as improvement of its quality and accuracy. Transport category airplane fuselage master geometry parametrical modeling method was developed basing on specification requirements, general view drawing, generalised theoretical drawing and fuselage parameters matrix. The method takes in account main geometrical parameters and allows their variation in wide range with automatic model update for providing required characteristics. Proposed method is verified using data of airplanes made by SC “Antonov”. Fuselage parameters influence on its volume and washed surface area was studied. Fuselage master geometry parametrical modeling method implementation in frontend design of passenger airplane for local airlines allowed quick geometrical characteristics determination on early design stage and provide meeting the requirements of specification and aviation regulations.

Keywords: Fuselage · Master geometry · Parametrical modeling · Surface mathematical model · Transport category airplane

1 Introduction

Aircraft configuration development, its geometric parameters coordination and description are among of the most important tasks of preliminary design stage [1, 8]. Fuselage is one of the major airplane’s units. Its size and shape closely interconnected with airplane Technical Requirements Specification (TRS).

Accordingly to the methodology of aircraft integrated design and modeling [1], the master geometry model is one of the airplane configuration automated development results. Traditional descriptive geometry methods of fuselage surface mathematical modeling [1, 10] involve its outlines definition and shape alignment in two stages. At first stage, fuselage shape is designed using graphical methods, at the second stage analytical methods are used for the shape description and mathematical modeling. Fuselage nose and tail sections shape are set using table method, by points coordinates or by power function equation [1] with consequent interpolation by conic or spline curves.

This approach results significant labor and time costs, low accuracy and involves possible human factor errors. The impact of such errors can be significant for the further project development. In addition, this approach makes difficult later edits because it requires repetition of all the works based on theoretical models.

In modern open information technology world, integrated design systems developers provided the tools for parametric modeling of airplane parts, assemblies and units using analytical geometry methods while preserving the complete history of operations [1, 3, 9, 11]. Parametric modeling requires additional formalization of the initial data [3], which leads to some additional time cost of modeling stage, but it provides higher accuracy and associativity which reduces human factor errors and enables automated rebuild, so changes at subsequent design stages can be done quickly and automatically taken in account for all models [3, 4, 12]. In addition, parametric model editing automation allows speeding up studies and expanding the scope of researches using numerical methods of engineering analysis. As result, the most rational geometric parameters of the fuselage can be found and efficiency of the airplane increased.

Thus, parametric modeling allows us to solve the problem of definition, coordination and description of fuselage mathematical model parameters at qualitatively new level, reducing the number of stages and iterations in the design of the airplane and increasing its efficiency.

2 The Aim of the Study

The aim of the study is to develop a transport category airplane fuselage master geometry parametric modeling method basing on its generalized theoretical drawing and geometric parameters matrix, which can be edited in given ranges.

Tasks:

1. Analysis of transport category airplane fuselage requirements and its design features.
2. Development of fuselage generalized theoretical drawing and geometric parameters matrix.
3. Creating fuselage master geometry parametric model in Siemens NX integrated design system.
4. Determination of the model geometric characteristics in given parameters ranges, results verification.

3 Features of Fuselage Master Geometry Parametrical Modeling Method

There are number of requirements to fuselage to be considered during design and parametric modeling of transport category airplane [2, 11]. The main of them are:

- accommodation of payload crew and equipment prescribed by the airplane technical specifications;

- efficient structural integration of aircraft units providing minimum mass, high reliability and maintainability of airframe;
- high quality of external surface providing low aerodynamic drag;
- compliance with the requirements of regulatory documents (AP-25, CS-25, ICAO SRP A6, AC-25, OST), particularly in placement of passengers, crew and a view from the cockpit.

The parametric model of fuselage master geometry should provide [3–5]:

- a single source of information about the fuselage surface configuration for further aircraft design, preparation for its production and operation;
- fuselage parameters and its surface shape associative coupling;
- high surface quality, the absence of gaps or tangent discontinuities which were unspecified structurally, smooth and streamlined contours.

In this study, a simplified parametric model of fuselage master geometry is described. It is mainly intended for frontend design stage application. It can be used for preliminary analysis of the fuselage and the whole aircraft characteristics, as well as for the exchange of preliminary data on its configuration between developers.

Let's consider transport category airplane fuselage typical functional areas (Fig. 1), their purposes and features.

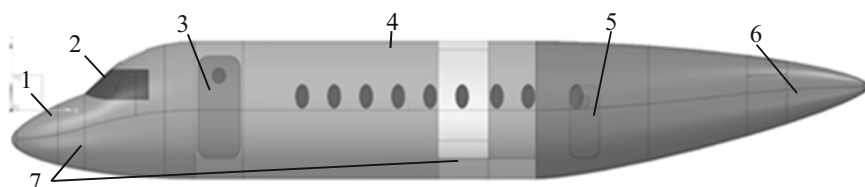


Fig. 1. Fuselage functional areas. 1 – radar dome, 2 – cockpit, 3 – entrance door, 4 – passengers cabin, 5 – cargo door, 6 – auxiliary power unit compartment, 7 – landing gear bays

Most transport category airplanes are equipped with a meteorological radar [6, 8]. It allows to increase flight safety and expand the range of possible operating conditions for the aircraft. To ensure unhindered scanning of the front hemisphere, the meteorological radar compartment is located in the fuselage nose part and it is covered by a radiotransparent dome made of composite materials.

The cockpit purpose is to accommodate the crew, elements of the flight control system and navigation equipment. Cockpit conventional location in fuselage nose section best fits requirements of the field view and ensure the safety of flights. The pilots cabin is separated from the passenger compartment by protective bulkhead that helps to protect the crew in case of possible terrorist attack. As well a cargo compartment bulkhead should effectively protect the crew during an emergency landing. According to ICAO and FAR-25 requirements in most cases the cockpit should be equipped with an emergency hatch.

The location of the entrance doors in the nose and tail sections of the fuselage is typical for medium and heavy aircraft. Landing gear bays usually are not pressurized and located in the nose and the central parts of the fuselage.

In the tail section of cargo planes the ramp type cargo door is often required, which significantly changes the fuselage shape [8, 10]. It is conventional to use the fuselage tail section for auxiliary power unit compartment. In the central and rear parts of the fuselage wing and stabilisers fittings are provided. The design in these areas requires additional reinforcement due to the high intensities of the transmitted loads.

Basing on above mentioned requirements to the fuselage, it's master geometry model and functional features the method of parametrical master geometry modeling method is developed (Fig. 2). It can be used in design of different transport category airplane types in a wide range of take-off masses.

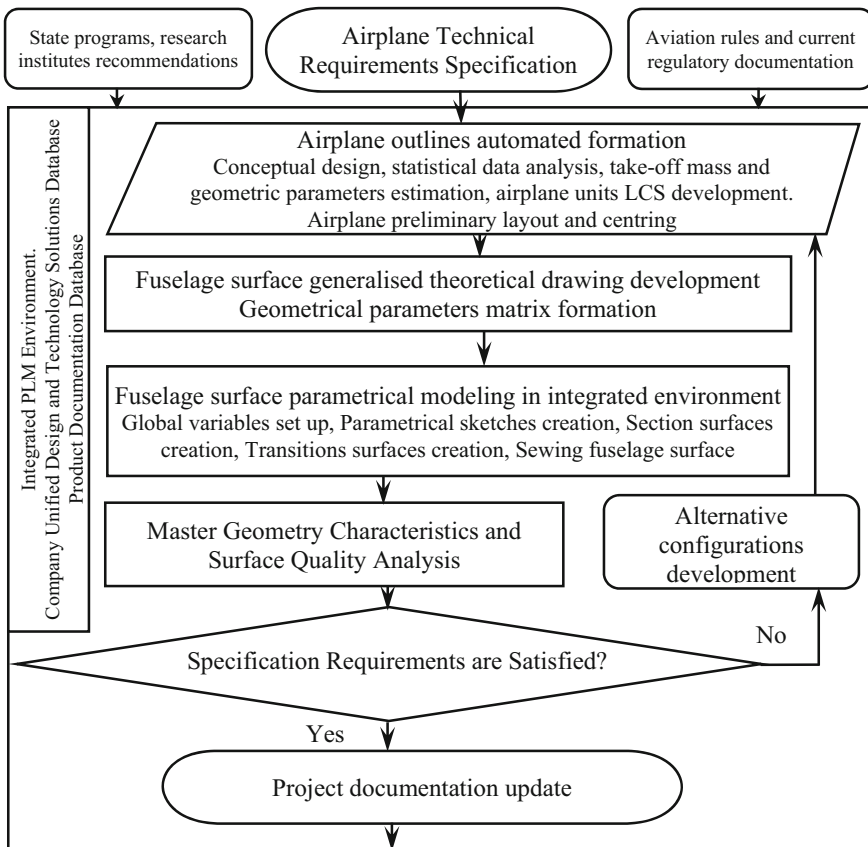


Fig. 2. Fuselage master geometry parametrical modeling method scheme.

4 Fuselage Surface Generalized Theoretical Drawing and Master Geometry Parametrical Model

The initial data for creating the fuselage master geometry are: aircraft design specification, general view drawing, generalized theoretical drawing (Figs. 3, 4) and fuselage geometric parameters matrix (Table 1). Conventionally, the fuselage is divided into the nose section, the central and the tail sections.

Table 1. The matrix of fuselage geometric parameters

Parameter, designation, unit	Value range	
	From	To
Absolute		
Fuselage cross-section equivalent diameter, d_f , mm	3000	4500
Wind screen angle, ϕ_{ws} , deg	40	50
Sight angle, ϕ_v , deg	15	25
Relative		
Nose section relative deflection, y_n	0.22	0.24
Tail section relative deflection, y_h	0.4	0.2
Fuselage fineness ratio, λ_f	7	9.3
Nose section fineness ratio, λ_n	1.4	2
Tail section fineness ratio, λ_h	2.8	3.3
Fuselage cross-section aspect ratio, k_{hb}	1.126	1.175
Cross-section shape coefficient, η_m	0.927	0.78

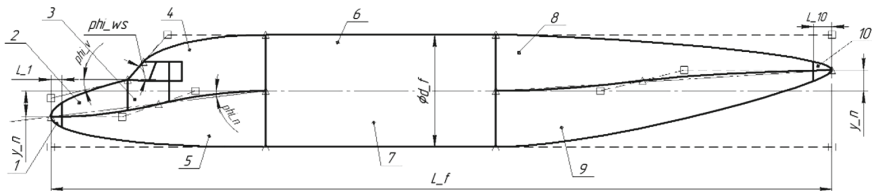


Fig. 3. Fuselage generalized theoretical surface drawing fragment.

The matrix of fuselage geometric parameters includes main absolute and relative geometric parameters which influence on aircraft performance is studied at the considered project stage.

The parameters matrix is represented in the model as global variables expressions. These variables are used in the sketches instead of dimensions values, which subsequently allows unified changing specified parameters, without switching between different layers and sketches.

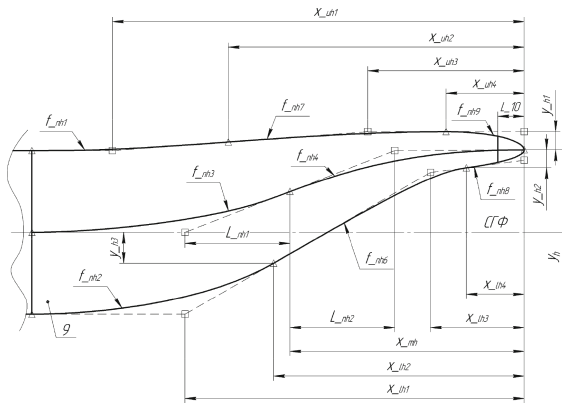


Fig. 4. Fuselage generalized theoretical surface drawing fragment for airplane with ramp type cargo door.

Fuselage geometric parameters are determined during its automated configuration synthesis. In the future design the matrix of geometric parameters is refined and supplemented during internal layout development, optimization of aerodynamic, technical and economic performance of the aircraft.

The diameter and fineness ratio of the fuselage is determined basing on passenger or cargo cabin layout requirements and condition of providing minimal aerodynamic drag.

For modern passenger airliners its conventional to use cylindrical fuselage shape, since this shape works the best for pressurized cabins, which is decisive for cruising altitudes above 10,000 m. Large airplane pressurized fuselages may have two or three decks then cross-section shape consists of arcs combination. Light transport turboprop aircraft may have not pressurized cabin then the fuselage cross-section shape close to rectangular is usually used which provides the best use of the internal space. For medium and heavy cargo planes with unpressurized cabins sometimes elliptical cross-section is used. For other than round fuselage cross-sections equivalent (same area) cross-section diameter should be found.

Nose section is greatly influenced by ergonomic requirements, which could be considered by placing virtual pilot mannequins with reference geometric parameters. The type of control system and command levers also affects the overall dimensions of the pilots' cabin. Stabilizers configuration and rear cargo door define the shape of the tail section.

Parameters of fuselage nose and tail section features such as cockpit windscreen, ramp, fairings or cargo doors, are determined accordingly to requirements of field of view, bird strike resistance, minimum weight and drag of the fuselage. To build a parametric model, the data of the fuselage geometric parameters matrix should be supplemented with derivative, reference and additional parameters. These parameters are defined mainly by fuselage surface smoothness and continuity considerations, their values are linked to the parameters presented in the matrix, and their effect on aircraft performance is not significant.

One of the military cargo airplane features is configuration of the fuselage tail section with a ramp type cargo door. Consideration of providing required height of the cargo doorway opening and at the same time reduce its length leads to increase in the tail section deflection angle and the use of the characteristic S-shape of the surface upper and lower guides. A fragment of the cargo plane tail section theoretical drawing is shown on Fig. 4.

Typically theoretical surfaces of fuselage nose and tail section have double curvature. To create 2, 4, 5, 6, 7, 8, 9 segments (Fig. 4) Section Surface Siemens NX feature is used [12], which provides high precision, well controlled cross-sections shape and flexibility in setting their parameters. Each of these segments surfaces are formed by movement of the conic section curve along the guides and apex which are created as sketch curves.

To build 1, 3, 10 segments Through Curve Mesh feature is used, which provides converging and transition surfaces creating. The outer surfaces of doors, access doors and nose landing gear bay doors fit into the fuselage theoretical contour.

Discriminants of the generator curves determine the fuselage cross-section shape, in the proposed model the values are specified for the initial and final sections of the segments which could be interpolated by linear or cubic laws. A section generator discriminant should correspond to cross-section shape coefficient sent by the parameters matrix (Table 1).

To complete the fuselage master geometry model building it is necessary to mirror the created surfaces (Mirror Feature) and sew segments into one surface (Sew).

If the model is created correctly and resulting surface is continuous, the sewing operation will result fuselage solid model. This model can be used in further engineering analysis applications. Fuselage model volume and washed surface area can be measured using the Measure Bodies command.

The quality of the master geometry model mostly depends on continuity and smoothness of the surface. Continuity is provided by using the same curves by different connected segments of the surface. Smoothness of the model surface can be checked visually using reflections of parallel lines. Continuity and smoothness on the reflections approve surface smoothness. Surface smoothness can be checked on the edges of the segments as first G1 (tangency) and second G2 (curvature) derivatives of its guides. Corresponding tools are available within Analysis menu of Siemens NX.

5 Testing and Results

The method was tested using data of An-1X8 transport category airplanes family, as well as in frontend designing of new airplane for local airlines. Side views of created fuselage models presented on the Fig. 5 in unified scale.

As result of the testing, the possibility of automated model editing in given ranges of geometric parameters values is confirmed.

Influence of geometric parameters on the characteristics of the model, in particular on its volume and the area of the washed surface, is studied. Obtained graphs are shown in Fig. 6.

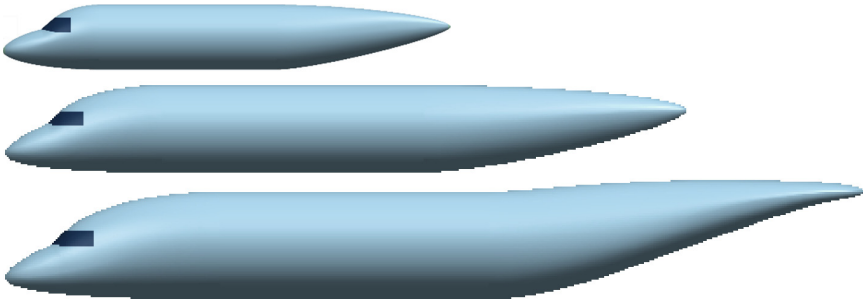


Fig. 5. Fuselage master-geometry parametric models.

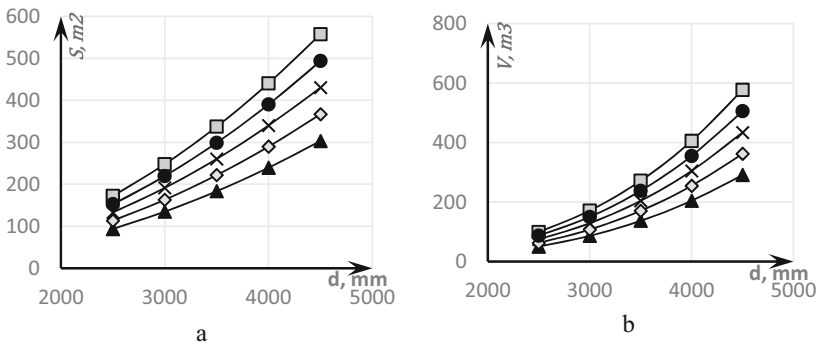


Fig. 6. Fuselage diameter and fineness ratio influence on model characteristics: a – washed surface area, b – internal space volume

□ – $\lambda_f = 6$; ● – $\lambda_f = 7$; × – $\lambda_f = 8$; ◇ – $\lambda_f = 9$; ▲ – $\lambda_f = 10$.

Obtained results were approximated by linear dependencies. Mathematical model based on cylinder geometric characteristics was formed. The model is describing dependency of fuselage volume (V) and washed surface (S) from its diameter, fineness ratio, nose and tail sections fineness ratios:

$$S = \pi d^2 \lambda_f (1 - 0.033(\lambda_n + \lambda_h));$$

$$V = \frac{\pi}{4} d^3 \lambda_f (1 - 0.051(\lambda_n + \lambda_h)).$$

Accuracy of the model is estimated by comparison of the results to characteristics of existing airplanes fuselages, error is shown for correspondent data in Table 2 in braces. Errors caused by simplifications of generalized drawing are not exceeding 5%. This result is acceptable for frontend design stage.

Table 2. Resulted model geometrical properties

Meter, designation, unit	Airplane model		
	X-174	An-1x8 Passenger	An-178 Cargo
Fuselage diameter d_f , mm	2300	3350	3900
Fuselage fineness ratio, λ_f	7	8,2	7,69
Nose section fineness ratio, λ_n	1.9	1,65	1,8
Tail section fineness ratio, λ_h	3	3,15	3
Fuselage length, L, mm	16100	27470 (0.11%)	30000 (0.11%)
Fuselage washed surface, S, m ²	98	243 (0.69%)	299 (0.24%)
Fuselage internal volume, V, m ³	50	183 (0.35%)	271 (4.06%)

In design of the local airliner for 19 passenger seats, the fuselage diameter $d_f = 2300$ mm is chosen, which is the minimum diameter providing required passage width between the rows of passenger seats arranged in three rows in an economy class layout [7].

6 Conclusion

Transport category airplane fuselage master-geometry parametrical modeling method is developed. It provides definition coordination and description of fuselage mathematical model by means of up to date integrated design systems accordingly to the requirements of airplane design specification and regulatory documents. Better accuracy and editing automation allow reducing the number of design iterations in comparison to traditional methods of descriptive geometry and improve usability of the model in parametrical engineering analysis applications for aerodynamic and mass characteristics study and optimization on early design stages.

Parametrical master-geometry model represents main geometric features of the transport category aircraft fuselages (fuselage fineness ratio, nose and tail section fineness ratios, fuselage cross-section shape, cockpit windscreen, nose and tail axis deflections) and ensures the smoothness of the aerodynamic surface and the possibility of solid model creating.

The method was tested using data from the An-1x8 transport category aircraft family, as well as in the process of local airliner frontend design. Obtained results correspond to the data of prototype with good accuracy, acceptable for considered design stage.








The proposed method allows changing the model automatically in a given range of geometric parameters, which make it useful in further parametric studies of the aerodynamic, strength and mass characteristics of the fuselage with aid of computer engineering analysis systems.

References

1. Grebenikov, A.G.: Metodologiya integrirovannogo proyektirovaniya i modelirovaniya sbornykh samoletnykh konstruksiy. Methodology of integrated design and modeling of assembled aircraft structures. National aerospace university Kharkov Aviation Institute, Kharkov (2006)
2. Aviation Rules. Airworthiness standards transport category aircraft. Part 25. IAC (2014)
3. Balabuev, P.V., Matusevich, V.I.: Strategija i praktika ANTK «Antonov» v sozdanii samoletov «AN» na osnove polnogo jelektronnogo opredelenija izdelija. Strategy and Practice of AN Airplanes Creation Based on Full Electronic Product Definition. Informacionnye tehnologii v naukoemkom mashinostroenii: Komp'juternoe obespechenie industrial'nogo biznesa, pp. 84–97 (2001)
4. Matusevich, V.I., Bojko, Ju.R.: Koncepcija i plany kompleksnogo reshenija zadach avtomatizirovannogo proektirovaniya, tehnologicheskoy podgotovki i upravlenija samoletoostroitel'nym proizvodstvom. The concept and plans for a comprehensive solution of the computer-aided design, technological preparation and management of aircraft manufacturing problems. Tehnologicheskie sistemy **1**, 77–82 (1999)
5. Bratuhina, A.G., Ivanova, Ju.L.: Sovremennye tehnologii aviastroenija. Modern aircraft manufacturing technology. Mashinostroenie, Moscow (1999)
6. Geremes, Ju.N., Grebenikov, A.G., Gumennyj, A.M.: Koncepcija sozdaniya passazhirskogo samoljota dlja mestnyh vozdušnyh linij. The concept of passenger aircraft for local airlines creating. Otkrytye informacionnye i komp'juternye integrirovannye tehnologii **47**, 20–33 (2010)
7. Grebenikov, A.G., Gumennyj, A.M., Serdjukov, A.A., Chumak, A.S.: Optimizacija geometricheskikh parametrov samoleta mestnyh vozdušnyh linij. Optimization of local airlines aircraft geometric parameters. In: Mizhnarodna naukoivo-tehnicna konferencija on Problemi stvorennja ta zabezpechennja zhittevoogo ciklu aviacijnoi tehniki, p. 12. National Aerospace University Kharkov Aviation Institute, Kharkov (2011)
8. Torenbeek, E.: Advanced Aircraft Design: Conceptual Design Analysis and Optimization of Subsonic Civil Airplanes. Delft University of Technology, Netherlands (2013)
9. Roskam, J., Anemaat, W.A.: General aviation aircraft design methodology in a PC environment. SAE International (1996)
10. Grebenikov, A.G., Zhelodochenko, V.N., Kobyljanskiy, A.A.: Osnovy obshchego proyektirovaniya samoletov s gazoturbinnymi dvigatelyami. Fundamentals of the general design of aircraft with gas turbine engines. Part 2. National Aerospace University Kharkov Aviation Institute, Kharkov (2003)
11. Kiva, D.S., Grebenikov, A.G.: Nauchnye osnovy integrirovannogo proektirovaniya samoletov transportnoj kategorii. Scientific basis for integrated design of transport category aircraft. National Aerospace University Kharkov Aviation Institute, Kharkov (2014)
12. Goncharov, P.S., El'cov, M.Ju., Korshikov, S.B.: NX dlja konstruktora-mashinostroitelja. NX for the structural engineer. DMK Press, Moscow (2010)



Circular Layout in Thermal Deburring

Sergiy Plankovskyy¹ , Olga Shypul¹ , Yevgen Tsegelnyk¹  ,
Alexander Pankratov² , Tatiana Romanova² ,
and Igor Litvinchev³ 

¹ National Aerospace University “Kharkiv Aviation Institute”, 17 Chkalova Street, Kharkiv 61070, Ukraine

y. tsegelnyk@khai.edu

² A. Pidgorny Institute of Mechanical Engineering Problems of the National Academy of Sciences of Ukraine, 2/10 Pozharskogo Street, Kharkiv 61046, Ukraine

³ Faculty of Mechanical and Electrical Engineering, Nuevo Leon State University, 66450 Monterrey, NL, Mexico

Abstract. Balanced circular layout with a prohibited zone is considered. The problem is motivated by thermal deburring, the modern technology of removing burrs from machine parts by detonating gas mixtures in a cylindrical deburring chamber. To achieve a uniform distribution of thermal and pressure effects, the chamber zone with extremely high temperature/pressure has to be excluded from the feasible layout thus giving rise to the layout with the prohibited zone. Cylindrical parts have to be placed as distant as possible one from another, as well as from the border of the cylindrical container and from the prohibited cylindrical zone. Mechanical stability of the layout is assured by introducing balancing conditions. The overall problem is then reduced to the two-dimensional sparsest circular packing problem subject to the prohibited circular zone and the balancing conditions. Using ANSYS CFX software, a numerical simulation for the temperature/pressure distribution in the chamber is presented. Three cases are considered: the sparsest balanced and unbalanced layouts, as well as the sparsest balanced layout with the prohibited zone. The latter layout is considered to be the most promising for the uniform thermal processing.

Keywords: Thermal effect · Thermal deburring · Circular layout problem · Balancing conditions · Distance constraints · Mathematical model

1 Introduction

Thermal energy method (thermal deburring) [1, 2] is a process of removing production-related burrs from machine parts [3]. The burrs are removed by a chemical reaction between the material and a mixture of combustible gases [4]. Thermal deburring is frequently used in additive technologies (3D printing) for cleaning complex shaped parts from particles of non-sintered powder [5].

The parts are located in a hydraulically closed cylindrical bell-shaped deburring chamber. The gas mixture is then ignited by a spark and a high-pressure zone is created

with the temperature of the subsequent combustion ranging from 2500 to 3300 °C. The burr reaches its ignition temperature and completely combusted in a few milli seconds [4]. Due to shock waves arising from explosive combustion of the gas mixtures, the heat fluxes acting in the chamber can differ significantly [6]. Consequently, the thermal factors have to be taken into account along with balancing geometrical/mechanical conditions.

In this paper the optimized packing circular objects in a circular container (de-burring chamber) is considered. The objective is optimizing uniformity of heat fluxes subject to shock waves. The objects have to be arranged subject to variable location of the prohibited area (high pressure zone).

2 Related Works

Layout problems (also known as packing objects in a container) are often presented in modern science and technology, e.g. in design of aircraft structures [7], adaptive machining with virtual localization [8], laser cutting of sheet materials [9], topological optimization of parts obtained by additive manufacturing (3D printing) [10, 11] to mention a few. In contrast to the dense packing (e.g., [12]), where the objective is to minimize the unused space (waste), in the applications mentioned above the layout is affected by various technological constraints. Consequently, the objects have to be located more distantly one from another than in the dense packing.

Circular packing is an important part of the general layout stream. These problems are NP-hard [13]. The bibliography on regular packing is quite extensive, see, e.g., [14–17] and the references therein. Circular packing problems with balancing conditions are considered, e.g., in [18] using phi-functions for modeling and global/local optimizers to solve arising optimization problems.

Optimized packing objects in a convex bounded domain is considered in [19] subject to constraints on the gravity center position and moments of inertia of the system. The layout is aimed to optimize pairwise distances between the objects and distances between each object and the boundary of the domain. Using the phi-function technique a mathematical model is derived and a solution strategy is proposed for the corresponding non-linear optimization problem.

The concept of the sparsest packing for irregular shaped objects was introduced in [20]. The sparsest packing is aimed to place the objects in the container as distant as possible.

3 Balanced Circular Layout

The circular layout problem is studied in the following formulation. Denote by C a circular container centred at point $(0, 0)$ and having its radius r and weight w_0 . Let a collection of circles C_i having their radii r_i and weights w_i , $i \in I_n = \{1, 2, \dots, n\}$ be given. Assume that the gravity centre of each circle is placed in its centre. Denote by (x_i, y_i) a variable centre of C_i . A circular prohibited area (high pressure zone) is defined by $\hat{C} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 - \hat{r}^2 \geq 0\}$, where \hat{r} is the radius of \hat{C} .

The Euclidian distance between each pair of the circles C_i and C_j is denoted by τ_{ij} for $i > j \in I_n$, while τ_i denotes the distance between the circle C_i and the boundary of the container C and $\hat{\tau}_i$ denotes the distance between the circle C_i and the prohibited area \hat{C} for $i \in I_n$. Let τ denote $\min\{\tau_{ij}, i > j \in I_n, \tau_i, i \in I_n, \hat{\tau}_i, i \in I_n\}$.

A Circular Balanced Layout Problem. Place the collection of circles $C_i, i \in I_n$ into the given container C provided that the gravity centre (x_c, y_c) of the collection of circles be located inside the δ -neighborhood of the container centre and τ will reach its maximum value.

The balanced layout of the collection of circles $C_i, i \in I_n$, can be formulated as the following nonlinear programming problem:

$$\max_{\tau \geq 0, x, y, x_c, y_c} \tau \quad (1)$$

subject to

$$(x_i - x_j)^2 + (y_i - y_j)^2 - (r_i + r_j + \tau)^2 \geq 0, i < j \in I_n, \quad (2)$$

$$x_i^2 + y_i^2 - (r_i + \hat{r} + \tau)^2 \geq 0, i \in I_n, \quad (3)$$

$$(r - r_i - \tau)^2 - x_i^2 - y_i^2 \geq 0, i \in I_n, \quad (4)$$

$$-\delta \leq x_c \leq \delta, -\delta \leq y_c \leq \delta, \quad (5)$$

where $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, $x_c = \sum_{i=1}^n w_i x_i / \sum_{i=0}^n w_i$, $y_c = \sum_{i=1}^n w_i y_i / \sum_{i=0}^n w_i$.

The problem (1)–(5) has totally $2n + 3$ unknown variables τ, x, y, x_c, y_c .

Constraints (2) state that the distance between the circles C_i and C_j has to be at least τ for $i < j \in I_n$. Constraints (3) state the similar condition for the circle C_i and the prohibited circular area \hat{C} , while constraints (4) assure at least τ -distance between the circle $C_i, i \in I_n$ and the boundary of the circular container C . Balancing conditions are represented by (5).

To solve the problem (1)–(5) the approach introduced in [20] is used. The multistart strategy is applied for the non-linear nonconvex optimization problem (1)–(5). A good feasible starting point is obtained by the homothetic transformations of circles and then the local optimization procedure is implemented combined with a decomposition technique [21]. The IPOPT code (<https://projects.coin-or.org/Ipopt>) is used to solve arising nonlinear optimization subproblems. The algorithm returns a local maximal value τ^* and the corresponding solution point $(\tau^*, x^*, y^*, x_c^*, y_c^*)$.

4 Thermal Problem Formulation

In this section basic mathematical tools used for the heat transfer simulation are presented. The main assumptions are as follows:

- heat transfer between the gas (air ideal gas) with an initial temperature of 1000 °C and the circular objects defined by their boundaries with a constant temperature of 25 °C is considered;
- the layout area (container) boundary is set as an adiabatic wall;
- the computing time is limited by the shock-wave attenuation time, which is caused by the pressure pulse at the initial time;
- 2D axisymmetric case is considered;
- Reynolds decomposition is used to simplify Navies-Stokes equations.

A commercial code ANSYS CFX is used for numerical simulations.

The flow of gases is described by the model presented in [6]. Excluding the mass forces, radiation, baro- and thermal diffusion and chemical reaction the corresponding equations are as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \cdot \mathbf{u}) = 0, \quad \frac{\partial (\rho \cdot \mathbf{u})}{\partial t} + \nabla \cdot (\rho \cdot \mathbf{u} \cdot \mathbf{u}) = -\nabla P + \nabla \cdot \boldsymbol{\tau}_{eff},$$

$$\frac{\partial (\rho h)}{\partial t} + \nabla \cdot (\mathbf{u} \rho h) = \frac{\partial P}{\partial t} + \mathbf{u} \cdot \nabla P + \nabla (\lambda_{eff} \nabla T + \boldsymbol{\tau}_{eff} \cdot \mathbf{u}),$$

where ρ is the density; \mathbf{u} is the velocity vector; t is the time; P is the pressure; $\boldsymbol{\tau}_{eff} = (\mu + \mu_t)[\nabla \mathbf{u} + (\nabla \mathbf{u})^T - 2/3 \mathbf{I} \cdot (\nabla \mathbf{u})]$ is the effective stress tensor (i.e., the sum of the viscous and turbulent stresses); μ is the viscosity; μ_t is the turbulent viscosity; \mathbf{I} is the unit tensor; h is the enthalpy of the mixture; $\lambda_{eff} = \lambda + \lambda_t$ is the effective thermal conductivity; λ is the laminar heat conductivity; $\lambda_t = C_p \mu_t \text{Pr}_t^{-1}$ is the turbulent heat conductivity; C_p is the specific heat of the gas; Pr_t is turbulent Prandtl number; T is the temperature.

For calculation of μ_t the SAS SST turbulence model [22] is used.

For convective heat flux calculations the formulation of Kader is applied [23]:

$$T^+ = \text{Pr} \cdot \tilde{y}^+ \exp(-\Gamma) + [2.12 \ln(1 + \tilde{y}^+) + \beta] \exp(-1/\Gamma), \quad (6)$$

where $\beta = (3.85 \text{Pr}^{1/3} - 1.3)^2 + 2.12 \ln(\text{Pr})$; $\Gamma = 0.01(\text{Pr} \cdot \tilde{y}^+)^4 / (1 + 5 \text{Pr} \cdot \tilde{y}^+)$.

In the Eq. (6) dimensionless temperature is used and defined as:

$$T^+ = [\rho C_p \tilde{u}_{\tan}(T_w - T_f)]/q_w, \quad (7)$$

where T_w is the wall temperature; T_f is the temperature of combustion products in the flow core; q_w is the wall heat flux; $\tilde{u}_{\tan} = [(\tilde{u}_{\tan}^{vis})^4 + (\tilde{u}_{\tan}^{\log})^4]^{0.25}$ is the velocity profile in the boundary layer; $\tilde{u}_{\tan}^{vis} = U_1/\tilde{y}^+$ is the velocity in the viscous sublayer; $\tilde{u}_{\tan}^{\log} = U_1/(1/k \cdot \ln(\tilde{y}^+) + c)$ is the velocity in the logarithmic sublayer; y^+ is the dimensionless distance from the first grid node to the wall; $\tilde{y}^+ = \max(y^+, 11.067)$; the limiting value $y^+ = 11.067$ corresponds to the intersection of the logarithmic and the linear profiles; U_1 is the flow rate in the node which is the nearest to the wall.

The value of the heat flux is calculated by following Eq. (8)

$$q_w = [\rho c_p \tilde{u}_{\text{tan}}(T_w - T_f)]/T^+. \quad (8)$$

To set the initial pressure pulse, its distribution in the calculation domain (container) is defined by the expression:

$$P_0 = 1 \text{ [atm]} + 10 \text{ [atm]} \cdot \text{step}((\hat{r}/1 \text{ [mm]})^2 - (x/1 \text{ [mm]})^2 - (y/1 \text{ [mm]})^2), \quad (9)$$

where *step* is the record of the Heaviside step function, adopted in ANSYS CFX; \hat{r} is the radius of the circular prohibited area \hat{C} centered at the origin, which defines a region with high pressure. In the calculations $\hat{r} = 3$ mm is used.

Equation (9) defines the following initial pressure distribution:

$$P_0 = \begin{cases} 1 \text{ [atm]}, & \hat{r}^2 - x^2 - y^2 < 0, \\ 11 \text{ [atm]}, & \hat{r}^2 - x^2 - y^2 > 0. \end{cases}$$

Dimensions of the mesh elements are chosen to provide the value $y^+ \approx 1$, where $y^+ = \rho \Delta y u_\tau / \mu$ is the dimensionless distance from the wall, u_τ is the friction velocity, ρ is the density and μ is the viscosity of gas, Δy is the distance of nearest mesh node from the wall. The mesh is further milled in an area in which the initial time was set high pressure. The mesh is created by the ANSYS mesh generator – plain unstructured tetra with wall layers (15...20 elements in thickness). Since the problem is a two-dimensional total number of elements in the simulation was less than 200000. The time step is chosen to assure that the Curret number is less than 5. The solver settings use a high-resolution advection scheme and the second-order backward Euler transient scheme. The stopping criterion is $P_{\max}/P_{\min} \leq 1.05$.

5 Computational Results

Three realizations of the circular layout problem (1)–(5) are studied:

- *case 1* – the circular layout excluding constraints (3), i.e. *balanced layout without the prohibited area*;
- *case 2* – the circular layout excluding constraints (3) and (5), i.e. *circular layout without the prohibited area and the balancing conditions*;
- *case 3* – the circular layout including constraints (3) and (5), i.e. *circular layout with the prohibited area and the balancing conditions*.

We consider layout of three circles C_i , $i = 1, 2, 3$ with the weights w_i , $i = 1, 2, 3$ proportional to their radii r_i , $i = 1, 2, 3$ inside the container C of radius $r = 50$ mm with the prohibited area $\hat{C}(\hat{r} = 3$ mm) centered at $(0, 0)$.

Results for the optimized layouts obtained for these three cases by the algorithm described in Sect. 3 are summarized in Table 1 and illustrated in Fig. 1.

Table 1. Circular layout obtained by the algorithm from Subsection 3

Input data r_i , mm	Output data		
	x_i^* , mm	y_i^* , mm	τ^* , mm
<i>Balanced layout for $(x_c^*, y_c^*) = (0, 0)$</i>			
$r_1 = 10$	11.7741100	11.7739800	21.9314070
$r_2 = 5$	-23.0635050	-0.4844600	
$r_3 = 5$	-0.4847160	-23.0635000	
<i>Layout without balance conditions</i>			
$r_1 = 10$	-6.6368890	16.4001780	22.3077950
$r_2 = 5$	20.9526650	-8.7113325	
$r_3 = 5$	-8.9957540	-20.8329690	
<i>Balanced layout for $(x_c^*, y_c^*) = (0, 0)$ and the prohibited area \hat{C} with $\hat{r} = 3$</i>			
$r_1 = 10$	-18.6071430	-18.8686040	13.5000000
$r_2 = 5$	8.3990950	28.9564910	
$r_3 = 5$	28.8151920	8.7807170	

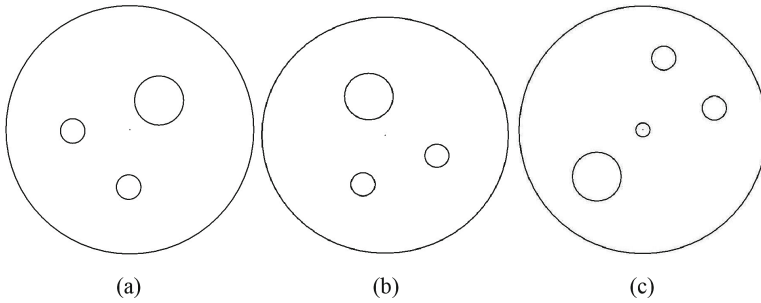


Fig. 1. Local optimal layout of circles: a) case 1; b) case 2; c) case 3.

These three cases are compared with respect to the average value of the heat flux at the circular objects' boundary.

The averaged heat fluxes (8) as a function of time for the case 1 (Fig. 1a) and case 2 (Fig. 1b) are presented in Fig. 2 and Fig. 3, correspondingly. Dynamics of the heat fluxes for the circles C_2 and C_3 for the case 1 are practically the same. However, the average relative difference in the heat fluxes between the circles C_1 and $C_2(C_3)$ is around 1.54 times.

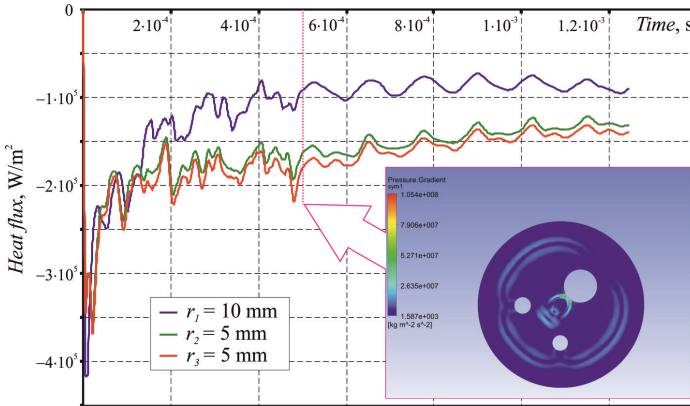


Fig. 2. Averaged heat fluxes for case 1 as a function of time and the pressure gradient distribution at $t = 5 \times 10^{-4}$ sec.

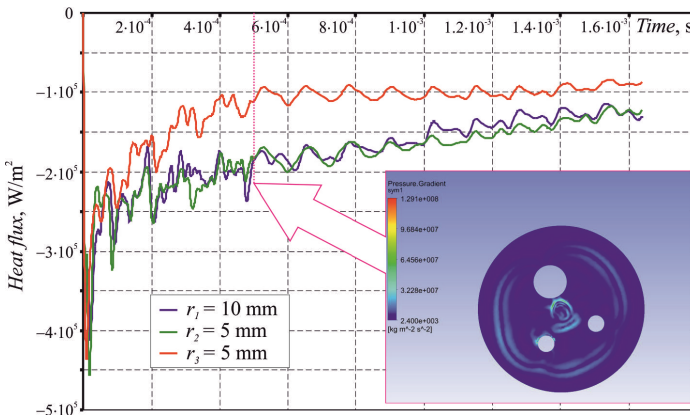


Fig. 3. Averaged heat fluxes for case 2 as a function of time and the pressure gradient distribution at $t = 5 \times 10^{-4}$ sec.

For the case 2 dynamics of heat fluxes for the circles C_1 and C_2 are very similar. The average relative difference in the heat fluxes between C_3 and C_1 (C_2) is around 1.45 times. Comparing to the case 1 (difference in 1.54), the heat distribution for the layout without balancing conditions is somewhat more uniform.

We may assume that the difference in the heat fluxes is caused by significantly different distances between the circles and the boundary of the high pressure zone (prohibited area \hat{C}) which is the origin of the shock-wave propagation. For the case 1 (Fig. 1a) the relative difference in these distances is up to 4.127 times, while for the case 2 (Fig. 1b) the difference is up to 3.131 times.

To verify this assumption, the values of the heat fluxes are calculated for the sparsest balance layout with the prohibited zone of high pressure (prohibited area \hat{C}). For this case the relative difference in the distances between the circles and the boundary of the prohibited zone is up 1.67 times. Numerical results are shown in Fig. 4.

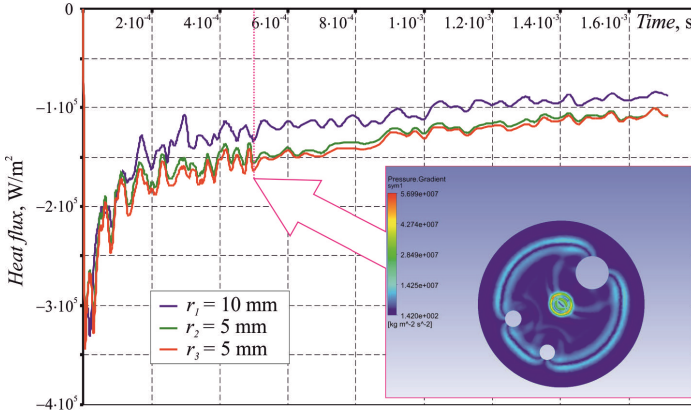


Fig. 4. Averaged heat fluxes for case 3 as a function of time and the pressure gradient distribution at $t = 5 \times 10^{-4}$ sec.

As can be seen from Fig. 4 the average heat fluxes for the circles are much closer than before. The relative average difference in heat fluxes between C_1 and $C_2(C_3)$ is less than 1.21 times. To verify that the layout with the prohibited zone \hat{C} provides a local minimum for the difference in the heat fluxes for the circles, three additional cases were considered. They were obtained from the layout given in Fig. 1c by moving the center of the prohibited zone from the origin to the points with the coordinates (2, 2), (-2, -2) and (2, -2) [mm], correspondingly. For these cases the average relative difference in the heat fluxes between C_1 and $C_2(C_3)$ is 1.231, 1.221 and 1.223 times, correspondingly.

Thus, we may conclude that to get the most uniform distribution of the heat fluxes under the shock waves, the sparsest layout has to be combined with the location of the high-pressure (prohibited) zone at the initial moment.

6 Conclusion

Balanced circular layout with prohibited zone arises in thermal deburring, where burrs are eliminated from machine parts by detonating gas mixtures in a deburring chamber. The prohibited zone is introduced to avoid placing parts in the central area of the chamber with extremely high temperature/pressure. To achieve more uniform distribution of thermal and pressure effects, the parts have to be placed in the chamber as

distant as possible one from another, as well as from the boundary of the chamber and from the prohibited zone. Balancing conditions are introduced to assure mechanical stability of the layout. Corresponding balanced sparsest layout is obtained by local optimization techniques. The numerical simulation for the temperature/pressure distribution in the chamber is performed using ANSYS CFX software. Three cases are considered: the sparsest balanced and unbalanced layouts, as well as the sparsest balanced layout with prohibited zone. Numerical analyses demonstrate that the later case is the most promising for the uniform thermal processing.

In this paper only circular objects were considered. However, in many applications more sophisticated shapes of the objects and/or prohibited zone are used. Generalizing the original model and solution/simulation techniques for more general, e.g., polygonal shapes, is an interesting area for the future research.











References

1. Benedict, G.F.: Thermal energy method: deburring (TEM). In: Benedict, G.F. (ed.) *Nontraditional Manufacturing Processes*, pp. 349–361. CRC Press, Boca Raton (2017). <https://doi.org/10.1201/9780203745410-22>
2. Fritz, A., Sekol, L., Koroskenyi, J., et al.: Experimental analysis of thermal energy deburring process by design of experiment. In: *Proceedings of the ASME 2012 International Mechanical Engineering Congress and Exposition*, vol. 3, pp. 2035–2041. ASME, New York (2012). <https://doi.org/10.1115/IMECE2012-88411>
3. Jin, S.Y., Pramanik, A., Basak, A.K., et al.: Burr formation and its treatments – a review. *Int. J. Adv. Manuf. Technol.* **107**(5–6), 2189–2210 (2020). <https://doi.org/10.1007/s00170-020-05203-2>
4. Lamikiz, A., Ukar, E., Tabernero, I., Martinez, S.: Thermal advanced machining processes. In: Davim, J.P. (eds.) *Modern Machining Technology*, pp. 335–372. Woodhead Publishing, Oxford (2011). <https://doi.org/10.1533/9780857094940.335>
5. Sibanda, P.S., Carr, P., Ryan, M., Bigot, S.: State of the art in surface finish of metal additive manufactured parts. In: Jin, Y., Price, M. (eds.) *Advances in Manufacturing Technology XXXIII. Advances in Transdisciplinary Engineering*, vol. 9, pp. 221–225. IOS Press, Amsterdam (2019). <https://doi.org/10.3233/ATDE190039>
6. Plankovskyy, S., Teodorczyk, A., Shypul, O., et al.: Determination of detonable gas mixture heat fluxes at thermal deburring. *Acta Polytechnica* **59**(2), 162–169 (2019). <https://doi.org/10.14311/AP.2019.59.0162>
7. Zhu, J.-H., Guo, W.-J., Zhang, W.-H., Liu, T.: Integrated layout and topology optimization design of multi-frame and multi-component fuselage structure systems. *Struct. Multidiscip. Optim.* **56**(1), 21–45 (2017). <https://doi.org/10.1007/s00158-016-1645-5>
8. Petrakov, Y., Shuplietsov, D.: Contour milling programming technology for virtual basing on a CNC machine. *Eastern-Eur. J. Enterp. Technol.* **98**(21), 54–60 (2019). <https://doi.org/10.15587/1729-4061.2019.162673>
9. Plankovskyy, S., Tsegelnyk, Y., Shypul, O., et al.: Cutting irregular objects from the rectangular metal sheet. In: Nechyporuk, M., et al. (eds.) *Integrated Computer Technologies in Mechanical Engineering. AISC*, vol. 1113, pp. 150–157. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37618-5_14

10. Araújo, L.J.P., Özcan, E., Atkin, J.A.D., Baumers, M.: Analysis of irregular three-dimensional packing problems in additive manufacturing: a new taxonomy and dataset. *Int. J. Prod. Res.* **57**(18), 5920–5934 (2019). <https://doi.org/10.1080/00207543.2018.1534016>
11. Romanova, T., Stoyan, Y., Pankratov, A., et al.: Optimal layout of ellipses and its application for additive manufacturing. *Int. J. Prod. Res.* (2019). <https://doi.org/10.1080/00207543.2019.1697836>
12. Machchhar, J., Elber, G.: Dense packing of congruent circles in free-form non-convex containers. *Comput. Aided Geom. Des.* **52–53**, 13–27 (2017). <https://doi.org/10.1016/j.cagd.2017.03.006>
13. Chazelle, B., Edelsbrunner, H., Guibas, L.J.: The complexity of cutting complexes. *Discrete Comput. Geom.* **4**(2), 139–181 (1989). <https://doi.org/10.1007/BF02187720>
14. Akeb, H., Hifi, M., Negre, S.: An augmented beam search-based algorithm for the circular open dimension problem. *Comput. Ind. Eng.* **61**(2), 373–381 (2011). <https://doi.org/10.1016/j.cie.2011.02.009>
15. Hifi, M., M'Hallah, R.: A literature review on circle and sphere packing problems: models and methodologies. *Adv. Oper. Res.* **2009** (2009). <https://doi.org/10.1155/2009/150624>. Article ID 150624
16. Stoyan, Y., Yaskov, G.: Packing equal circles into a circle with circular prohibited areas. *Int. J. Comput. Math.* **89**(10), 1355–1369 (2012). <https://doi.org/10.1080/00207160.2012.685468>
17. Litvinchev, I., Infante, L., Ozuna Espinosa, E.L.: Approximate circle packing in a rectangular container: integer programming formulations and valid inequalities. In: González-Ramírez, R.G., et al. (eds.) *Computational Logistics. ICCL 2014*. LNCS, vol. 8760, pp. 47–60. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-11421-7_4
18. Kovalenko, A.A., Romanova, T.E., Stetsyuk, P.I.: Balance layout problem for 3D-objects: mathematical model and solution methods. *Cybern. Syst. Anal.* **51**, 556–565 (2015). <https://doi.org/10.1007/s10559-015-9746-5>
19. Plankovskyy, S., Nikolaev, A., Shypul, O., et al.: Balance layout problem with the optimized distances between objects. In: Vasant, P., et al. (eds.) *Data Analysis and Optimization for Engineering and Computing Problems. EAISICC*, pp. 82–91. Springer, Cham (2020)
20. Romanova, T., Pankratov, A., Litvinchev, I., et al.: Sparsest packing of two-dimensional objects. *Int. J. Prod. Res.* (2020). <https://doi.org/10.1080/00207543.2020.1755471>
21. Romanova, T., Stoyan, Y., Pankratov, A., et al.: Decomposition algorithm for irregular placement problems. In: Vasant, P., et al. (eds.) *Intelligent Computing and Optimization. ICO 2019. AISC*, vol. 1072, pp. 214–221. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-33585-4_21
22. Menter, F.R.: Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA J.* **32**(8), 1598–1605 (1994). <https://doi.org/10.2514/3.12149>
23. Kader, B.A.: Temperature and concentration profiles in fully turbulent boundary layers. *Int. J. Heat Mass Transf.* **24**(9), 1541–1544 (1981). [https://doi.org/10.1016/0017-9310\(81\)90220-9](https://doi.org/10.1016/0017-9310(81)90220-9)



Mathematical Simulation of a Buck Quasi-resonant Converter as a Part of Autonomous Voltage Inverter

Alexey Gorodny¹  , Andrii Dymereets¹  ,
Roman Yershov¹  , Yurii Denisov¹  , and Serhii Boiko²  

¹ Chernihiv National University of Technology, Chernihiv, Ukraine
aleksey.gorodny@gmail.com, andrey.dymereets@gmail.com,
roman.d.yershov@gmail.com, den711td@gmail.com

² National University of Internal Affairs, Kremenchuk, Ukraine
bsn1987@i.ua

Abstract. A comparative analysis between the characteristics of a buck converter driven by pulse-width modulation and a buck half-wave zero-current switching quasi-resonant converter driven by pulse-frequency modulation for application in the power supply systems of the unmanned aerial vehicles is presented. All the characteristics are obtained both analytically by dint of a generalized method for commutation processes calculation in high-frequency switched-mode power supply converters and numerically by dint of simulation in MATLAB Simulink, and coincide each other that is proof of adequacy. The essential advantages of the quasi-resonant converter as a part of autonomous voltage inverter, which are consisting in a wider regulation range of output voltage, a lower output voltage ripple and in more near-to-linear output characteristic, are shown.

Keywords: Mathematical simulation · Converter · Buck · Quasi-resonant · Voltage inverter · Brushless motor · BLDC · Unmanned aerial vehicle · UAV

1 Introduction

Nowadays, unmanned aerial vehicles (UAV) are widely used both in civilian and in military application areas. Electric motors are consuming most of the electrical energy in UAV. Their operation modes, as well as power supply system, are directly affect the flight time of the UAV [1, 2]. In modern UAV, brushless DC (BLDC) motors are most commonly used due to their longevity and efficiency.

Figure 1 shows the structural diagram of the typical BLDC power supply system for the UAV [3, 4], where: B – battery; DC-DC – direct voltage converter/stabilizer (is often absent or replaced with bidirectional ones); DC-AC – autonomous voltage inverter.

The target of presented work is to improve the parameters and characteristics of the DC-AC inverter for supplying the BLDC in UAVs.

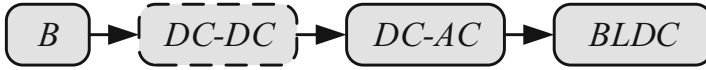


Fig. 1. Typical BLDC power supply system of the UAV

2 Overview

There are different ways to improve the performance of the power stage of power supply system: improvement of the components quality; optimization of control and energy circulation process [5, 6]; the usage of much advanced topologies for SMPS units (DC-DC and DC-AC), such as quasi-resonant converters (QRC) [7–14] instead of the conventional hard-switched SMPS driven by pulse-width modulation (PWM).

Three-phase voltage inverter is used as the DC-AC converter that drives the BLDC in the UAV [1–4, 12–14]. Each half-bridge stage (in particular, a high-side transistor switch) in such inverter in relation to the other elements of the power stage can be considered as a buck converter. It can be embodied as a conventional SMPS with PWM or zero-current switching (ZCS) QRC.

Comparison of energy efficiency between SMPS with PWM and ZCS-QRC in case of conventional buck and boost topologies is described in [15]. The researches show a reduction (3–7 times) of commutation losses in ZCS-QRC in comparison with SMPS with PWM.

The mathematical simulation of the commutation processes in QRCs are very complicated cause their nonlinearity and high-order of differential equation. In [16, 17], the linearized mathematical models of buck and boost QRCs are described, their transfer functions are calculated and their frequency response are obtained. The advantages of above method are in its simplicity, but it does not allow getting the dependences of voltages and currents of the converter in time domain. Therefore, it is not possible to estimate the internal transient processes that are occurring in it.

In [18, 19], a more sophisticated method is used, according to which the period of converter operation is divided onto intervals, at each of which the electromagnetic processes of the converter are calculated. However, in these works, transistor was considered as an ideal switch due to its parasitic parameters have not been taken into account, as well as the switching-on and switching-off intervals, which at the same time are very important when evaluating the energy efficiency of converters. A similar, but improved method, is described in detail in [20] and applied for the switching processes calculation in boost SMPS driven by pulse-width modulation (PWM).

In connection with the foregoing, the generalized method of commutation processes calculation in high-frequency switched-mode power converters [20] is selected for the analysis in this paper.

3 Mathematical Simulation of Electromagnetic Processes in Buck Half-Wave ZCS-QRC by the Generalized Method

Figure 2 shows a circuit diagram of the buck half-wave ZCS-QRC, where: U_{IN} – input voltage; VT – transistor switch; i_{VT} – transistor drain current; u_{VT} – transistor drain-source voltage; u_c – control voltage; VD1 – series diode that specifies a half-wave converter operation mode; L_r – resonant inductance; C_r – resonant capacitance; u_{Cr} – resonant capacitor voltage; VD2 – reverse diode; L_f – filter inductance; i_{Lf} – filter inductor current; u_{Lf} – filter inductor voltage; C_f – filter capacitance; R_{LOAD} – load resistance; i_{LOAD} – load current; u_{LOAD} – load voltage.

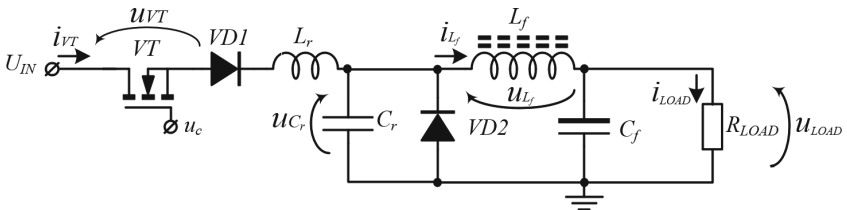


Fig. 2. Scheme of buck half-wave ZCS-QRC

Figure 3 shows the theoretical (qualitative) timing diagrams of converter operation [7] and their division onto specific operation intervals:

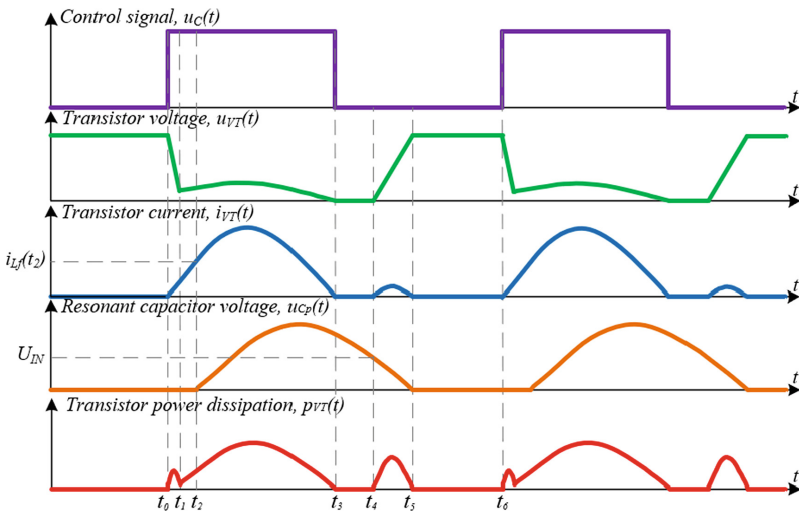


Fig. 3. Theoretical timing diagrams of buck half-wave ZCS-QRC operation

- t_0-t_1 – transistor switching-on;
- t_1-t_2 – transistor current increasing;
- t_2-t_3 – resonance;
- t_3-t_4 – resonant capacitor discharging;
- t_4-t_5 – transistor switching-off;
- t_5-t_6 – transistor close-state.

According to the generalized method, an equivalent circuit is composed for each operation interval, a transform to the Laplace-domain equivalent circuit is made and, on its basis, a matrix Eq. (1) is composed:

$$[Y] \cdot [U] = [J], \tag{1}$$

where $[Y]$ – matrix of conductivities; $[U]$ – vector-column of unknown voltages; $[J]$ – vector-column of assigning currents.

The solution of Eq. (1) is a vector-column of unknown node voltages:

$$[U] = [Y]^{-1} \cdot [J], \tag{2}$$

where $[Y]^{-1}$ – matrix inverse to the $[Y]$.

3.1 Transistor Switching-on Interval (T_0-T_1)

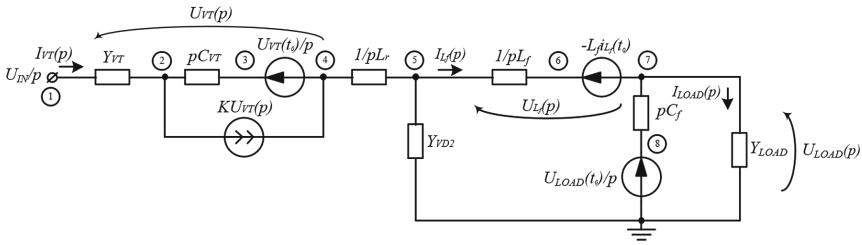


Fig. 4. Laplace-domain equivalent circuit for transistor switching-on interval

For transistor switching-on interval, Eq. (1) has the form (Fig. 4):

$$\begin{bmatrix} pC_{VT} + Y_{VT} & -pC_{VT} - K & 0 & 0 \\ -pC_{VT} & pC_{VT} + \frac{1}{pL_r} + K & -\frac{1}{pL_p} & 0 \\ 0 & -\frac{1}{pL_r} & \frac{1}{pL_r} + Y_{VD2} + \frac{1}{pL_f} & -\frac{1}{pL_f} \\ 0 & 0 & -\frac{1}{pL_f} & \frac{1}{pL_f} + pC_f + Y_{LOAD} \end{bmatrix} \times \begin{bmatrix} U_2(p) \\ U_4(p) \\ U_5(p) \\ U_7(p) \end{bmatrix} = \begin{bmatrix} \frac{U_{IN}(Y_{VT} - K) + U_{VT}(t_0)C_{VT}}{p} \\ \frac{U_{IN}K - U_{VT}(t_0)C_{VT}}{p} \\ -\frac{i_{L\phi}(t_0)}{p} \\ U_{LOAD}(t_0)C_\phi + \frac{i_{L\phi}(t_0)}{p} \end{bmatrix} \tag{3}$$

Solving Eq. (3), the unknown node voltages can be obtained and the voltages and currents at the converter elements can be found:

$$U_{VT}(p) = \frac{U_{IN}}{p} - U_4(p), \tag{4}$$

$$U_{Cr}(p) = U_5(p), \tag{5}$$

$$U_{Lf}(p) = U_5(p) - U_7(p), \tag{6}$$

$$U_{LOAD}(p) = U_7(p), \tag{7}$$

$$I_{VT}(p) = \left(\frac{U_{IN}}{p} - U_2(p) \right) \cdot Y_{VT}. \tag{8}$$

$$I_{Lf}(p) = \frac{U_5(p) - U_7(p) + L_f \cdot i_{Lf}(t_0)}{p \cdot L_f}, \tag{9}$$

$$I_{LOAD}(p) = U_7(p) \cdot Y_{LOAD} \tag{10}$$

Substituting in (4)–(10) the expressions of the node voltages that obtained from (3) and by applying the inverse Laplace transform, the expressions for the converter voltages and currents in the time domain are obtained. By these expressions it is possible to draw the quantitative timing diagrams of the converter operation, to calculate electrical and energy parameters, to obtain the dependences of electrical and energy characteristics on the parameters of converter elements.

The calculation for all other intervals is carried out in the same way.

3.2 Transistor Current Increasing Interval (T₁–T₂)

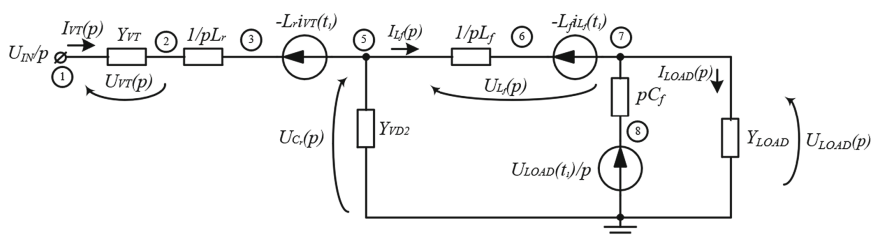


Fig. 5. Laplace-domain equivalent circuit for transistor current increasing interval

Matrix equation for transistor current increasing interval (Fig. 5):

$$\begin{bmatrix} Y + \frac{1}{pL_r} & -\frac{1}{pL_r} & 0 \\ -\frac{1}{pL_p} & \frac{1}{pL_r} + \frac{1}{pL_f} + Y_{VD} & -\frac{1}{pL_f} \\ 0 & -\frac{1}{pL_f} & \frac{1}{pL_f} + pC_f + Y_{LOAD} \end{bmatrix} \times \begin{bmatrix} U_2(p) \\ U_5(p) \\ U_7(p) \end{bmatrix} = \begin{bmatrix} \frac{U_{IN}Y_{VT} - i_{VT}(t_1)}{p} \\ \frac{i_{VT}(t_1)}{p} - \frac{i_{L\phi}(t_1)}{p} \\ U_{LOAD}(t_1)C_f + \frac{i_{L_f}(t_1)}{p} \end{bmatrix} \quad (11)$$

3.3 Resonance Interval (T₂-T₃)

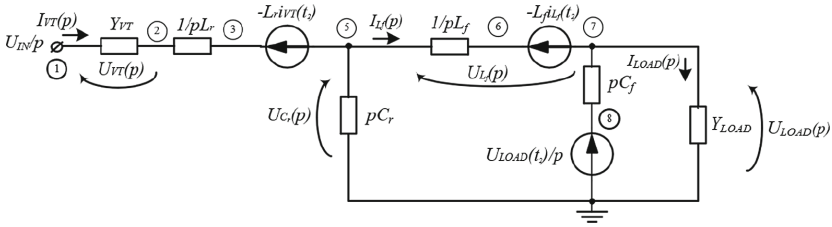


Fig. 6. Laplace-domain equivalent circuit for resonance interval

Matrix equation for resonance interval (Fig. 6):

$$\begin{bmatrix} Y_{VT} + \frac{1}{pL_r} & -\frac{1}{pL_r} & 0 \\ -\frac{1}{pL_r} & pC_r + \frac{1}{pL_r} + \frac{1}{pL_f} & -\frac{1}{pL_f} \\ 0 & -\frac{1}{pL_f} & \frac{1}{pL_f} + pC_f + Y_{LOAD} \end{bmatrix} \times \begin{bmatrix} U_2(p) \\ U_5(p) \\ U_7(p) \end{bmatrix} = \begin{bmatrix} \frac{U_{IN}Y_{VT} - i_{VT}(t_2)}{p} \\ \frac{i_{VT}(t_2)}{p} - \frac{i_{L_f}(t_2)}{p} \\ U_{LOAD}(t_2)C_f + \frac{i_{L_f}(t_2)}{p} \end{bmatrix} \quad (12)$$

3.4 Resonant Capacitor Discharging Interval (T_3 – T_4)

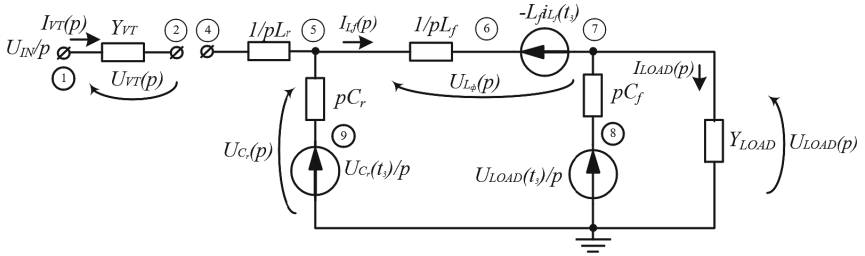


Fig. 7. Laplace-domain equivalent circuit for resonant capacitor discharging interval

Matrix equation for resonant capacitor discharging interval (Fig. 7):

$$\begin{bmatrix} Y_{VT} & 0 & 0 & 0 \\ 0 & \frac{1}{pL_r} & -\frac{1}{pL_r} & 0 \\ 0 & -\frac{1}{pL_r} & \frac{1}{pL_r} + pC_r + \frac{1}{pL_f} & -\frac{1}{pL_f} \\ 0 & 0 & -\frac{1}{pL_f} & \frac{1}{pL_f} + pC_f + Y_{LOAD} \end{bmatrix} \cdot \begin{bmatrix} U_2(p) \\ U_4(p) \\ U_5(p) \\ U_7(p) \end{bmatrix} = \begin{bmatrix} \frac{U_{IN}Y_{VT}}{p} \\ 0 \\ U_{C_r}(t_3)C_r - \frac{i_{L_f}(t_3)}{p} \\ U_{LOAD}(t_3)C_f + \frac{i_{L_f}(t_3)}{p} \end{bmatrix}. \quad (13)$$

3.5 Transistor Switching-off Interval (T_4 – T_5)

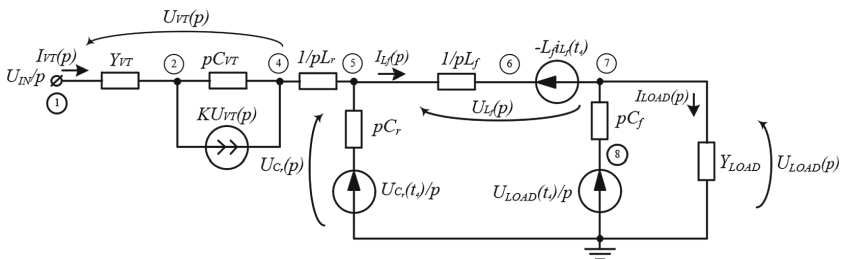


Fig. 8. Laplace-domain equivalent circuit for transistor switching-off interval

Matrix equation for transistor switching-off interval (Fig. 8):

$$\begin{bmatrix} pC_{VT} + Y_{VT} & -pC_{VT} - K & 0 & 0 \\ -pC_{VT} & pC_{VT} + \frac{1}{pL_r} + K & -\frac{1}{pL_r} & 0 \\ 0 & -\frac{1}{pL_r} & pC_r + \frac{1}{pL_r} + \frac{1}{pL_f} & -\frac{1}{pL_f} \\ 0 & 0 & -\frac{1}{pL_f} & \frac{1}{pL_f} + pC_f + Y_{LOAD} \end{bmatrix} \times \begin{bmatrix} U_2(p) \\ U_4(p) \\ U_5(p) \\ U_7(p) \end{bmatrix} = \begin{bmatrix} \frac{U_{IN}(Y_{VT} - K)}{p} \\ \frac{U_{IN}K}{p} \\ U_{Cr}(t_4)C_r - \frac{i_{Lr}(t_4)}{p} \\ U_{LOAD}(t_4)C_r + \frac{i_{Lr}(t_4)}{p} \end{bmatrix} \quad (14)$$

3.6 Transistor Close-State Interval (T₅–T₆)

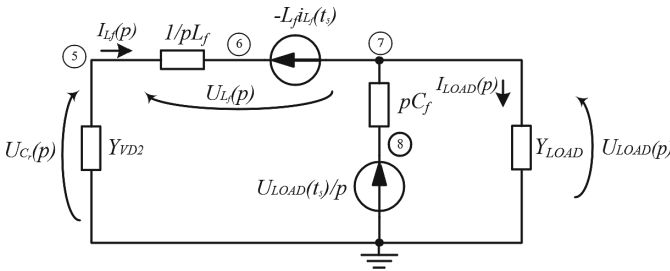


Fig. 9. Laplace-domain equivalent circuit for transistor close-state interval

Matrix equation for transistor close-state interval (Fig. 9):

$$\begin{bmatrix} Y_{VD2} + \frac{1}{pL_f} & -\frac{1}{pL_f} \\ -\frac{1}{pL_f} & Y_{LOAD} + \frac{1}{pL_f} + pC_f \end{bmatrix} \times \begin{bmatrix} U_5(p) \\ U_7(p) \end{bmatrix} = \begin{bmatrix} -\frac{i_{Lr}(t_s)}{p} \\ U_{LOAD}(t_s)C_f + \frac{i_{Lr}(t_s)}{p} \end{bmatrix}. \quad (15)$$

4 Comparison of SMPS with PWM and ZCS-QRC Characteristics

A similar calculation was performed for conventional buck converter with the aim of comparative analysis and identifying the main advantages and disadvantages versus ZCS-QRC. The parameters of the converter elements used during the investigation are collected in Table 1.

Figure 10 shows the regulation characteristics of the buck SMPS and ZCS-QRC, particularly, dependences of the output voltage (Fig. 10a) and the voltage ripple factor (Fig. 10b) on the control signal period under a constant pulse duration.

Table 1. Symbols and nominal values of SMPS and ZCS-QRC parameters

Parameter	Symbol	Value
Input voltage	U_{IN}	30 V
Load impedance	R_{LOAD}	10 O
Operation period	T_{OP}	1.0 us
Control pulses duration	τ_c	200 ns
Filter inductance	L_f	50 uH
Filter capacitance	C_f	220 nF
Resonant inductance	L_r	1.0 uH
Resonant capacitance	C_r	2 nF
Inductors resistance	R_L	20 m Ω
Capacitors conductivity	G_C	100 μ S
Transistor resistance	R_{VT}	1.6 O
Transistor drain-source capacitance	C_{VT}	150.. 2300 pF

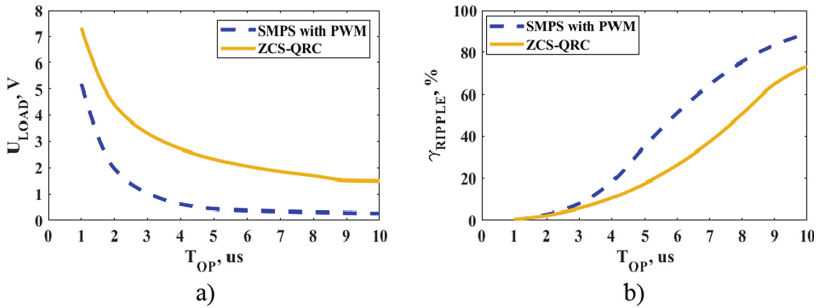
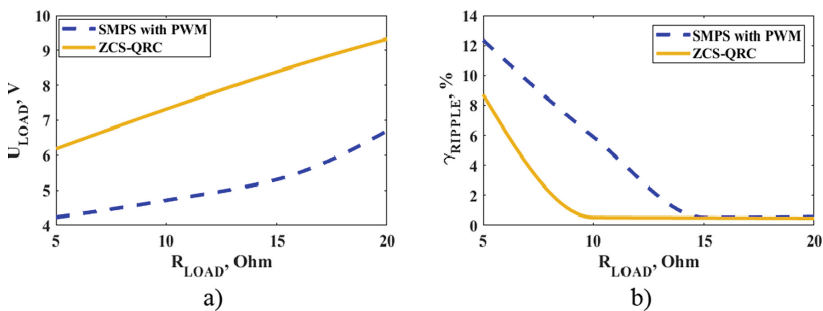

Fig. 10. Buck SMPS and ZCS-QRC regulation characteristics: output voltage (a) and ripple factor (b) graphs under operation period variation

Figure 11 shows the output characteristics of the buck SMPS and ZCS-QRC, particularly, dependences of the output voltage (Fig. 11a) and the voltage ripple factor (Fig. 11b) under variation a load impedance.


Fig. 11. Buck SMPS and ZCS-QRC output characteristics: output voltage (a) and ripple factor (b) charts under variation of a load impedance

5 Conclusions

- The mathematical simulation of electromagnetic processes in buck half-wave ZCS-QRC was made by the generalized method of commutation processes calculation;
- The characteristics of ZCS-QRC driven by PFM and SMPS driven by PWM with the same parameters of their elements was compared;
- It is shown that in ZCS-QRC the output voltage value is greater than in SMPS with PWM (1.5–3 times), and the output voltage ripple factor is lower (up to 30%) under the same control pulse duty cycle;
- It is shown that ZCS-QRC has a much linear output characteristic than SMPS with PWM, a higher output voltage value (up to 50%), and a lower output voltage ripple (up to 6%) under the same load impedance.

Acknowledgement. The Ministry of Education and Science of Ukraine as a part of the scientific research project “Systems of electric drives with improved energy and dynamic characteristics for special applications” (registration number **0119U000421**) has supported this work.

References

1. Cwojdzinski, L., Adamski, M.: Power units and power supply systems in UAV. *Aviation* **18**, 1–8 (2014)
2. Lee, B., Kwon, S., Park, P., Kim, K.: Active power management system for an unmanned aerial vehicle powered by solar cells, a fuel cell, and batteries. *IEEE Trans. Aerosp. Electron. Syst.* **4**(50), 3167–3177 (2014)
3. Lee, B.K., Fahimi, B., Ehsani, M.: Overview of reduced parts converter topologies for AC motor drives. In: *IEEE 32nd Annual Power Electronics Specialists Conference*, vol. 4, pp. 2019–2024, Vancouver, BC (2001)
4. Boukoberine, M.N., Zhou, Z., Benbouzid, M.: Power supply architectures for drones - a review. In: *45th Annual Conference of the IEEE Industrial Electronics Society (IECON 2019)*, pp. 5826–5831. Lisbon, Portugal (2019)
5. Sai, P., Rani, Ch., Nelakuditi, U.: Implementation of power optimization technique for UAVs. *Mater. Today Proc.* **5**, 132–137 (2018)
6. Shilov, Y.: Development of automatic control systems for unmanned aerial vehicle multi-tare type. *Trudy MFTI* **4**(6), 139–151 (2014)
7. Liu, K.: High frequency quasi-resonant converter techniques. Ph. D. Dissertation. Electrical Engineering Department, Virginia Polytechnic Institute and State University (1986)
8. Zheng, T., Chen, D., Lee, F.C.: Variations of quasiresonant dc-dc converter topologies. In: *IEEE Power Electronics Specialists’ Conference Record*, pp. 381–392 (1986)
9. Liu, K., Oruganti, R., Lee, F.C.: Resonant switches — topologies and characteristics. *IEEE Trans. Power Electron.* **1**(2), 106–116 (1987)
10. Tomioka, S., Abe, S., Shoyama, M., Ninomiya, T., Firmansyah, E.: Zero-current-switched quasi-resonant boost converter in power factor correction application. In: *24th Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, pp. 1165–1169 (2009)
11. Kim, K.-H., Youn, M.-J.: Performance comparison of PWM inverter and variable DC link inverter schemes for high-speed sensorless control of BLDC motor. *Electron. Lett.* **21**(38), 1294–1295 (2002)

12. Zhi, Y.P., Fang, L.L.: Novel soft-switching inverter for brushless DC motor variable speed drive system. *IEEE Trans. Power Electron.* **2**(19), 280–288 (2004)
13. Pan, Z.Y., Luo F.L.: Transformer based resonant DC link inverter for brushless DC motor drive system. In: *IEEE 35th Annual Power Electronics Specialists Conference*, vol. 5, pp. 3866–3872. Aachen, Germany (2005)
14. Khalilian, M., Farzanehfard, H., Adib, E.: A novel quasi-resonant three-phase soft-switching inverter. In: *3rd Power Electronics and Drive Systems Technology (PEDSTC 2012)*, pp. 471–476, Tehran (2012)
15. Denisov, Y., Gorodny, A., Gordienko, V., Stepenko, S., Yershov, R., Tepla, T.: Comparison of power losses in switches of increasing number of common-state directives with parallel and series resonant circuits. *Tekhnichna elektrodynamika* **4**, 44–46 (2016)
16. Szabo, A., Kansara, M., Holmes, P.G., Ward, E.S.: Mathematical modelling of quasi-resonant converters. *Electron. Lett.* **31**, 2057–2058 (1995)
17. Costa, J.M.F.D., Silva, M.M.: Small-signal models of quasi-resonant converters. In: *IEEE International Symposium on Industrial Electronics (ISIE 1997)*, vol. 2, pp. 258–262 (1997)
18. Wong, L.K., Leung, F.H., Tam, P.K.S.: A simple large-signal nonlinear modeling approach for fast simulation of zero-current-switch quasi-resonant converters. *IEEE Trans. Power Electron.* **3**(12), 437–442 (1997)
19. Nejadpak, A., Tahami, F.: Stabilizing controller design for quasi-resonant converters described by a class of piecewise linear models. *IEEE Trans. Circ. Syst. I: Regul. Pap.* **61**, 312–323 (2013)
20. Gorodny, A., Dymereys, A., Kuts, Y., Denisov, Y., Denisova, N.: Generalized method of commutation processes calculation in high-frequency switched-mode power converters. In: *14th ISP Conference Mathematical Modeling and Simulation Systems (MODS2019)*, pp. 71–80 (2019)



Tensor Modeling of Spatial Rigidity of Metal-Cutting Machine Spindles

Volodymyr Chupryna^(✉)  and Andriy Rudyk 

State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Striletska Street 1, Chernihiv 14033, Ukraine
voldchu@ukr.net, andrei.rudik@gmail.com

Abstract. The paper proposes a tensor approach to determining the spatial rigidity of spindle unit of a metal-cutting machine in the form of 3D surfaces. A tensor-geometric model of rigidity of a spindle unit in the form of an ellipsoid of rigidity (or pliability) is developed, which allows to obtain the value of rigidity of a spindle unit in cutting zone in any direction of the working space. With the known magnitudes of the force load on this model it is possible to determine the deformation of the spindle in the directions that affect the accuracy of the workpiece. A mathematical simulation of the elastic system of a spindle unit for a boring processing variant is carried out, which establishes a direct relation between the shape of the ellipsoid of rigidity (pliability) of the spindle unit and the elliptical (not round) shape of the resulting trajectory of the cutting tool. This allows you to predict the achievable accuracy of machining on a machine with this spindle assembly when designing.

Keywords: Spindle unit · Tensor model · Ellipsoid of rigidity · Ellipsoid of pliability · Accuracy of processing

1 Introduction

The intensification of progressive manufacturing processes in mechanical engineering requires high static and dynamic rigidity of metal-cutting equipment, including spindle units (SU).

Recently, researchers have used tensor calculus to describe the characteristics of processes and systems, including tensor analysis in mechanics [8, 12]. So, in [9], on the basis of the tensor idea of the structure of the processing machine, the notion of the ellipsoid of rigidity is formulated, which is used to characterize the rigidity of the entire machine, for example machine tools or industrial robots [15]. The ellipsoid reflects the rigidity anisotropy of the machine in three-dimensional space. The rigidity of the elastic system of the machine-tool is usually evaluated in the area of processing parts (cutting area).

However, the overall rigidity of the machine is determined by the rigidity of the structural elements - parts and components [5]. Due to the extension of the aggregate-modular principle of machine layout, decomposition methods, including diacoptics, became more important [6]. This requires the identification and optimization of the characteristics of individual nodes, in particular their rigidity.

SU is an important element of a metal-cutting machine that largely determines the static and dynamic rigidity of the system in the cutting area [10]. The rigidity of the SU depends on many parameters (outer and inner diameters of the spindle, the distance between the supports, the lengths of the console and the shank, the number of supports of the spindle, their type and other factors). The high rigidity of the structure provides the necessary precision of machining of parts according to standards.

To obtain the specified geometric accuracy, it is also necessary to select efficient machining modes and optimal tool parameters.

Many works have been devoted to the study of elastic properties and accuracy of SU on supports of different types, in particular of the spindles on rolling supports [3], on hydraulic supports [14] and others (including electro-spindles and motor-spindles [11]). Elastic characteristics of the SU are usually evaluated in the cutting zone by the radial rigidity of the elastic system C_{SU} (N/m), or its inverse value - the pliability δ_{SU} (m/N). The radial stiffness of the joint is normalized on the spindle flange, and the radial stiffness of the node is less than the axial one.

The required structural rigidity is provided at the design stage in CAD systems. Thus, virtual modeling of SU with definition of characteristics is carried out in CAI systems [2], often with the subsequent optimization of their parameters [1, 13].

The elastic properties of machines today describe the spatial values of rigidity with the use of second-order tensors (in the form of ellipsoids of rigidity). Tensors allow you to determine the value of rigidity or durability of the structure at a given point, and in any direction of three-dimensional space. With known power loads, the ellipsoids of rigidity make it possible to calculate the magnitudes of machine deformation in the direction of forming processing errors. This makes it possible to apply tensor mathematical rigidity models to evaluate the accuracy of machining on the machines when designing them.

However, each node of the machine has its own rigidity characteristics, in particular the SU. A comprehensive assessment of the spatial rigidity of the SU shows that a reliable model for the rigidity of the SU has not been definitively identified today.

The purpose of the article is to theoretically substantiate and construct a tensor-geometric model of the spatial rigidity of the SU to evaluate the elastic-deformation properties of the structure and the accuracy of machining of parts.

2 Tensor-Geometric Model of Spindle Unit Elastic System

When machining parts, it is important to ensure a high rigidity of the elastic system of the machine in the cutting area, which is significantly dependent on the rigidity of the SU and the additional technological equipment (mandrel, tool, etc.).

Preliminary studies of the spatial rigidity of the SU indicate a rather complex shape of its surface. However, in the first approximation, it can be approximated by the surface of the ellipsoid.

To analyze the elastic-deformation properties of SU, we develop a mathematical model of SU with bringing the rigidity to the cutting zone.

The tensor-geometric model will be made on the example of the SU (Fig. 1). An instrumental mandrel (with a boring cutter) is also attached to the SU.

Theoretically, the rigidity of the SU in the cutting area can be calculated using the scheme shown in Fig. 2. After defining the bends x, y, z of the spindle from the components of the cutting force P , it is possible to calculate the principal rigidity along the major axes c_{x0}, c_{y0}, c_{z0} and construct the ellipsoid of rigidity.

However, let us consider the more general case of the location of the rigidity ellipsoid with respect to the coordinate axes and solve the problem in a generalized state using tensors.

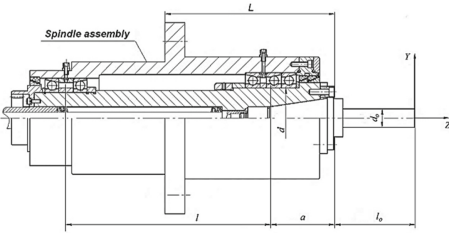


Fig. 1. Spindle assembly (with mandrel)

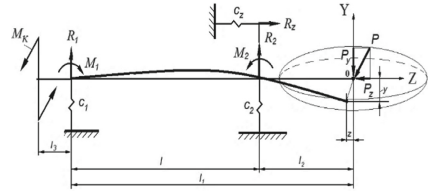


Fig. 2. Calculation scheme of a spindle node

To analyze the elastic-deformation properties of SU when cutting as an invariant we will take the work of the generalized forces on small displacements [16]

$$Inv = A_f = \sum (f \cdot dq), \tag{1}$$

where A_f - is the work of the generalized forces; f - is the vector of generalized forces; dq - is the vector of generalized elastic displacements.

When moving to another basis, the invariant, as a scalar value, does not change

$$Inv = f^T \cdot dq = P^T \cdot dX = const, \tag{2}$$

where $dX = [dx, dy, dz, d\phi_x, d\phi_y, d\phi_z]$ - vector of generalized coordinates in the cutting area;

$P = [P_x, P_y, P_z, M_x, M_y, M_z]$ - vector of the generalized forces applied in the cutting zone.

Using the transposed Jacobi matrix J^T , we have

$$f = J^T \cdot P. \tag{3}$$

The spatial rigidity matrix of the SU structure can be represented as a diagonal matrix with principal rigidity: $C_{SU} = diag [c_{x0}, c_{y0}, c_{z0}, c_{\phi x0}, c_{\phi y0}, c_{\phi z0}]$, where c_{x0}, c_{y0}, c_{z0} - translational rigidity of the node; $c_{\phi x0}, c_{\phi y0}, c_{\phi z0}$ - rotary rigidity of the node. Then,

$$P = C_{SU} \cdot X. \tag{4}$$

The equation for the ellipsoid of rigidity is written as

$$L^T \cdot C_{SU} \cdot L = 1, \tag{5}$$

where L - vector that sets the direction in space.

Thus, the rigidity ellipsoid reflects the surface of the level of a given quadratic shape and is similar to the geometric image of a second order tensor (for example, an inertia tensor).

Along its principal axes of rigidity, it sets its own orthogonal basis– the coordinate system of the ellipsoid, which is oriented in some way in space, and the magnitudes of the axes determine their numbers – major rigidities of the ellipsoid.

When analyzing the accuracy of the machine in the cutting zone, mainly translational deviations of the elements of the elastic system and, accordingly, the translational rigidity of the SU are taken into account. That is,

$$C_{SU} = \begin{bmatrix} c_{x0} & 0 & 0 \\ 0 & c_{y0} & 0 \\ 0 & 0 & c_{z0} \end{bmatrix}. \tag{6}$$

In geometric interpretation, the rigidity matrix of the SU can be represented in the form of an ellipsoid with major axes c_{x0}, c_{y0}, c_{z0} , which in the three-dimensional coordinate system is a quadratic shape.

The direction of the principal axes of rigidity is given by its own orthogonal basis, the coordinate system of the ellipsoid, which is oriented in space in some way. The magnitudes of the half axes determine the eigenvalues the principal rigidities of the ellipsoid. However, for the vast majority of SU structures, one of the major rigidity axes coincides with the spindle axis.

An ellipsoid of rigidity can always be constructed at an arbitrary point O if the directions of the principal rigidity axes x_0, y_0, z_0 in the XYZ basis are known, and the values of the translational rigidity coefficients c_{x0}, c_{y0}, c_{z0} in these directions. In a typical form, the ellipsoid of rigidity is shown in Fig. 3.

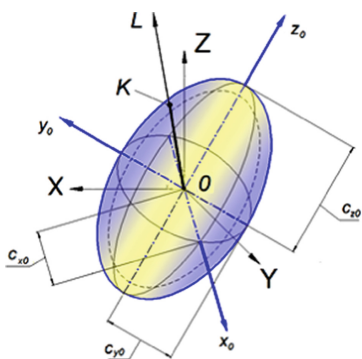


Fig. 3. A typical ellipsoid of rigidity

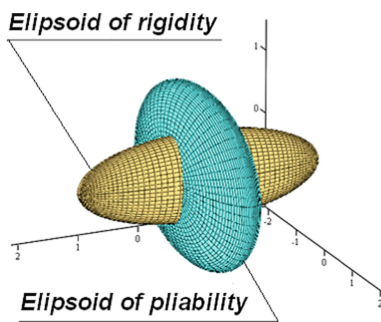


Fig. 4. Two conjugate ellipsoids

The rigidity ellipsoid reflects the change in rigidity of the elastic system in space, that is the distribution of rigidity values of the system in different directions of the SU workspace from the center of the ellipsoid. The amount of rigidity in a given direction is determined by the length of the radius-vector of the ellipsoid. The developed model allows to determine the value of the rigidity of the SU in an arbitrary direction of the working space. If a beam L is drawn from the center of the ellipsoid (point O), then it intersects the surface of the ellipsoid at point K . The OK radius-vector module will determine the rigidity of the SU structure in this direction.

In practice, it is often more convenient to handle the characteristics of compliance than rigidity. A rigidity ellipsoid is always associated with a return ellipsoid of inertia. It is also depicted as an ellipsoid, but with a different ratio of major axis values. For example, Fig. 4 shows MathCAD-built interconnected (conjugate) ellipsoids of rigidity (E_C) and pliability (E_δ). The ratio of the rigidity on the major axes is 0.5: 0.7: 2, and the pliability - 2: 1.4: 0.5, respectively.

The mathematical description of the ellipsoid as a second-order surface in Cartesian coordinates is as follows [4].

$$\begin{aligned} a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{23}yz \\ + 2a_{13}xz + 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} = 0, \end{aligned} \quad (7)$$

and at least one of the coefficients $a_{11}, a_{22}, a_{33}, a_{12}, a_{13}, a_{23}$ must be non-zero.

The invariants of this equation are: $I_1 = a_{11} + a_{22} + a_{33}$,

$$I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{23} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{33} & a_{13} \\ a_{13} & a_{11} \end{vmatrix}, I_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}. \quad (8)$$

Equation (7) can also be written in matrix form as

$$[X_4^T] \cdot [A] \cdot [X_4] = 0, \quad (9)$$

where $[X_4] = [x \ y \ z \ 1]^T$.

For center frames (which also include the ellipsoid surface), the center of the surface can be moved to the origin point.

$$[X_4^T] \cdot [Tr^T] \cdot [A_4] \cdot [Tr] \cdot [X_4] = 0, \quad (10)$$

where Tr - the corresponding transfer matrix.

Also, by turning the Euler angles, you can make the quadrilateral axes parallel to the coordinate axes. Then we get the standard (canonical) quadric equation in the form

$$[X_4^T] \cdot \begin{bmatrix} \bar{a}_{11} & 0 & 0 & 0 \\ 0 & \bar{a}_{22} & 0 & 0 \\ 0 & 0 & \bar{a}_{33} & 0 \\ 0 & 0 & 0 & \bar{a}_{44} \end{bmatrix} \cdot [X_4] = 0. \quad (11)$$

We separately distinguish the quadratic and linear groups of terms of Eq. (7)

$$[x \ y \ x] \cdot \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} + 2[a_{14} \ a_{24} \ a_{34}] \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} + a_{44} = 0. \quad (12)$$

Then, in a compact matrix entry we have

$$[X_3^T] \cdot [A_3] \cdot [X_3] + 2[B_3] \cdot [X_3] + a_{44} = 0, \quad (13)$$

where $[A_3] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$, $[B_3] = [a_{14} \ a_{24} \ a_{34}]$, $[X_3] = [x \ y \ z]^T$.

The linear group of members of Eq. (8) is responsible for moving the center of the surface of the second order from the origin of the coordinate system (point O), and the quadratic group - for the rotation of the principal coordinate system of the surface relative to the coordinate axes.

If the center of the second order surface coincides with the origin of the coordinate system (point O), then matrix Eq. (12) is simplified

$$[X_3^T] \cdot [A_3] \cdot [X_3] + a_{44} = 0. \quad (14)$$

The ellipsoid belongs to the central quadrics and has a central point of symmetry. In what follows, we consider the ellipsoid with the center at the origin.

If the principal axes of the ellipsoid are collinear to the axes of the coordinates X , Y , Z then the ellipsoid equation has the form (canonical form)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad (15)$$

where a, b, c —the numbers that determine the axis of the ellipsoid.

If as eigenvalues a, b, c of the ellipsoid we take the values of rigidity along the principal rigidity axes c_{x0}, c_{y0}, c_{z0} , then the equation of the ellipsoid of rigidity will be written as

$$\frac{x^2}{c_{x0}^2} + \frac{y^2}{c_{y0}^2} + \frac{z^2}{c_{z0}^2} = 1. \quad (16)$$

Or in matrix form

$$[x \ y \ z] \cdot \begin{bmatrix} \frac{1}{c_{x0}^2} & 0 & 0 \\ 0 & \frac{1}{c_{y0}^2} & 0 \\ 0 & 0 & \frac{1}{c_{z0}^2} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 1. \quad (17)$$

The rigidity ellipsoid corresponds to the diagonal rigidity matrix of the system. For SU, it has the form (6).

In the construction of machines, the major axis of rigidity SU c_{x0}, c_{y0}, c_{z0} often coincide with the coordinate axes X, Y, Z.

The spatial rigidity of the SU in the form of an ellipsoid with principal rigidity (half axes) c_{x0}, c_{y0}, c_{z0} is shown in Fig. 3. The directions of major rigidity are often collinear to the axes of the principal coordinate system X, Y, Z.

The center of rigidity is located on the axis of rotation, at the end of the mandrel (point O), which is conventionally applied cutting force $P = [P_x, P_y, P_z]^T$.

If the radial rigidity of the spindle in different radial directions are identical, the ellipsoid of rigidity SU is transformed into an ellipsoid of rotation. To construct such an ellipsoid, we must know only the values of c_{y0} (or c_{x0}) and c_{z0} .

According to the design scheme (Fig. 2) under the action of the cutting force component P_y displacement (deflection) of the console of the spindle on the value of y . The amount of deflection depends on the rigidity c_y system in the radial direction.

The bending in the cutting zone can be calculated by the traditional method, taking into account the rigidity of the spindle shaft for bending and the rigidity of the supports. The radial rigidity of the SU is defined as

$$c_{y0} = P_y/y. \quad (18)$$

In the axial direction of the SU acting axial force component of the cutting force P_z , which is perceived angular contact bearing. The spindle is moved to the value z . The axial rigidity of SU is defined analogically (as $c_{z0} = P_z/z$). The axial rigidity of SU is almost determined by the rigidity of the thrust bearing and habitually exceeds the radial.

Figure 3 shows the spatial rigidity of the SU in the form of a rigidity ellipsoid in the isometric image, with all the major rigidities c_{x0}, c_{y0}, c_{z0} , and the cutting force vector P , which is applied at the center of the ellipsoid.

The vector X the spatial deformation of the spindle from the force P can be found from the expression (4) as

$$X = C_{SU}^{-1} \cdot P \quad (19)$$

However, when solving deformation problem it is sometimes more convenient to use the values of pliability than rigidity. For the three major rigidities c_{x0}, c_{y0}, c_{z0} , there are three major flexures $\delta_{x0}, \delta_{y0}, \delta_{z0}$ of SU, as inverse quantities, as well as the ellipsoid of elasticity (Fig. 4).

It reflects the change in the pliability of the SU in space. The amount of pliability is determined by the length of the radius-vector of the ellipsoid of pliability in a given direction. The ellipsoid of pliability is mathematically described by formulas similar to the ellipsoid of rigidity, in which $1/\delta$ is used instead of c .

It should also be noted this characteristic of conjugate ellipsoids. If known values for the major pliability, you can find the ellipsoid of rigidity. Conversely, if the known values of the major rigidity, it is possible to find the ellipsoid of pliability.

From formula (16) can be written the equation of the ellipsoid of rigidity to the base coordinate system X_0, Y_0, Z_0 through the major pliability of SU

$$\frac{x_0^2}{1/\delta_{x_0}^2} + \frac{y_0^2}{1/\delta_{y_0}^2} + \frac{z_0^2}{1/\delta_{z_0}^2} = 1, \tag{20}$$

or

$$\delta_{x_0}^2 x_0^2 + \delta_{y_0}^2 y_0^2 + \delta_{z_0}^2 z_0^2 = 1. \tag{21}$$

In matrix form

$$[x \ y \ z] \cdot \begin{bmatrix} \delta_{x_0}^2 & 0 & 0 \\ 0 & \delta_{y_0}^2 & 0 \\ 0 & 0 & \delta_{z_0}^2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 1. \tag{22}$$

The compliance ellipsoid corresponds to the diagonal matrix of the under-sensitivity of the system. For SU, it has the following form

$$D_{SU} = \begin{bmatrix} \delta_{x_0} & 0 & 0 \\ 0 & \delta_{y_0} & 0 \\ 0 & 0 & \delta_{z_0} \end{bmatrix}. \tag{23}$$

In Fig. 5 shows a SU with an ellipsoid of pliability

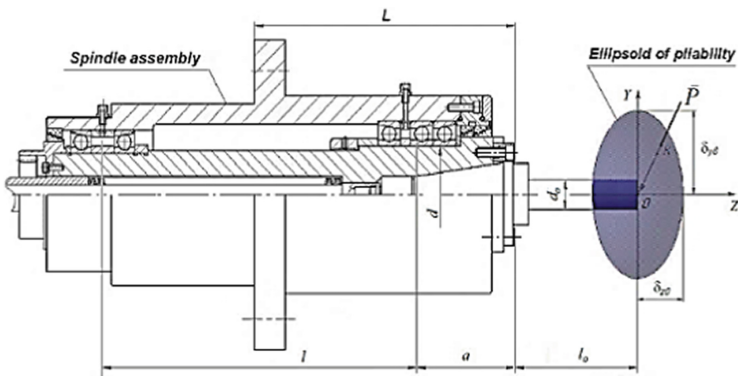


Fig. 5. SU with ellipsoid of pliability

The proposed ellipsoid of spatial rigidity (pliability) of the SU allows to determine the magnitudes of deformation of the spindle in any direction of the working space from the active force load.

The vector X of the spatial deformation spindle can be found as

$$X = D_{\text{SU}} \cdot P. \quad (24)$$

To find the magnitude of δ_k SU in a given direction, it is necessary to set a direct line L that passes through the center of the ellipsoid

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}, \quad (25)$$

where l, m, n – guiding coefficients; x_0, y_0, z_0 – coordinates of the ellipsoid's center.

If the point on the line coincides with the beginning of the coordinate system, then

$$x_0 = 0, y_0 = 0, z_0 = 0. \quad (26)$$

Line L can be defined as the intersection line of two non-coplanar planes

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases}. \quad (27)$$

We find the pliability of SU in this direction as the length (module) of the radius-vector \vec{K} drawn from the center of the ellipsoid to the point K

$$\delta_K = \left| \sqrt{(x_K^2 + y_K^2 + z_K^2)} \right|. \quad (28)$$

In the general case, the compliance ellipsoid is not an ellipsoid of rotation, that is, $\delta_{x0} \neq \delta_{y0}$ due to errors in processing and assembly (e.g., ellipse or ovality of openings for bearings in the housing, in the spindle cone, irrational mounting). This is usually the cause of errors in the workpiece.

3 Mathematical Modeling of Spindle Unit Elastic System

In order to study the deformation characteristics of the SU elastic system with the ellipsoid of rigidity (pliability), mathematical modeling of the system (19) in MatLab (Simulink module) was performed [7].

In order to perform system modeling, it is first necessary to develop different variants of models of application of forces: static and rotational.

The static force vector P can be specified in an arbitrary direction in the workspace (Fig. 6). In modeling, the direction of the force in space is determined in the spherical coordinate system by two angles ρ and φ .

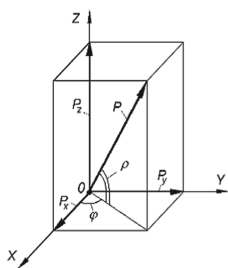


Fig. 6. The force vector P

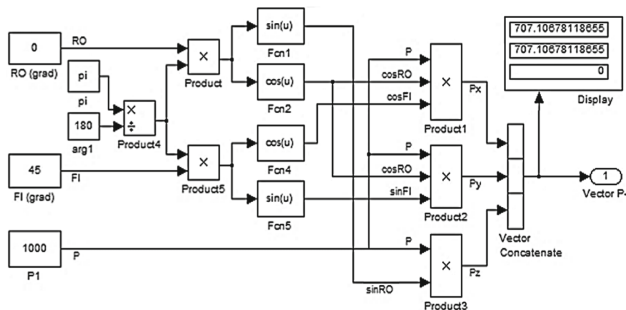


Fig. 7. Static force vector P simulation

The force vector P is the projection of the force on the coordinate axis

$$P = \begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \begin{bmatrix} P \cdot k_{Px} \\ P \cdot k_{Py} \\ P \cdot k_{Pz} \end{bmatrix} = \begin{bmatrix} P_x = P \cdot \cos \rho \cdot \cos \phi \\ P_y = P \cdot \cos \rho \cdot \sin \phi \\ P_z = P \cdot \sin \rho \end{bmatrix} \quad (29)$$

and when modeling the static force vector, the force value is given in the input constant P (here $P = 1000\text{N}$). The direction of the force vector in space is given by the angles ρ and ϕ (in degrees) - the input constants RO and FI (Fig. 7).

Modeling of the rotational force vector is shown in Fig. 8. The rotational force of cutting appears at boring operations on machines, and also at milling (processing by end mills in a circle). The force P rotates with respect to the spindle axis (Z axis)

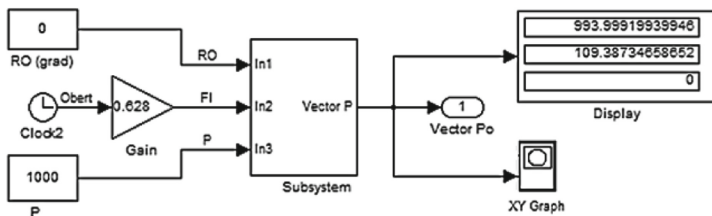


Fig. 8. Modeling of rotational force vector

The rotation of force relative to the axis is ensured in the model bath by the introduction of a block $Clock2$ in the scheme, which changes the angle ϕ from 0° to 360° .

After constructing the necessary models for the power loads, we proceed to the simulation of the investigated SU elastic system according to expression (19). To determine the displacement vector X under the action of rotational force $P = 1000\text{ N}$, a mathematical simulation of a static system was performed.

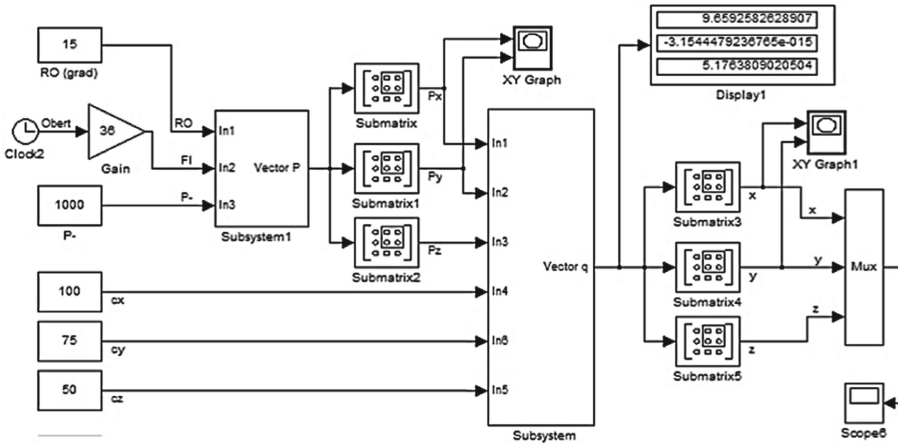


Fig. 9. Mathematical model of the elastic SU system with rotational force

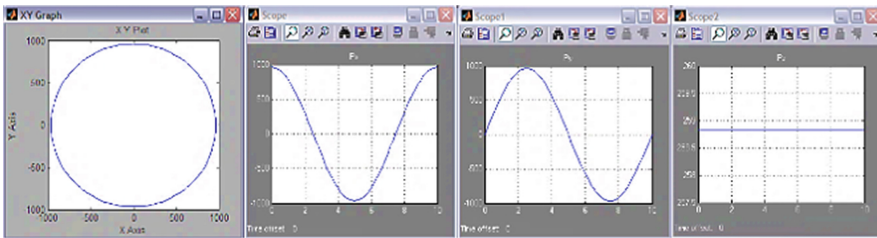


Fig. 10. The trajectory of the force vector with components

When the SW is loaded with a rotational force P (for example, during a boring operation on a machine), the displacement vector of the system X will also rotate synchronously and continuously change in magnitude (provided that $c_{x0} \neq c_{y0}$).

According to the ellipsoid of rigidity (or pliability) of the elastic system of the SU, the displacement vector X must describe the elliptical trajectory. Mathematical model of the system with rotational (around the spindle axis OZ) force, shown in Fig. 9.

Figure 10 shows the resulting circular trajectory of the vector of rotational force, which he describes in one revolution. The oscillograms of changes of all three components of the force along the coordinate axes are also given (the axial force Pz remains unchanged).

The displacement vector (shown in Fig. 11) describes an elliptical trajectory in the XY plane. Changing the component force Pz due to the angle ρ (constant RO) changes the modulus of the vector of displacements X , but the shape of the trajectory in the plane XY remains unchanged (elliptical).

And only in the case of $c_{x0} \neq c_{y0}$, the shape of the trajectory will be represented as an ideal circle, since the rigidity ellipsoid will be an ellipsoid of rotation.

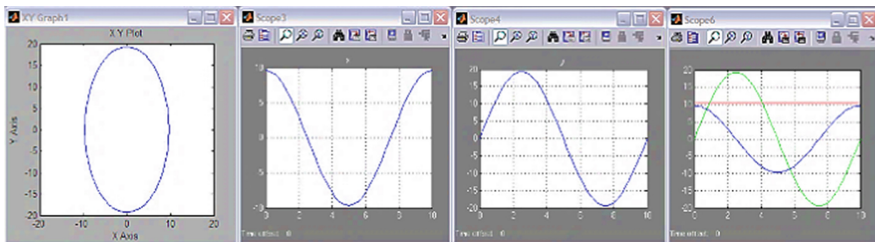


Fig. 11. The trajectory of the displacement vector with components

When machining on boring machines, it is necessary to ensure minimal machining errors, that is, the ellipsoid of rigidity (pliability) of the SU must be an ellipsoid of rotation. However, due to the structural or technological features of the machines and their components, this is not always possible.

The presence of an ellipsoid of rigidity of the SU elastic system and the corresponding elliptical trajectory of the cutting tool leads to the appearance of macro errors of machining of parts - ellipticity, ovality, beating of the treated surfaces.

Similar ellipsoids of rigidity or pliability can be constructed for other knots of metal-cutting machines (tables, revolving heads, rotary tables, etc.).

The interaction of knots (and their ellipsoids) in cutting zone is the subject of further research.

4 Conclusions

A tensor approach to the estimation of the spatial rigidity of the SU of the machine in the form of surfaces of ellipsoids of rigidity or pliability is proposed.

The theoretical substantiation of the tensor-geometric model of the spatial rigidity of the SU allows to estimate the elastic-deformation properties of the structure and to construct an appropriate ellipsoid of the rigidity (pliability) of the SU. With the known magnitudes of force loads with the help of an ellipsoid it is possible to find the spindle deformations and to estimate the magnitudes of the machining errors.

A mathematical simulation establishes a direct relationship between the shape of the rigidity (pliability) ellipsoid and the elliptical shape of the cutting tool trajectory during boring, which affects the precision of machining - the ellipse of the parts.

References

1. Altintas, Y., Cao, Y.: Virtual design and optimization of machine tool spindles. *Ann. CIRP* **54**, 379–382 (2005)
2. Cao, Y., Altintas, Y.: A general method for the modeling of spindle-bearing system. *J. Mech. Des. ASME* **126**, 1089–1104 (2004)
3. Danilchenko, Yu., Kuznetsov, Yu.: Precision spindle nodes on rolling bearings. *Economic Thought, Ternopil* (2003). (in Ukrainian)

4. Ilyin, V., Poznyak, E.: Analytical Geometry: A Textbook for High Schools. Fizmatlit, Moscow
5. Ito Y.: Modular Design for Machine Tool. McGraw-Hill, New-York (2008). 504 p.
6. Kochinev, N.A., Sabirov, F.S.: Quasi-static method of measuring the balance of elastic displacements of the supporting system of machine tools. *Meas. Tech.* **49**(6), 572–578 (2006)
7. Klee, H., Allen, R.: Simulation of Dynamic Systems with MATLAB and Simulink. CRC Press (2011). 815 p.
8. Lebedev, L.P., Cloud, M.J., Veremejev, V.: Tensor Analysis with Applications in Mechanics. World Scientific Publishing Company (2010). 363 p.
9. Poduraev, Yu.: Mechatronics: foundations, methods, application: educational. Pos. for high schools. Mechanical engineering, Moscow (2006). (in Russian)
10. Push, A.V.: Spindle units: quality and reliability. Mechanical Engineering, Moscow (1992). (in Russian)
11. Badrawy, S.: Dynamic modeling and analysis of motorized milling spindles for optimizing the spindle cutting performance. *Slash Documents, Science & Technology*, 4th October 2012 (2012). 18 p.
12. Strutinsky, V.: Tensor mathematical models of processes and systems. ZhDTU, Zhytomyr (2005). (in Ukrainian)
13. Strutinsky, V., Chupryna, V.: Multicriterial optimization of layouts of spindle knots of metal-cutting machine tools on rolling bearings. *Armament Syst. Mil. Equip.* **3**(47), 106–111 (2016). (in Ukrainian)
14. Strutynsky, V., Fedorynenko, D.: Statistical dynamics of spindle units for hydrostatic bearings. LLC “Publishing” Aspect-Polygraph, Nizhin (2011). (in Ukrainian)
15. Strutinsky, V., Kirichenko, A.: Theoretical analysis of the stiffness of a six-coordinate mechanism of a parallel structure. *Bull. Natl. Tech. Univ. Ukraine “Kyiv Polytech. Inst.”* Kyiv, 198–207 (2009). (in Ukrainian)

**Mathematical Modeling and Simulation
of Systems in Information Technology
and Information Security**



Hybrid Convolutional Neuro-Fuzzy Networks for Diagnostics of MRI-Images of Brain Tumors

Yuriy Zaychenko¹  and Galib Hamidov²

¹ Igor Sikorsky Kyiv Polytechnic Institute, Kiev, Ukraine

zaychenkoyuri@ukr.net

² Azershig, Baku, Azerbaijan

galib.hamidov@gmail.com

Abstract. The problem of medical images of brain analysis and classification of detected tumors into two classes: benign and malignant is considered.

For brain tumors recognition hybrid convolutional; network was developed in which CNN VGG-16 and ResNetV2_50 were used for feature extraction and FNN ANFIS- for classification of detected tumors. For ANFIS training the adaptive stochastic gradient method was suggested and implemented and its efficiency was explored. For preventing overfitting two layers of dropout were added. As a loss-function binary cross-entropy function was used. As the optimization algorithm Adam W with technique Warm-up was used. The experimental investigations of the suggested hybrid CNN-ANFIS network in the problem of classification real images were carried out on the special data set Brain MRI images for brain tumor detection. The comparisons of classification accuracy of the suggested CNN-ANFIS network with results of convolutional networks with classification algorithms SVM, NN and dimensionality reduction were performed which confirmed the reasonableness of application of hybrid networks for medical images recognition. In general the classifier SVM has shown the best results for considered problem.

Keywords: Medical diagnostics · Brain tumor · Classification · ANFIS · CNN · Hybrid network

1 Introduction. State-of-Art Analysis of Brain Tumor Classification

One of the actual spheres of information technologies application is brain tumors diagnostics based on medical images processing. In general, diagnostics of brain tumors is complicated due to different localizations of tumors in the brain. Finally diagnosis of oncological disease in a brain is made after histological conclusion as without histological investigation diagnosis isn't valid [2]. Diagnosis of brain tumor is made step by step: first in a dispensary, then it's confirmed in a hospital, Usually 3 stages can be distinguished in diagnostics process.

At the first stage the ill patient addresses to a doctor in polyclinic with some claims on bad state of brain activity or pain. The doctor makes preliminary observation of a

patient state and in dependence on his state transfer him to a hospital or prescribe dispensary treatment. Under detecting some neurological symptoms a patient is directed to neurologist for a consultation. Computer tomography of brain is made for detection of possible oncological pathology. After analysis of KT-image of brain tumor the neurologist makes preliminary diagnosis and after additional tests and observation he puts a clinical diagnosis and directs the patient to an oncologic dispensary.

To mandatory methods of observation of a patient in a dispensary are CT and MRT investigations. Now MRT with contrastive amplification is a “golden standard” in brain tumors diagnostics [3, 5]. In a case of detection at the MRT image a tumor the patient is directed to a hospital for the stationary treatment. After entering a patient into a hospital or oncological dispensary the several observation are made and decision is taken concerning the strategy of treatment: is it necessary to make surgical operation or use radiant or chemical therapy.

Till now the most of works based on histology brain cancer analysis were performed on not large datasets. Some improvement presents data set with 7909 breast images obtained from 82 patients [7]. In this research the authors estimated various texture descriptors and various classifiers and carried out the experiments with accuracy from 82% to 85%.

The alternative to this approach is the application of convolutional neural networks (CNN) for medical images processing and diagnostics, which is considered and developed in the present research.

It was shown that CNN is able to overcome the conventional texture descriptors [2, 9, 10]. Besides traditional approach to detection of features based on descriptors demands much efforts and high level knowledge of experts and usually is specific for every task that prevents its direct application for another similar tasks.

In the previous work new hybrid CNN_FNN network for classification medical images of breast cancer was developed in which CNN VGG 16 was applied for informative features extraction while fuzzy neural network NEF Class was used as classifier [12]. The experimental investigations of the efficiency of hybrid CNN-FNN network were carried out and comparison with CNN using conventional classifiers was performed. The experiments had shown the rather high accuracy of the suggested approach.

The interest represents investigation of application of other FNN for classification of detected tumors on medical images. In the present paper new hybrid convolutional-fuzzy neural network is suggested in which CNN is used for features extraction while fuzzy neural network ANFIS is used for classification of detected tumors and their classification into benign and malignant ones.

The main goal of this paper is the development of hybrid CNN-ANFIS network and investigation its efficiency in the problem of classification medical images of brain tumors and comparison its efficiency with known classifiers on the standard dataset of medical images [13].

2 CNN Model for Medical Images Classification

In the Fig. 1 the architecture of CNN VGG-16 is presented which is used as a detector of informative features. It was trained by stochastic gradient method at a dataset Image net. As alternative approach was used another CNN ResNetV2_50 (Fig. 2). This architecture is more deep and has better performance.

As a classifier in this paper it is suggested to use FNN ANFIS unlike known works in which MLP and SVM were used as classifiers.

FNN ANFIS is based on fuzzy logic inference system Takagi-Sugeno [11]. This network was developed at the beginning of 1990 years. As this network principles of neural networks are integrated with principles of fuzzy logic systems it enables to unite their advantages in one structure. The inference of this network is based on set of fuzzy rules “if-then” which may be trained that to attain desired output. Therefore, ANFIS is a universal approximator like other fuzzy logic systems. For efficient application of ANFIS parameters training by genetic algorithm may be used.

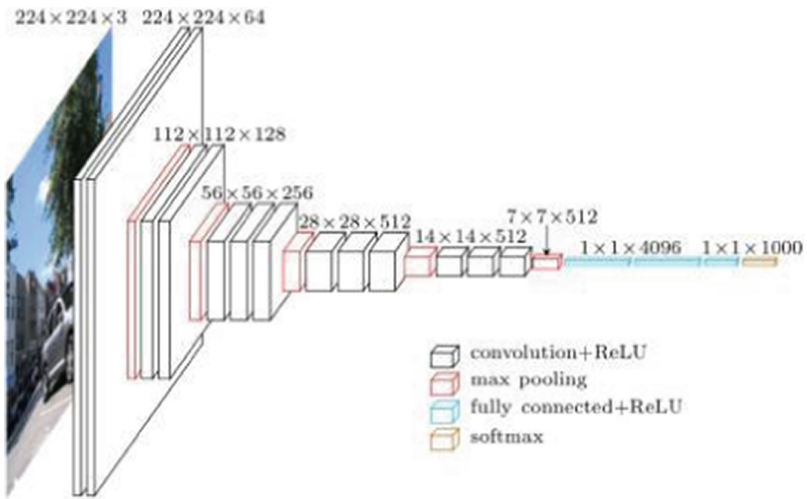


Fig. 1. Convolutional network VGG-16

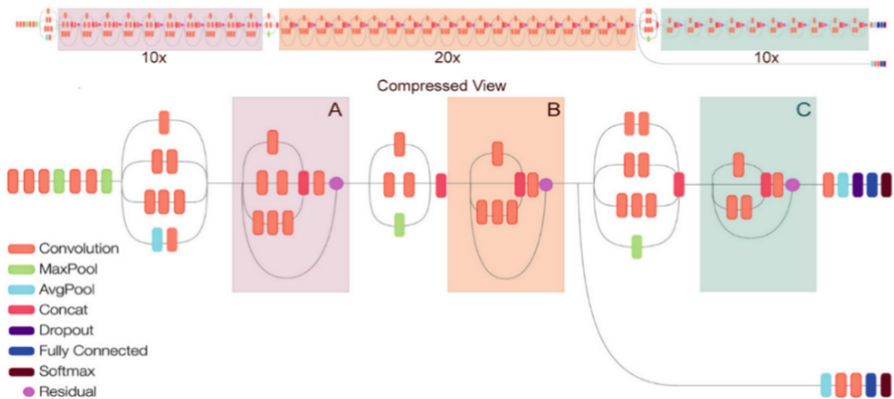


Fig. 2. Convolutional network ResNetV2

3 Dataset

All the Initial data for experiments were taken from site Kaggle [13]. As a whole 253 images were used which contained 98 MRI brain images without tumors (Fig. 3) and 155 images with brain tumors (Fig. 4). All the sample was divided into training and test subsamples in ratio 80%/20%.

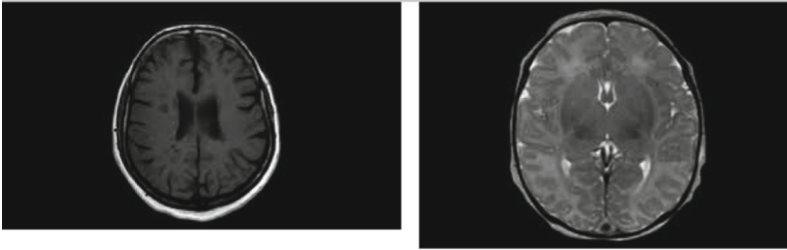


Fig. 3. MRI brain images without tumor

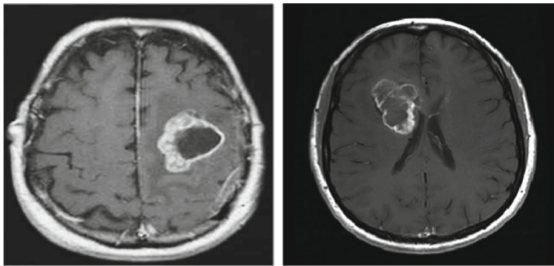


Fig. 4. MRI brain images with tumor

For augmentation were used: data shift in horizontal and vertical directions, brightness change, horizontal mirroring and rotation up to 15%.

4 Experimental Investigations and Their Analysis

In this investigation preliminary training of convolutional networks CNN VGG 16 and ResNet_V2 was performed. There are two main ways of training CNN:

- **Feature extraction.** In this case full-connected layers are dropped out and the rest network is used as feature extractor in new datasets;
- **Fine Tuning.** In this case new dataset is used for fine tuning of preliminary trained neural network.

In our paper CNN VGG-16 was used for features extraction in medical images of tumors. Thereafter the obtained features were entered into ANFIS. For its training stochastic gradient method and its modifications were applied.

Numerous experiments with suggested hybrid CNN network were carried out and comparison with known results was performed. The obtained results are presented below.

In the Table 1 the best results of FNN ANFIS are presented and alternative network structures. Volume of training was limited by 50 epochs.

Table 1. Classification results for different CNN architectures

№	Architecture of CNN	Accuracy	Specificity	Precision	Recall	F1 score
1.	VGG16 + NN + Warm-up AdamW	0.8526	0.8649	0.8607	0.9052	0.8514
2.	VGG16 + ANFIS + Warm-up AdamW	0.8368	0.7671	0.8839	0.8462	0.8378
3.	VGG16 + SVM + Warm-up AdamW	0.8737	0.8784	0.8710	0.9310	0.8722
4.	ResnetV2_50 + DR + AdamW	0.7789	0.8052	0.7840	0.8673	0.7789

In the architecture №. 1 CNN VGG16 was used whose output was transferred to form batch size x number of features using operation Flatten. For obtained features as classifier was used NN whose structure is described in the Table 2. For convergence acceleration was added layer batch normalization. For preventing overfitting was added dropout layer with value of $p = 0.35$. For this network binary cross-entropy loss function was used as criterion optimization. As optimization algorithm Adam W [15] with application of technique Warm-up [16, 17] was applied.

Table 2. Architecture of applied CNN

Network layer	Form of layer output	Number of parameters
VGG16 (Model)	(None, 7, 7, 512)	14714688
Flatten	(None, 25088)	0
Batch Normalization	(None, 25088)	100352
Dropout	(None, 25088)	0
Dense	(None, 1)	25089

Total number of parameters: 14840129

Number of trainable parameters: 75265

Number of non- trainable parameters: 14764864

In the architecture №2 as a base CNN VGG16 was applied whose output was transferred to a form batch size x number of features by operation Flatten. After this that to cut number of features full-connected layer was added. To these features analysis FNN ANFIS was applied as classifier (see Table 3).

Rules number was set equal to 32. For preventing overfitting two layers of dropout were added with parameters values 50% and 25% correspondingly. As a loss-function binary cross-entropy function was used. As optimization algorithm Adam W [16] with technique Warm-up [6, 17] was used.

Table 3. ANFIS classifier

Network layer	Form of layer output	Parameters number
VGG16 (Model)	(None, 7, 7, 512)	14714688
Flatten	(None, 25088)	0
Dropout	(None, 25088)	0
Dense	(None, 8)	200712
Dropout	(None, 8)	0
ANFIS	(None, 1)	544
Dense	(None, 1)	2

Total number of parameters: 14915 946

Number of trainable parameters: 201258

Number of non- trainable parameters: 14714668

In the architecture №3 CNN VGG 16 was used whose output was transferred to a form batch size x number of features by the operation Flatten. As a classifier was used SVM (Table 4).

Table 4. SVM classifier.

Network layer	Form of layer output	Parameters number
VGG16 (Model)	(None, 7, 7, 512)	14714688
Flatten	(None, 25088)	0
Dense	(None, 1)	25089

Total number of parameters: 14739777.

Number of trainable parameters: 25089.

Number of non-trainable parameters: 14714688.

Parameter “regularization force” was set to 0.01 and a loss function Hinge Loss [14] was used. As optimization algorithm Adam W [15] with technique Warm-up was applied [16, 17].

In the architecture №4 as a base CNN ResNet50_v2 was used, whose output was transferred to a form batch size x number of filters with operation GlobalAveragePooling_2D, which for 3xD kernel of filter finds average for each 2xD-space that to obtain 1D-kernel.

To these features convolutional layers were applied for cutting dimension. After output dimension reduction neural network was used as classifier (see Table 5). For convergence improvement layers of batch-normalization were added. For preventing overfitting a dropout layer was added with parameter $p = 25\%$ at convolutional layers

and $p = 30\%$ at full-connected layer. For this network binary cross-entropy loss function was utilized. As optimization algorithm Adam W was used.

Table 5. Architecture of CNN ResNet50_v2.with convolutional layers for dimension reduction

Network layer	Form of output layer	Parameters number
resnet50v2 (Model)	(None, 7, 7, 2048)	23564800
Global Average Pooling 2D	(None, 2048)	0
Reshape	(None, 2048, 1)	0
Conv1D	(None, 512, 8)	264
Max Pooling 1D	(None, 256, 8)	0
Batch Normalization	(None, 256, 8)	32
Dropout	(None, 256, 8)	0
Conv1D	(None, 64, 16)	4112
Max Pooling 1D	(None, 32, 16)	0
Batch Normalization	(None, 32, 16)	64
Dropout	(None, 32, 16)	0
Conv1D	(None, 16, 32)	16416
Max Pooling 1D	(None, 8, 32)	0
Batch Normalization	(None, 8, 32)	128
Dropout	(None, 8, 32)	0
Flatten	(None, 256)	0
Dropout	(None, 256)	0
Dense	(None, 1)	257

Total number of parameters: 23586073

Number of trainable parameters: 21161.

Number of non- trainable parameters: 23564912.

After experimental results analysis it was established that the best results were obtained with SVM classifier. This method has the best accuracy by criteria Accuracy, Specificity, Recall and F1 Score. But the best result by criterion (metrics) Precision was obtained by ANFIS classifier.

The application of technique of dimension reduction has shown weak results and this technique demands more detailed investigations and more data. The application of neural network as classifier has shown middle results.

In a whole for improving results and more reliable conclusions it's desirable to increase the data volume.

5 Conclusions

- The problem of medical images of brain analysis and classification of detected tumors into two classes: benign and malignant is considered.

- For brain tumors recognition hybrid convolutional; network was developed in which CNN VGG-16 was used for feature extraction and FNN ANFIS- for classification of detected tumors.
- For ANFIS training the adaptive stochastic gradient method was suggested and implemented and its efficiency was explored;
- The experimental investigations of the suggested hybrid CNN-ANFIS network in the problem of classification real images were carried out on the special data set Brain MRI images for brain tumor detection.
- The comparison of classification accuracy of the suggested CNN-ANFIS network with results of convolutional networks with classification algorithms SVM, NN и dimensionality reduction confirmed the reasonableness of application of hybrid networks for medical images recognition.
- In general the classifier SVM has shown the best results for considered problem.
- But the results may be substantially improved with extending dataset volume.

References

1. Boyle, P., Levin, B. (eds.): World Cancer Report 2012. IARC, Lyon (2012). http://www.iarc.fr/en/publications/pdfs-online/wcr/2008/wcr_2012.pdf
2. Dong, H., Yang, G., Liu, F., Mo, Y., Guo, Y.: Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks (2017)
3. Arya, P., Malviya, A.K.: A Survey on Brain Tumor Detection and Segmentation from Magnetic Resonance Image (2019)
4. Zhang, Y., Zhang, B., Coenen, F., Xia, J., Lu, W.: One-class kernel subspace ensemble for medical image classification. EURASIP J. Adv. Signal Process. **2014**(17), 1–13 (2014)
5. Rajan, P.G., Sundar, C.: Brain Tumor Detection and Segmentation by Intensity Adjustment (2017)
6. Singh, A., Mansourifar, H., Bilgrami, H., Makkar, N., Shah, T.: Classifying Biological Images Using Pre-trained CNNs. <https://docs.google.com/document/d/1H7xVK7nwXcv11CYh7h15F6pM0m218FQloAXQODP-Hsg/edit?usp=sharing>
7. Spanhol, F.A., Oliveira, L.S., Petitjean, C., Heutte, L.: A dataset for breast cancer histopathological image classification. IEEE Trans. Biomed. Eng. **63**(7), 1455–1462 (2016)
8. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**, 436–444 (2015)
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, vol. 25, pp. 1097–1105 (2012)
10. Nauck, D., Kruse, R.: New learning strategies for NEFCLASS. In: Proceedings of Seventh International Fuzzy Systems Association World Congress IFSA 1997, vol. IV, pp. 50–55. Academia Prague (1997)
11. Zgurovsky, M., Zaychenko, Y.: The Fundamentals of Computational Intelligence: System Approach, p. 308. Springer, Switzerland (2016)
12. Zaychenko, Y., Hamidov, G., Varga, I.: Medical images of breast tumors diagnostics with application of hybrid CNN–FNN network. Syst. Res. Inf. Technol. (4), c. 37–47 (2018)
13. Brain MRI Images for Brain Tumor Detection. <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>

14. Understanding loss functions: Hinge loss. <https://medium.com/analytics-vidhya/understanding-loss-functions-hinge-loss-a0ff112b40a1>
15. Decoupled weight decay regularization. <https://arxiv.org/pdf/1711.05101.pdf>
16. A Closer Look at Deep Learning Heuristics: Learning rate restarts, Warmup and Distillation. <https://arxiv.org/abs/1810.13243>
17. On the Difficulty of Warm-Starting Neural Network Training. <https://arxiv.org/pdf/1910.08475.pdf>



Method for Automatic Analysis of Compliance of Settlements with Suppliers and Settlements with Customers by Neural Network Model of Forecast

Tatiana Neskorodieva¹(✉)  and Eugene Fedorov² 

¹ Donetsk National University named Vasyl Stus,
600-Richchia str., 21, Vinnitsa 21021, Ukraine
t.neskorodieva@donnu.edu.ua

² Cherkasy State Technological University, Shevchenko blvd., 460,
Cherkasy 18006, Ukraine
fedorovee75@ukr.net

Abstract. The problem of analyze automation of the audit data of the pre-requisite “Compliance of costs and incomes” based on the forecast is considered. A neural network model for forecast based on a gateway recurrent unit is proposed. For parametric identification of this model, adaptive cross entropy is proposed. This allows you to increase the forecast efficiency by reducing computational complexity and improving the forecast accuracy.

Software was developed using the Matlab package that implements the proposed method. The developed software is studied when solving the problem of forecasting indicators in the task of analyzing the data mapping “settlements with suppliers - settlements with customers”.

Keywords: Automatic analysis · Audit data · “Settlements with suppliers - settlements with customers” mapping · Forecast · Neural network · Gateway recurrent unit

1 Introduction

In the process of development of international and national economies and industry of IT in particular, it is possible to distinguish the following basic tendencies: realization of digital transformations, forming of digital economy, globalization of socio-economic processes and of IT accompanying them [1]. These processes result in the origin of global, multilevel hierarchical structures of heterogeneous, multivariable, multifunction connections, interactions and cooperation of managing subjects (objects of audit), the large volumes of information about them have been accumulated in the informative systems of account, management and audit.

In relation to the enterprises of Ukraine [2] it is marked that in the conditions of swift development of infrastructure and deepening of informatization of economic processes efficiency of activity of enterprises, establishments and organizations are all more depend on the information technologies (IT) used in management system. Now

environment of information technologies (IT-environment) as a structural constituent of organization is the difficult system, that unites various informative, programmatic, technical, human and other types of resources for the achievement of aims of enterprise.

These tendencies determine conceptions of development of modern audit and corresponding to them information technologies of treatment of financial and economic information about activity of enterprise and decision support systems in an audit on the basis of the automated analysis of the large volumes of data about financial and economic activity and states of enterprises with the multi-level hierarchical structure of heterogeneous, multivariable, multifunction connections, intercommunications and cooperation of objects of audit with the purpose of expansion of functional possibilities, increase of efficiency and universality of IT-audit [3].

Automated DSS audit means the automatic forming of recommendable decisions, based on the results of the automated analysis of data, that improves quality process of audit. Unlike the traditional approach, computer technologies of analysis of data in the system of audit accelerate and promote the process accuracy of audit, that extremely critical in the conditions of plenty of associate tasks on lower and middle levels, and also amounts of indexes and supervisions in every task.

According to the international standard of audit № 520 "Analytical procedures" [4], analytical procedures are the procedures, suggesting the analysis of financial indexes and tendencies with the subsequent study of rejections and relations that conflict with other corresponding information or deviate from the forecast sums.

In the last few years among analytical procedures of exposure of financial fraud the class of procedures, based on the methods of intellectual analysis of data, was formed [5]. In practice different methods of extraction of data, namely: K-nearest neighbors, decision tree [6], fuzzy logic [7], logistic model, Bayesian belief networks, naive Bayes algorithm, Beneish M-Score, Benford's law, Altmann Z-Score, are being used for the improvement of accuracy of the fraud finding out [5].

One of the elements of analysis among the tasks of audit is a forecast of economic indicators. Methods of short-term forecast, such as regressive [8], structural [9], logical [10] conduct forecast on the basis of indexes in current moment to time. Methods of long-term forecast, such as autoregressive [11], exponential smoothing out [12] can conduct a forecasts on the basis of long temporal row.

A compromise between the above-mentioned groups of methods is connectionist [13], that can use neural networks both for a short-term and for long-term forecasts. At the same time connectionist methods usually use parametric identification on the basis of local search, that reduces forecast accuracy.

Connectionist methods conduct a forecast on the basis of nonlinear model of neural network. The advantages of these methods are scalability, high adaptivity. The disadvantages of these methods are absence of design transparency; complexity of choice of model structure; hard requirements to training sample; problem of choice of method of parametric identification; that results in insufficient forecast accuracy high computational complexity of parametric identification. The general feature of all the methods mentioned above is that they possess a high computational complexity and/or do not give high forecast accuracy. Therefore an actual task is an increase of forecast efficiency by reducing computational complexity and improving forecast accuracy.

The aim of study is to increase the efficiency of automatic data analysis in DSS audit based on the forecast method due to the model of the gateway recurrent unit and the method of its parametric identification.

For the achievement of the aim it is necessary to solve the following tasks:

- to offer the forecast neural network model due to the model of the gateway recurrent unit;
- to choose the criterion of estimation of efficiency of forecast neural network model;
- to offer a method for parametric identification of a neural network prediction model based on adaptive cross-entropy;
- to offer the method for automatic analysis of audit data;
- to conduct numeral researches.

2 Formal Problem Statement

In the DSS of audit, the verification of pre-condition ‘‘Compliance of expenses and income’’ of accounting provisions (standards) [14] at the average level is decomposed on the verification tasks of mapping of generalized values indicators: ‘‘settlements with suppliers - settlements with customers’’, ‘‘accounts payable - debt’’.

The conversion of the data of settlements with suppliers into the data on the receipt of funds from customers in the accounting system can be represented as a functional structure of the data subsets transformation at the stages of operating activities (Fig. 1).

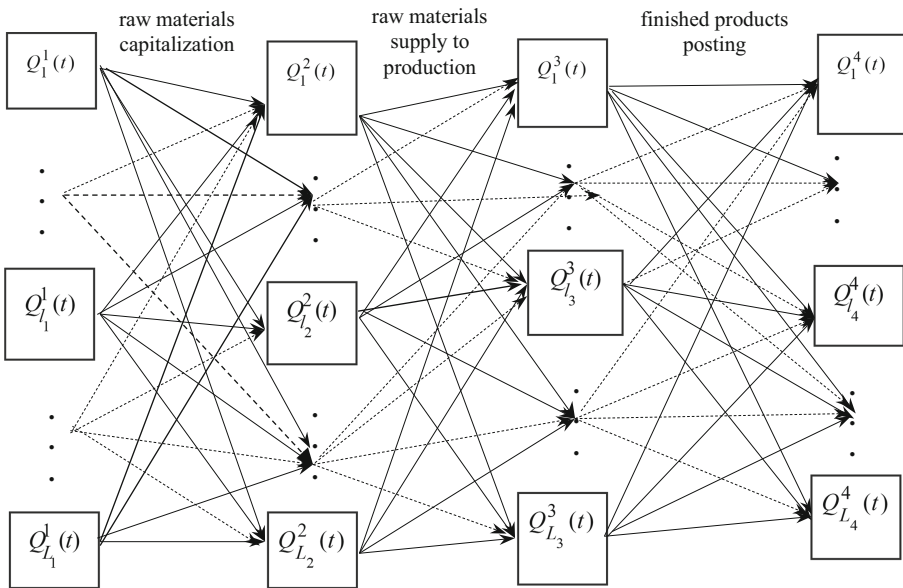


Fig. 1. Functional structure of the data subsets transformation at the stages on operating activities

Denote:

$Q_{l_1}^1(t)$, $l_1 = \overline{1, L_1}$ are data sets of calculations by type of supplier l_1 ,

$Q_{l_2}^2(t)$, $l_2 = \overline{1, L_2}$ are data sets of stocks by type of raw material l_2 ,

$Q_{l_3}^3(t)$, $l_3 = \overline{1, L_3}$ are production datasets by type l_3 ,

$Q_{l_4}^4(t)$, $l_4 = \overline{1, L_4}$ are data sets of calculations by type of finished product l_4 ,

$t \in \{t_{j_m}, T_m, j = \overline{1, J_m}, m = \overline{1, M}, T\}$

The verification of mappings of sets of the functional structure (Fig. 1) is preceded by an analysis of the values of quantitative indicators for the verification period with the purpose of exposure of values that deviate from the forecast. The forecast values are determined on the basis of conformities to law, that is formed from the tested data for other (as a rule, preceding) periods that is imported from the database of the system of account of enterprise in DSS of audit.

We will distinguish the first mapping «settlements with suppliers - settlements with customers». The values of the audit data at the verification stage of this mapping are formed on the basis of a set of quantitative indicators characterizing the transfer of funds to suppliers (raw materials and/or components) and the receipt of funds from customers for goods sold:

$$(V_u, \Delta_u), u \in U, (V_g, \Delta_g), g \in G,$$

where:

V - quantitative indicator (in physical units),

Δ - monetary indicator (amounts transferred by suppliers/customers),

u - type of raw material,

U - set of types of raw material,

g - type of finished product,

G - set of types of finished products.

At the analysis of this mapping (in direct and reverse direction) as elements of pairs of learning set can be chosen:

$$x = (V_u(t), V_u(t-1), \dots, V_u(t-M)), u \in U_g, y = (V_g(t)), g \in G,$$

where:

x is input signal,

M is time of delay,

U_g - set of raw materials that are used in the manufacture of finished products of the type g ,

G_u - finished products set, the production of which uses raw materials of the type u ,

y - is output signal.

As the value of the lag M or lead can be selected the value of the indicator of the difference of the operating cycle (the average time between the operation of purchasing raw materials and the sale of finished products made from this raw material) and the

production cycle. This indicator depends on the industry, type of production and range of products ($1 < M < 30$). Information for prior to the verification period is considered reliable. Therefore, the corresponding data can be selected as a training set.

A comparison of the predicted value for the model and the value of the control sample will allow us to identify the types of finished products and stocks and quantization periods for which there are significant deviations and which will be recommended by the decision maker for detailed study at the lower level.

So, for a forecast a learning set is given $S = \{(x_\mu, d_\mu)\}$, $\mu \in \overline{1, P}$. Then, problem of increase of forecast accuracy on the model of the gateway recurrent unit (GRU) $g(x, W)$, here x - an input signal, W is a vector of parameters, appears as a problem of being for this model of such vector of parameters W^* , that satisfies to the criterion

$$F = \frac{1}{P} \sum_{\mu=1}^P (g(x_\mu, W^*) - d_\mu)^2 \rightarrow \min. \quad (1)$$

3 Forecast Model Based on a Gateway Recurrent Unit

Gateway recurrent unit (GRU) [15] has two gateways. Instead of a hidden neuron, a hidden block h is used, which is connected to the reset gateway r and update z . Gateways determine how much information to skip [16].

If the reset gateway is close to 1 and the update gateway is close to 0, then we get an Elman neural network (ENN), in which there is no control over the amount of information. If the reset gateway and update gateway are close to 0, then we get a multilayer perceptron (MLP), in which information from the hidden block (long-term) is ignored. If the update gateway is close to 1, then the input (short-term) information is ignored.

The forecast model by the gateway recurrent unit (2)–(7) is defined as follows:

$$y_i^{in}(n) = x_i, \quad (2)$$

$$r_j(n) = f \left(b_j^r + \sum_{i=1}^M w_{ij}^{in-r} y_i^{in}(n) + \sum_{i=1}^{N^h} u_{ij}^{h-r} h_i(n-1) \right), j \in \overline{1, N^h}, \quad (3)$$

$$z_j(n) = f \left(b_j^z + \sum_{i=1}^M w_{ij}^{in-z} y_i^{in}(n) + \sum_{i=1}^{N^h} u_{ij}^{h-z} h_i(n-1) \right), j \in \overline{1, N^h}, \quad (4)$$

$$\tilde{h}_j(n) = g \left(b_j^{\tilde{h}} + \sum_{i=1}^M w_{ij}^{in-\tilde{h}} y_i^{in}(n) + \sum_{i=1}^{N^h} u_{ij}^{h-\tilde{h}} r_i(n) h_i(n-1) \right), j \in \overline{1, N^h} \quad (5)$$

$$h_j(n) = z_j(n) h_j(n-1) + (1 - z_j(n)) \tilde{h}_j(n), j \in \overline{1, N^h} \quad (6)$$

$$y^{out}(n) = f\left(b^{out} + \sum_{i=1}^{N^h} w_i^{h-out} h_i(n)\right), \quad (7)$$

were:

- M – number of unit delays of the input layer,
- N^h – hidden layer neurons number,
- w_{ij}^{in-r} – weights between the input layer and the reset gateway,
- w_{ij}^{in-z} – weights between the input layer and the update gateway,
- $w_{ij}^{in-\tilde{h}}$ – weight between input and candidate layer,
- u_{ij}^{h-r} – weights between the hidden layer and the reset gateway,
- u_{ij}^{h-z} – weights between the hidden layer and the update gateway,
- $u_{ij}^{h-\tilde{h}}$ – weight between hidden and candidate layer,
- w_i^{h-out} – weight between the hidden and the output layer,
- in and out - input and output layer, respectively,
- h and \tilde{h} - hidden and candidate layer, respectively,
- r and z - reset and update gateway, respectively.

In paper for model training GRU (2)–(7) the target function is selected, which means the choice of such values of the parameter vector $W = (w_{11}^{in-r}, \dots, w_{MN^h}^{in-r}, u_{11}^{h-r}, \dots, u_{N^h N^h}^{h-r}, w_{11}^{in-z}, \dots, w_{MN^h}^{in-z}, u_{11}^{h-z}, \dots, u_{N^h N^h}^{h-z}, w_{11}^{in-\tilde{h}}, \dots, w_{MN^h}^{in-\tilde{h}}, u_{11}^{h-\tilde{h}}, \dots, u_{N^h N^h}^{h-\tilde{h}}, w_1^{h-out}, \dots, w_{N^h}^{h-out})$, which deliver a minimum of the standard error (1) (the difference between the model output and the test output).

4 Method for Parametric Identification of a Model of a Gateway Recurrent Unit Based on Adaptive Cross Entropy

Cross entropy method (CE) [17] consists of two phases - the generation of a new population of potential solutions and the modification of the Gaussian distribution parameters. In this paper, to control the convergence rate of CE and to ensure that the entire search space is investigated at the initial iterations and the search becomes directional at the final iterations, the iteration number is taken into account when generating potential solutions.

The proposed method consists of the following steps:

1. Initialization
 - 1.1. Setting a parameter α that controls the rate of change of the Gaussian distribution parameters, and $0 < \alpha < 1$.
 - 1.2. Setting a parameter β for generating a standard deviation vector, and $0 < \beta < 1$.

- 1.3. Setting the maximum number of iterations N , population size K , solution length M (corresponds to the length of the GRU model parameter vector), the maximum number of best solutions selected.
- 1.4. Random creating a vector of mathematical expectations
 $\mu = (\mu_1, \dots, \mu_M)$, $\mu_j = x_j^{\min} + (x_j^{\max} - x_j^{\min})U(0, 1)$, where $U(0, 1)$ – function that returns a uniformly distributed random number in a range $[0, 1]$.
- 1.5. Create a standard deviation vector randomly

$$\sigma = (\sigma_1, \dots, \sigma_M), \sigma_j = \beta(x_j^{\max} - x_j^{\min})U(0, 1).$$

- 1.6. Determine the best solution (best GRU model parameter vector) $x^* = \mu$.
2. Iteration number $n = 1$.
3. Create a current population of potential solutions P .
 - 3.1. Solution number $k = 1$, $P = \emptyset$.
 - 3.2. Generate a new potential solution x_k .
 $x_{kj} = \mu_j + \sigma_j \left(\frac{N-n}{N}\right)N(0, 1)$, $j \in \overline{1, M}$, where $N(0, 1)$ – function that returns a standard normally distributed random number.
 - 3.3. If $k < K$, then $P = P \cup \{x_k\}$, $k = k + 1$, go to step 3.2.
 - 3.4. Sort P by target function, i.e. $F(x_k) < F(x_{k+1})$.
 - 3.5. $k^* = \arg \min_k F(x_k)$.
 - 3.6. If $F(x_{k^*}) < F(x^*)$, then $x^* = x_{k^*}$.
4. Modification of the Gaussian distribution parameters (based on the first B , i.e. best, new potential solutions from the population P).
 - 4.1. $\mu_j = \alpha\mu_j + (1 - \alpha)\tilde{\mu}_j$, $\tilde{\mu}_j = \frac{1}{B} \sum_{k=1}^B x_{kj}$, $j \in \overline{1, M}$.
 - 4.2. $\sigma_j = \alpha\sigma_j + (1 - \alpha)\tilde{\sigma}_j$, $\tilde{\sigma}_j = \frac{1}{B} \sum_{k=1}^B (x_{kj} - \tilde{\mu}_j)^2$, $j \in \overline{1, M}$.
5. If $n < N$, then $n = n + 1$, go to step 3.

The result is x^* .

5 Experiments and Results

Numerical experiments were carried out using the package Matlab. The number of single delays (time lag) $M = 10$, neurons number of hidden layer $N^h = 2M$.

To determine the structure of the forecast model based on the gateway recurrent unit (GRU), a number of experiments were carried out, the results of which are presented in Fig. 2. As the initial data for determining the values of the parameters of the forecast model over the neural network, calculation indicators with suppliers and customers of a machine-building enterprise were used with a two-year sampling depth with daily time intervals, the values of which are indicators representing trade secrets, and they are scaled. The criterion for choosing the structure of the neural network model was the minimum mean square error of the forecast. As can be seen from Fig. 2, with an increase in the number of hidden neurons, the error value decreases. For

prediction, it is enough to use 10 time delays in the input layer and 20 hidden neurons, since with a further increase in the number of delays and hidden neurons, the error value changes insignificantly.

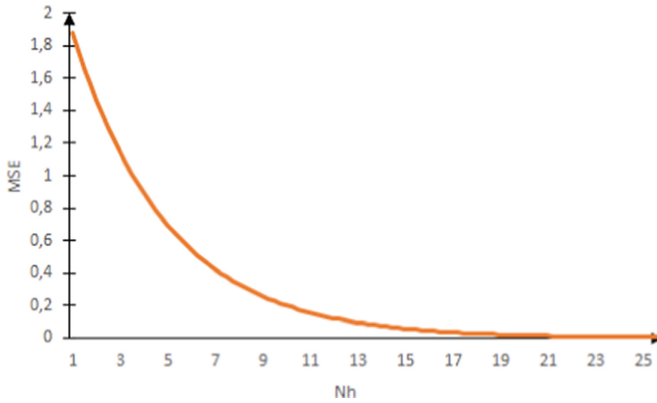


Fig. 2. The dependence of the MSE of the forecast on the hidden neurons number

In this work, we studied neural networks for prediction by the criterion of the minimum mean square error (MSE) of the forecast. The following neural networks were used in the experiments (Table 1): JNN (Jordan neural network), ENN (Elman neural network) also called SRN (Simple recurrent network), NARMA (nonlinear autoregressive-moving average), BRNN (bidirectional recurrent neural network), LSTM (long short-term memory), GRU (gated recurrent unit). According to Table 1, LSTM and GRU have the greatest prediction accuracy, however, LSTM requires more computational complexity due to more connections between neurons. This is due to the fact that the LSTM and GRU neural networks use control over the amount of information passed through.

Table 1. Comparative characteristics of neural networks for prediction

Network	NAR	JNN	ENN (SRN)	NARMA	BRNN	LSTM	GRU
Minimum MSE of forecast	0.2	0.14	0.17	0.16	0.13	0.08	0.08

This model was tested on the values of category *X* indices (coefficient of variation is in the range from 0 to 10%), *Y* (coefficient of variation is in the range from 10 to 25%) and *Z* (coefficient of variation is more than 25%). Comparing the values of the coefficient of variation and the forecast error in the GRU network, which amounted to 5%, for the resource category *X* for the forecast period from 1 to 26 days, for the resource category *Y* for the forecast period from 1 to 12 days, for the resource category *Z* for the forecast period from 1 to 5 days, we can conclude that this network is suitable for forecasting indicators in the tasks of DSS audit.

6 Conclusions

The article considers the problem of increasing the efficiency of automatic data analysis in DSS of audit based on the forecast due to the gateway recurrent unit (GRU). Reached the further development of the method of parametric identification of the model GRU, which is based on adaptive cross entropy (ACE), which increases the accuracy of the forecast, because at the initial iterations the entire search space is investigated, and at the final iterations the search becomes directional. Software implementing the proposed method in the package Matlab, was developed and researched on the indicators of supply and payment of stocks of a machine-building enterprise with a two-year sampling depth with daily time intervals. The experiments confirmed the operability of the developed software and allow us to recommend it for use in practice in the automated analysis subsystem of the DSS of audit to verify the totality of the displays of the sets of calculation sets with suppliers and customers. The prospects for further research are to test the proposed methods on a wider set of test databases.

References

1. The World Bank: World Development Report 2016: Digital Dividends (2016). <https://www.worldbank.org/en/publication/wdr2016>. Accessed 12 Feb 2020
2. The State Audit Service of Ukraine: Praktychna metodolohiia IT-audytu. (Practical methodology of IT-audit) (2015). <http://dkrs.kmu.gov.ua/kru/doccatalogdocument/>. Accessed 12 Feb 2020
3. Zhu, B.: Research on the application of big data in audit analysis program. Paper presented at the International Seminar on Automation, Intelligence, Computing, and Networking, Paris, 29 April–2 May 2019. <https://doi.org/10.25236/isaicn.2019.047>
4. The International Auditing and Assurance Standards Board: International standards of account and audit (2016). <https://www.iaasb.org/standards-pronouncements>. Accessed 12 Feb 2020
5. Alex, G., Pereira, G., Pappa, L.: A customized classification algorithm for credit card fraud detection. *Eng. Appl. Artif. Intell.* **72**, 21–29 (2018)
6. Zgurovsky, M.Z., Zaychenko, Y.P.: *The Fundamentals of Computational Intelligence: System Approach*. Springer, Heidelberg (2017)
7. Vlasenko, A., Vlasenko, N., Vynokurova, O., Bodyanskiy, Y., Peleshko, D.: A Novel ensemble neuro-fuzzy model for financial time series forecasting. *Data* **4**(3), 126 (2019). <https://doi.org/10.3390/data4030126>
8. Khrystyanivsky, V.V., Neskorodieva, T.V., Polshkov, Yu.N.: *Ekonomiko-matematicheskie metody i modeli: praktika primeneniya v kursovyih i diplomnyih rabotah* (Economic and mathematical methods and models: application practice in term papers and dissertations). DonNU, Donetsk (2013)
9. Grimshaw, S., Alexander, W.: Markov chain models for delinquency: transition matrix estimation and forecasting. *Appl. Stochast. Models Bus. Ind.* **27**(3), 267–279 (2011). <https://doi.org/10.1002/asmb.827>
10. Emmerich, M., Lytvyn, V., Yevseyeva, I., Fernandes, V.B., Dosyn, D., Vysotska, V.: Preface: modern machine learning technologies and data science (MoML&T&DS-2019). In: *CEUR Workshop Proceedings*, vol. 2386 (2019)

11. Hu, Z., Bodyanskiy, Y., Tyshchenko, O.: Self-Learning and Adaptive Algorithms for Business Applications. Emerald Publishing Limited (2019) <https://doi.org/10.1108/978-1-83867-171-620191001>
12. Bidyuk, P., Prosyankina-Zharova, T., Terentiev, O.: Modelling nonlinear nonstationary processes in macroeconomy and finances. In: Hu, Z., Petoukhov, S., Dychka, I., He, M. (eds.) Advances in Computer Science for Engineering and Education. Advances in Intelligent Systems and Computing, vol. 754, pp. 735–745. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-91008-6_72
13. Tkachenko, R., Izonin, I., Greguš, M., Tkachenko, P., Dronyuk, I.: Committee of the SGTM neural-like structures with extended inputs for predictive analytics in insurance. In: Younas, M., Awan, I., Benbernou, S. (eds.) Big Data Innovations and Applications. Innovate-Data 2019. Communications in Computer and Information Science, vol. 1054. pp 121–132 Springer, Cham (2019). <https://doi.org/10.1007/978-3-030-27355-2>
14. Neskorođieva, T.V.: Postanovka elementarnykh zadach audytu peredumovy polozhen bukhholderskoho obliku v informatsiinii tekhnolohii systemy pidtrymky rishen (Formulation of elementary tasks of audit subsystems of accounting provisions precondition IT DSS). Mod. Inf. Syst. 3(1), 48–54 (2019). <https://doi.org/10.20998/2522-9052.2019.1.08>
15. Haykin, S.S.: Neural Networks and Learning Machines. Pearson, Delhi (2016)
16. Bengio, Y., Goodfellow, I.J., Courville, A.: Deep Learning. The MIT Press Cambridge, Massachusetts (2016). <http://www.deeplearningbook.org>
17. Cho, K., Merriënboer, B., Gulcehre, C., Bougares, F, Schwenk, H., Bengio, Y.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, pp. 1724–1734 (2014)



Mathematical Modeling of Non-Gaussian Dependent Random Variables by Nonlinear Regression Models Based on the Multivariate Normalizing Transformations

Sergiy Prykhodko^(✉)  and Natalia Prykhodko 

Admiral Makarov National University of Shipbuilding,
Mykolaiv 54025, Ukraine
{sergiy.prykhodko,natalia.prykhodko}@nuos.edu.ua

Abstract. The technique to construct nonlinear regression models based on the multivariate normalizing transformations and prediction intervals is considered. We demonstrate that there may be data sets for which the results of constructing the nonlinear regression models depend on firstly, which normalizing transformation is applied, univariate, or multivariate, and, secondly, are there any outliers in the data set. The application of multivariate normalizing transformations allows us to take into account the correlation between variables in the nonlinear regression model. We use the prediction intervals of nonlinear regressions to detect the outliers in the process of building the nonlinear regression models. In general, this process is iterative since we repeat constructing the model for new data after the outlier cutoff. The example of constructing a three-factor nonlinear regression model to simulate the non-Gaussian dependent random variable depend on three predictors is given. There are six iterations in the process of constructing this nonlinear regression model. We performed comparing the constructed model to the linear regression one and nonlinear regression models that are built by the univariate normalizing transformations. This model has larger values of multiple coefficient of determination, and prediction percentage, smaller values of the magnitude of relative error, and widths of the confidence and prediction intervals of regression compared to other models, both linear and nonlinear ones. Suitable values of prediction accuracy metrics for the constructed model may be explained better multivariate normalization of the data set and the outlier cutoff, which are used to build one.

Keywords: Mathematical modeling · Dependent random variable · Nonlinear regression model · Normalizing transformation · Prediction interval

1 Introduction

It is known, the acceptance-rejection method [1] may be used to simulate the non-Gaussian dependent random variable by the non-Gaussian joint distribution in the form of the probability density function. However, constructing the non-Gaussian joint

distribution is a complex problem, which makes it challenging to use this method for mathematical modeling and simulation of non-Gaussian dependent random variables.

Also, nonlinear regression models may be applied for mathematical modeling and simulation of non-Gaussian dependent random variables. According [2], a nonlinear regression model is defined as “a model for the relationship between a response and predictor(s) in which at least one parameter does not enter linearly into the model.” A nonlinear regression model in its structure contains an error term (a random variable), which is included additively [2–7] or multiplicatively [6, 7].

In [3], Bates and Watts indicate that the assumption of additivity of error is closely tied to the assumption of the constant variance of the disturbances (residuals). It may be the case that the residuals can be considered as having a constant variance, but as entering the model multiplicatively. In either case, one of the corrective actions is to take a transformation of the response (dependent variable). If the error is multiplicative, we can treat the nonlinear regression model as intrinsically linear and use the normalizing transformation [8]. That is, we define the nonlinear regression model as in [9] $Y = f(x, \theta, \varepsilon)$, where f is a nonlinear function; x is a vector of regressors (independent variables); θ is a vector of parameters; ε is the error term that has the same properties as in linear regression, i.e., the Gaussian random variable to describe residuals, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

In [9], we considered the techniques to build the models, confidence, and prediction intervals of nonlinear regressions based on the bijective multivariate normalizing transformations. However, there may be data sets for which the results of creating nonlinear regression models depend on, firstly, which normalizing transformation is applied, univariate, or multivariate, and, secondly, are there any outliers in the data set. In this paper, we propose the technique to construct nonlinear regression models based on the multivariate normalizing transformations and prediction intervals. We use the prediction intervals of nonlinear regressions to detect the outliers in the process of building the nonlinear regression models. In general, this process is iterative since we repeat constructing the model for new data after the outlier cutoff.

2 The Technique

We consider the technique to construct nonlinear regression models based on the multivariate normalizing transformations and prediction intervals. The technique consists of four steps. Firstly, we normalize a multivariate non-Gaussian data set by using a multivariate normalizing transformation. Secondly, we construct the nonlinear regression model based on the multivariate normalizing transformation [9]. Thirdly, we build the prediction interval of nonlinear regression, such as [9]. All in all, we check if there are data that exit the bounds of the prediction interval. If we detect the outliers, we discard them and repeat all the steps, starting with the first, for new data.

In the first step, multivariate non-Gaussian data of dependent and independent variables are normalized using a multivariate normalizing transformation of non-Gaussian random vector $P = \{Y, X_1, X_2, \dots, X_k\}^T$ to Gaussian one $T = \{Z_Y, Z_1, Z_2, \dots, Z_k\}^T$, which is given by [9]

$$T = \psi(P), \tag{1}$$

and which has the inverse transformation

$$P = \psi^{-1}(T). \tag{2}$$

In the second step, the nonlinear regression model is constructed based on the multivariate normalizing transformation (1), according to [9]. To do this, we first build the linear regression model for normalized data based on the transformation (1)

$$Z_Y = \hat{Z}_Y + \varepsilon = \bar{Z}_Y + (Z_X^+)^T \hat{b} + \varepsilon, \tag{3}$$

where \hat{Z}_Y is prediction linear equation result for values of components of vector $z_X = \{Z_1, Z_2, \dots, Z_k\}$; k is a number of dependent variables (regressors); Z_X^+ is the matrix of centered regressors that contains the values $Z_{1i} - \bar{Z}_1, Z_{2i} - \bar{Z}_2, \dots, Z_{ki} - \bar{Z}_k$; \hat{b} is a vector of estimators of parameters for of linear regression equation, $b = \{b_1, b_2, \dots, b_k\}^T$.

After that, we construct the nonlinear regression model using the inverse transformation (2) for the model (3)

$$Y = \psi_Y^{-1} [\bar{Z}_Y + (Z_X^+)^T \hat{b} + \varepsilon], \tag{4}$$

where ψ_Y is the component of the vector ψ of (1), $\psi = \{\psi_Y, \psi_1, \psi_2, \dots, \psi_k\}^T$.

In the third step, the prediction interval of nonlinear regression is built [9]

$$\psi_Y^{-1} \left(\hat{Z}_Y \pm t_{\alpha/2, v} S_{Z_Y} \left\{ 1 + \frac{1}{N} + (z_X^+)^T [(Z_X^+)^T Z_X^+]^{-1} (z_X^+) \right\}^{1/2} \right), \tag{5}$$

where $t_{\alpha/2, v}$ is a student's t -distribution quantile with $\alpha/2$ significance level and v degrees of freedom; z_X^+ is a vector for i -row, $z_X^+ = \{Z_{1i} - \bar{Z}_1, Z_{2i} - \bar{Z}_2, \dots, Z_{ki} - \bar{Z}_k\}$; $S_{Z_Y}^2 = \frac{1}{v} \sum_{i=1}^N (Z_{Yi} - \hat{Z}_{Yi})^2$, $v = N - k - 1$; N is a number of rows of data; $(Z_X^+)^T Z_X^+$ is a $k \times k$ matrix

$$(Z_X^+)^T Z_X^+ = \begin{pmatrix} S_{Z_1 Z_1} & S_{Z_1 Z_2} & \dots & S_{Z_1 Z_k} \\ S_{Z_1 Z_2} & S_{Z_2 Z_2} & \dots & S_{Z_2 Z_k} \\ \dots & \dots & \dots & \dots \\ S_{Z_1 Z_k} & S_{Z_2 Z_k} & \dots & S_{Z_k Z_k} \end{pmatrix},$$

where $S_{Z_q Z_r} = \sum_{i=1}^N [Z_{qi} - \bar{Z}_q][Z_{ri} - \bar{Z}_r]$, $q, r = 1, 2, \dots, k$.

Finally, we check if there are data that exit the bounds of the prediction interval. And, if we detect the outliers, we discard them and repeat all the steps, starting with the first, for new data.

3 Model Construction Example

We consider the example of constructing a three-factor nonlinear regression model to simulate the non-Gaussian dependent random variable Y depend on three predictors, which are denoted as $X_1, X_2,$ and X_3 . We construct an improved model for the data set from [10]. The dependent random variable Y is the efforts (in person-hours) of the mobile applications (apps) development in a planning phase. The predictors (regressors) $X_1, X_2,$ and X_3 are numbers of screens, functions, and files, respectively. The data set from [10] includes 38 data rows.

At first, the data set from [10] was checked for four-variate outliers. As in [10], we used the squared Mahalanobis distance (SMD) for detecting multivariate outliers [11, 12]. There are no four-variate outliers in the 38 data rows from [10] for a significance level of 0.005 since all SMD values are less than the Chi-Square distribution quantile.

The three-factor nonlinear regression model to simulate the non-Gaussian dependent random variable was constructed based on the Johnson four-variate transformation for the S_B family, such as [9]. The nonlinear regression model has the form [10]

$$Y = f(x, \hat{\theta}, \varepsilon) = \hat{\phi}_Y + \hat{\lambda}_Y \left[1 + e^{-(\hat{Z}_Y + \varepsilon - \hat{\gamma}_Y) / \hat{\eta}_Y} \right]^{-1}, \tag{6}$$

where $x = \{X_1, X_2, X_3\}$; $\hat{\theta}$ is the estimator of the vector of parameters θ , $\hat{\theta} = \{\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{\gamma}_Y, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3, \hat{\eta}_Y, \hat{\eta}_1, \hat{\eta}_2, \hat{\eta}_3, \hat{\phi}_Y, \hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \hat{\lambda}_Y, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3\}$; \hat{Z}_Y is a prediction result of the linear regression equation for data, which were normalized by Johnson’s four-variate transformation for the S_B family, $\hat{Z}_Y = \hat{b}_0 + \hat{b}_1 Z_1 + \hat{b}_2 Z_2 + \hat{b}_3 Z_3$; $Z_j = \gamma_j + \eta_j \ln \frac{X_j - \phi_j}{\phi_j + \lambda_j - X_j}$, $\phi_j < X_j < \phi_j + \lambda_j, j = 1, 2, 3$. In the first iteration, for the 38 data rows from [10], the parameter estimators of Johnson’s four-variate transformation for the S_B family are: $\hat{\gamma}_Y = 5.69898, \hat{\gamma}_1 = 0.524119, \hat{\gamma}_2 = 0.776179, \hat{\gamma}_3 = 0.540973, \hat{\eta}_Y = 2.40219, \hat{\eta}_1 = 0.743879, \hat{\eta}_2 = 0.79545, \hat{\eta}_3 = 0.534447, \hat{\phi}_Y = -114.5452, \hat{\phi}_1 = 1.7242, \hat{\phi}_2 = 1.6885, \hat{\phi}_3 = 0.90, \hat{\lambda}_Y = 3328.564, \hat{\lambda}_1 = 12.3743, \hat{\lambda}_2 = 12.091, \hat{\lambda}_3 = 8.30648$; the parameter estimators of the linear regression equation for normalized data are: $\hat{b}_0 = 0, \hat{b}_1 = 0.808152, \hat{b}_2 = -0.928296, \hat{b}_3 = 0.854262$.

The well-known prediction accuracy metrics are used to judge the prediction accuracy of regression models [13]. These metrics include a mean magnitude of relative error MMRE, a multiple coefficient of determination R^2 , and prediction percentage at the level of magnitude of relative error (MRE) of 0.25, PRED(0.25). The R^2 , MMRE, and PRED(0.25) values equal respectively 0.5789, 0.4933, and 0.5263 for nonlinear regression model (6) with the estimators of parameters which are calculated for the 38 data rows from [10]. In this case, the above values demonstrate to us unacceptable prediction results of the nonlinear regression model (6) approximately also as for the linear regression model [10].

Next, we calculated the prediction intervals of the three-factor nonlinear regression by (5) in the first iteration. In Table 1, the lower (LB) and upper (UB) bounds of the prediction interval found in the first iteration are marked as $LB_1,$ and $UB_1,$ respectively [10].

Table 1. Lower and upper bounds for nonlinear regression before and after outlier rejection.

No	Y	LB_1	UB_1	LB_6	UB_6	No	Y	LB_1	UB_1	LB_6	UB_6
1	192	60.5	377.3	141.6	216.6	20	198	70.8	402.2	160.4	237.6
2	272	60.5	377.3	–	–	21	146	49.2	353.3	124.8	197.8
3	288	88.6	524.1	233.6	323.2	22	191	66.2	392.5	151.1	227.6
4	116	51.1	352.9	115.6	185.7	23	99	24.7	290.0	76.4	136.2
5	372	54.5	362.4	–	–	24	382	140.1	624.9	326.6	397.3
6	504	90.1	453.3	–	–	25	270	93.4	477.2	218.5	301.1
7	28	–0.7	232.5	25.1	64.6	26	282	104.6	532.5	246.7	331.1
8	176	18.9	277.8	–	–	27	213	78.5	452.7	181.0	264.8
9	364	157.4	665.2	342.2	409.2	28	322	126.8	560.3	278.4	354.9
10	120	48.7	363.8	110.8	183.4	29	290	109.1	513.1	239.9	320.2
11	22	70.8	402.2	–	–	30	223	78.6	425.2	177.7	257.1
12	224	73.5	447.0	171.7	256.4	31	241	84.9	449.3	204.6	286.3
13	24	–23.9	170.9	15.4	49.1	32	87	17.1	267.3	53.0	103.8
14	200	106.5	511.6	–	–	33	36	–29.0	153.6	15.2	47.4
15	160	100.6	490.0	–	–	34	216	77.1	418.6	168.5	246.6
16	120	–23.9	170.9	–	–	35	67	1.4	233.2	29.0	69.6
17	96	–33.4	149.2	–	–	36	115	31.0	306.4	72.4	131.5
18	202	70.8	402.2	160.4	237.6	37	36	–23.9	170.9	15.4	49.1
19	145	49.2	353.3	124.8	197.8	38	98	24.7	290.0	76.4	136.2

We calculated the prediction intervals for a significance level of 0.05.

After that, we checked if there are data that exit the bounds of the prediction interval. We detected the three outliers (row 5, 6, and 11), discarded them, and repeated all the steps, starting with the first, for 35 data rows.

In the second iteration, we constructed the model by using data from the remaining 35 rows. For the model (6) with the parameter estimators obtained from the data of the 35 data rows, it appeared that the value of Y for row 17 exceed the prediction interval. There were six iterations; as a result, 29 data rows were left (1, 3, 4, 7, 9, 10, 12, 13, 18–38). At the sixth iteration, there were no outliers; the repetition of the steps was completed, the model (6) was constructed by using data from 29 data rows. In Table 1, the LB of prediction interval obtained in the sixth iteration is marked as LB_6 and the UB - as UB_6 . Also, in Table 1, a dash (-) depicts the elimination of the corresponding numbers of data in the relevant iteration. The row numbers with the outliers in data are bolded.

In the sixth iteration, for the data from 29 rows, the estimators of parameters of Johnson’s four-variate transformation for the S_B family are: $\hat{\gamma}_Y = 0.638164$, $\hat{\gamma}_1 = 0.387413$, $\hat{\gamma}_2 = 0.840380$, $\hat{\gamma}_3 = 0.477514$, $\hat{\eta}_Y = 1.12311$, $\hat{\eta}_1 = 0.659463$, $\hat{\eta}_2 = 0.831303$, $\hat{\eta}_3 = 0.632614$, $\hat{\phi}_Y = -28.4433$, $\hat{\phi}_1 = 1.82645$, $\hat{\phi}_2 = 1.59121$, $\hat{\phi}_3 = 0.657479$, $\hat{\lambda}_Y = 543.161$, $\hat{\lambda}_1 = 11.5548$, $\hat{\lambda}_2 = 12.9939$, $\hat{\lambda}_3 = 8.63368$; the estimators for parameters of the linear regression equation for normalized data are: $\hat{b}_0 = 0$, $\hat{b}_1 = 1.17702$, $\hat{b}_2 = -1.43269$, $\hat{b}_3 = 1.18398$.

In the sixth iteration, for 29 rows of data which normalized by using Johnson’s four-variate transformation for the S_B family, 3×3 matrix

$$(Z_X^+)^T Z_X^+ = \begin{pmatrix} 29.0 & 24.8 & 18.7 \\ 24.8 & 29.0 & 23.6 \\ 18.7 & 23.6 & 29.0 \end{pmatrix}.$$

The R^2 , MMRE, and PRED(0.25) values equal respectively 0.984, 0.103, and 0.862 for nonlinear regression model (6). These values indicate to us good prediction results of the three-factor nonlinear regression model (6) for parameter estimators calculated from the 29 data rows. Also, we compared the model (6) with other models, the linear regression model, and nonlinear regression models based on the univariate decimal logarithm transformation (Log10) and Johnson’s univariate transformation. These models were derived in [10]. The R^2 , MMRE, and PRED(0.25) values equal respectively 0.838, 0.237 and 0.733 for linear regression model, and equal respectively 0.789, 0.206 and 0.733 for the nonlinear regression model based on the Log10, and equal 0.878, 0.190 and 0.767 for the model (6) with parameter estimators for Johnson’s univariate transformation. The R^2 , MMRE, and PRED(0.25) values are better for the model (6) with parameter estimators obtained for Johnson’s four-variate transformation in comparison with the above models.

The LB and UB of prediction intervals of the linear regression and nonlinear regressions were also calculated by (5) based on the Log10 transformation, the Johnson univariate, and four-variate transformations for 0.05 significance level. The above bounds are given in Table 2.

Table 2. Lower and upper bounds of prediction intervals for regressions.

No	Y	Linear regression		Log10 univariate		Johnson univariate		Johnson four-variate	
		LB	UB	LB	UB	LB	UB	LB	UB
1	192	85.9	257.1	107.8	309.2	44.4	340.6	141.6	216.6
3	288	56.8	233.5	111.5	351.7	136.9	384.7	233.6	323.2
4	116	81.5	251.6	89.7	256.0	40.7	332.4	115.6	185.7
7	28	-72.1	115.9	24.9	78.1	23.0	162.5	25.1	64.6
9	364	234.9	420.4	164.5	495.7	181.8	385.1	342.2	409.2
10	120	77.1	260.2	95.4	281.0	41.5	340.4	110.8	183.4
12	224	128.7	314.9	100.3	317.6	68.2	370.7	171.7	256.4
13	24	-24.9	151.3	22.6	69.1	21.7	79.7	15.4	49.1
18	202	124.4	295.5	132.9	381.0	56.0	356.2	160.4	237.6
19	145	46.7	219.3	79.8	232.0	35.6	317.2	124.8	197.8
20	198	124.4	295.5	132.9	381.0	56.0	356.2	160.4	237.6
21	146	46.7	219.3	79.8	232.0	35.6	317.2	124.8	197.8
22	191	131.1	303.1	119.0	344.2	55.3	356.3	151.1	227.6

(continued)

Table 2. (continued)

No	Y	Linear regression		Log10 univariate		Johnson univariate		Johnson four-variate	
		LB	UB	LB	UB	LB	UB	LB	UB
23	99	15.0	188.4	48.1	140.4	27.3	256.1	76.4	136.2
24	382	221.9	409.1	160.7	485.7	71.6	375.0	326.6	397.3
25	270	198.8	379.5	144.9	432.9	98.2	376.9	218.5	301.1
26	282	167.2	353.5	144.1	440.2	290.2	387.7	246.7	331.1
27	213	142.1	325.8	126.0	398.9	61.6	364.9	181.0	264.8
28	322	239.9	420.5	243.4	723.3	197.5	384.9	278.4	354.9
29	290	202.4	381.1	215.0	640.8	124.4	379.7	239.9	320.2
30	223	168.7	342.5	141.0	410.4	72.7	368.0	177.7	257.1
31	241	187.9	371.8	159.0	481.3	82.6	373.3	204.6	286.3
32	87	-9.4	164.3	38.9	113.4	25.2	219.4	53.0	103.8
33	36	-36.5	138.6	19.1	58.3	21.4	50.7	15.2	47.4
34	216	151.0	321.6	141.8	406.2	67.3	364.3	168.5	246.6
35	67	-56.6	124.7	25.6	78.3	23.2	166.8	29.0	69.6
36	115	15.7	192.0	57.4	166.9	29.7	282.7	72.4	131.5
37	36	-24.9	151.3	22.6	69.1	21.7	79.7	15.4	49.1
38	98	15.0	188.4	48.1	140.4	27.3	256.1	76.4	136.2

Note that the width of the nonlinear regression prediction interval based on Johnson’s four-variate transformation is smaller than after Johnson’s univariate transformation for 28 from 29 data rows (except one with number 33), less than after Log10 transformation and smaller compared with the linear regression prediction interval width for all 29 data rows.

Such adequate prediction results for the created model might be explained by the better multivariate normalization of the non-Gaussian data, which we used to improve the three-factor nonlinear regression model based on the Johnson four-variate transformation for the S_B family and outlier rejection. Multivariate normality was tested by SMD [14]. A multivariate normality condition is only performed for the normalized data based on the Johnson four-variate transformation as for each row of the data; the SMD values are less than the Chi-Square distribution quantile, which equals to 11.14 for 0.025 significance level. Also, the values of the estimator of multivariate kurtosis β_2 [15] indicate good multivariate normality for the normalized data based on the Johnson four-variate transformation only. In our case $\beta_2 = 24$. The values of the estimator of multivariate kurtosis equal 34.82 and 25.09 for the 29 data rows, the normalized data based on the Johnson univariate and four-variate transformations, respectively.

4 Conclusions

Mathematical modeling of the non-Gaussian dependent random variable by nonlinear regression model using multivariate normalizing transformation is performed. The technique to construct nonlinear regression models based on the multivariate normalizing transformations and prediction intervals is further developed. A three-factor nonlinear regression model to simulate the non-Gaussian dependent random variable depend on three predictors is improved based on Johnson's four-variate transformation for the S_B family and discarding outliers by applying prediction intervals. This model, in comparison with other regression ones, both linear and nonlinear, has larger values of multiple coefficient of determination, and prediction percentage, less values of MMRE, and widths of the confidence and prediction intervals of regression. Suitable values of prediction accuracy metrics for the constructed model may be explained better multivariate normalization of the non-Gaussian data and the outlier rejection, which are applied to build one. Further research may be directed at the use of other multivariate normalizing transformations to derive the new multiple nonlinear regression models for mathematical modeling of the non-Gaussian dependent random variables.


References

1. Dieter, U.: Optimal acceptance-rejection methods for sampling from various distributions. In: Nelson, P.R. (ed.) *The Frontiers of Statistical Computation, Simulation, & Modeling* (1987). Am. Ser. Math. Manag. Sci
2. Chatterjee, S., Simonoff, J.S.: *Handbook of Regression Analysis*. Wiley, New York (2013)
3. Bates, D.M., Watts, D.G.: *Nonlinear Regression Analysis and Its Applications*, 2nd edn. Wiley, New York (1988). <https://doi.org/10.1002/9780470316757>
4. Drapper, N.R., Smith, H.: *Applied Regression Analysis*. Wiley, New York (1998)
5. Ryan, T.P.: *Modern Regression Methods*, 2nd edn. Wiley, New York (2008)
6. Seber, G.A.F., Wild, C.J.: *Nonlinear Regression*. Wiley, New York (1989). <https://doi.org/10.1002/0471725315>
7. Duncan, G.T.: An empirical study of jackknife constructed confidence regions in nonlinear regression. *Technometrics* **2**(20), 123–129 (1978). <https://doi.org/10.2307/1268703>
8. Freund, R.J., Wilson, W.J., Sa, P.: *Regression Analysis: Statistical Modeling of a Response Variable*, 2nd edn. Elsevier Academic Press, Burlington/London (2006)
9. Prykhodko, N.V., Prykhodko, S.B.: Constructing the non-linear regression models on the basis of multivariate normalizing transformations. *Electr. Model.* **6**(40), 101–110 (2018). <https://doi.org/10.15407/emodel.40.06.101>
10. Prykhodko, S., Prykhodko, N., Knyrik, K., Pukhalevych, A.: Mathematical modeling of effort of mobile application development in a planning phase. In: *Proceedings of the 1st International Workshop on Information-Communication Technologies & Embedded Systems*. CEUR Workshop Proceedings, Mykolaiv, Ukraine, 14–15 November 2019, vol. 2516, pp. 96–105. CEUR-WS.org (2019)
11. Johnson, R.A., Wichern, D.W.: *Applied Multivariate Statistical Analysis*. Pearson Prentice Hall, Upper Saddle River (2007)

12. Prykhodko, S., Prykhodko, N., Makarova, L., Pukhalevych, A.: Application of the squared Mahalanobis distance for detecting outliers in multivariate non-Gaussian data. In: Proceedings of the 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), pp. 962–965. IEEE, Lviv-Slavske (2018). <https://doi.org/10.1109/tcset.2018.8336353>
13. Port, D., Korte, M.: Comparative studies of the model evaluation criteria MMRE and PRED in software cost estimation research. In: Proceedings of the 2nd ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, pp. 51–60. ACM, New York (2008)
14. Olkin, I., Sampson, A.R.: Multivariate analysis: overview. In: Smelser, N.J., Baltes, P.B. (eds.) International Encyclopedia of Social & Behavioral Sciences, 1st edn. Elsevier, Pergamon (2001)
15. Mardia, K.V.: Measures of multivariate skewness and kurtosis with applications. *Biometrika* **3**(57), 519–530 (1970). <https://doi.org/10.1093/biomet/57.3.519>



Generation of Information Impacts Scenarios in Management Decision Support Systems

Andriy Boychenko¹ and Dmytro Lande^{1,2} 

¹ Institute for Information Recording of National Academy of Sciences of Ukraine, Kiev, Ukraine
boychenko.a@gmail.com, dwlande@gmail.com

² National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kiev, Ukraine

Abstract. The features of developing scenario models using cognitive maps and methods for developing of information impacts scenarios based on the analysis of the content of global computer networks are considered. This approach provides a solution to the problem of generating and ranking scenarios for impacting objects that correspond to the selected key concept analyzing the input text arrays working on a full-time basis. Improved information models and computer domain analysis tools have been developed. This makes it possible for an expert analyst to investigate these processes and generate results in a form convenient for decision-making. A method is proposed for constructing a domain model in the form of a semantic graph formed according to the monitoring of computer networks by determining the most significant concepts and the relationships between them. A method is proposed for the formation of optimal scenarios of informational impacts on target objects of a subject area based on finding many routes of influence distribution. Software and algorithmic tools for transferring data to the OWL format are developed. An example of application of the developed scenario approach is considered on the example of analysis of the movement of “Yellow vests” in France.

Keywords: Information impacts scenarios · Cognitive map · Ontology · Ant colony optimization · Management decision support system

1 Introduction

The growth and complexity of the information space requires scientists to immediately solve the problem of increasing the effectiveness of the information impact in the information-analytical component of modern computer networks. The aim of this activity is to ensure the information needs of society with the help of the most modern computer technologies, through the processing of enormous data sets and the acquisition of qualitatively new knowledge. As a result, analysts have to work with information resources that are unprecedented in their volumes, versatility, dynamism and growth rates. This forces specialists to improve methods and technologies for loading, structuring and analyzing various data.

Development of scenarios for the evolution of a situation is an important component of information security and computer systems and networks management. These scenarios provide an opportunity to investigate how significant is the impact of each influencing factor on the functioning and security of computer systems.

Without scenario modeling, it is impossible to achieve the following: support high-performance work and protect information in modern computer systems and networks; developing strategies and tactics at each level of management; determination of the effectiveness of management methods and their further improvement; rapid assessment of the effectiveness of control actions in a variety of directions and timely response to identified threats.

However, the construction and analysis of such models by traditional methods is ineffective in conditions of continuous change in the situation. Therefore, the task of developing new methods and scenario modeling techniques of information impacts by analyzing semantic models generated based on data obtained in the process of monitoring computer networks is prospective and practically useful.

2 Analysis of Recent Research and Publications

The current state of information impact modeling technologies lags behind the rapid development of the modern information society. This growth is characterized by the emergence of new forms of digital communication, the growth of data volumes and throughput of transmission channels, the improvement of the ways of interaction and mutual influence of the subjects of the information space.

This, in turn, significantly complicates the task of assessing the potential consequences of information operations in management decision support systems.

One of the ways to solve the problem of studying the nature of the distribution of influences in the information space is scenario modeling based on cognitive maps.

The first cognitive network was introduced in 1986 when Cosco (1986) improved the LCM approach proposed by Axelrod (1976) [1]. The most significant improvement made by Cosco is the integration of the concept of fuzzy logic. Scenario models are used in decision making, addressing perceptions of financial crisis policies, as well as the withdrawal, presentation, and analysis of mental models.

Carrying out these studies requires the use of artificial intelligence, and, first of all, systems and methods of decision support. Currently, many decision support systems have been developed and are being successfully applied in practice, the theoretical foundations of incentives and practical results of their implementation are reflected, for example, based on the use of cognitive maps in the study of political elites [1], preparation of decisions based on poorly structured data [2], ontological structures [3], development of scenarios using Text Mining [4].

In traditional decision support systems, a knowledge model (cognitive model) about a subject area is created with the assistance of knowledge engineers and is usually oriented toward specific tasks. The main tools for performing scenario modeling include the following: «KoCMoC»; «Канва»; PolyAnalyst; Deductor; Fuzzy Thought Amplifier; Cope; NIPPER; Gismo; iThink, Hyper; RESEARCH; FCMmapper etc.

Modeling of complex processes and systems remains a challenge. This is because such models need to generate networks that include numerous nodes and many different types of relationships.

3 Formulation of the Problem

A cognitive map is a signed oriented graph [5]: $G = \langle V, E \rangle$, where: V – the set of vertices V_i is $V, i = 1, 2, \dots, k$, which are elements of the system under study; E – the set of arcs e_{ij} is $E, i, j = 1, 2, \dots, N$, reflecting the relationship between the vertices V_i and V_j ; the effect of V_i on V_j can be positive when an increase (decrease) in one factor leads to an increase (decrease) in another; negative when an increase (decrease) in one factor leads to a decrease (increase) in the second, or to be absent (0).

The edges of the graph have weights +1 or -1, are abbreviated as “+” or “-”. A + sign denotes a positive relationship; a sign denotes a negative relationship. The weight of the path is equal to the product of the weights of its edges, that is, positive if the number of negative edges in it is even and negative if this number is odd. With a positive relationship, an increase in the factor-cause leads to an increase in the factor-effect, and with a negative connection, an increase in the factor-cause leads to a decrease in the factor-effect. If both positive and negative paths lead from vertex a_i to vertex b_j , the question of the nature of the influence of factor a_i on factor b_j remains uncertain.

The influence of the vertex a_i on a_j ($i, j = 1, 2, \dots, N$, where n is the number of vertices in the graph) is called the strongest if the influence on the final vertex a_j of the initial vertex a_i on the k th simple path is the largest in absolute the value among all actions on other simple paths connecting the vertices a_i and a_j .

When choosing scenario options, it is necessary to use the performance indicator for which the choice is made. Thus, the impact analysis provides:

1. The strength of the influence of one factor on another along a given path depends on the length of this path (that is, the number of edges in it).
2. The more parallel actions (in different ways) exist between concepts, the stronger the influence between them.

The task can be reduced to optimization tasks:

$$f(S_k) = \frac{\sum_{i=1}^{n-1} a_i w_{i+1}}{n} \rightarrow \max, \quad (1)$$

with restrictions: $n > 1$; $1 \geq a_i > 0$; $1 \geq w_{ij} > 0$, where:

a_i is the weight of the i -th concept;

w_{ij} – the value of the influence of concept a_i on concept a_j ;

n is the length of the k -th scenario.

4 Research Methodology

The key issue in scenario modeling is a cognitive map structure construction. For the subject area in question, the following groups of concepts can be distinguished in the map structure:

- factors characterizing the method (set of methods) of the assessment;
- destabilizing factors;
- indicators (factors reflect and explain the dynamics of the development of a problem situation).

When using cognitive maps, three options for modeling are possible:

- modeling with initial conditions without additional impact on the model;
- generation of scenarios for the evolution of a situation;
- selection of the value of target factors.

The list of concepts can be quite wide, however, taking into account changes in the state of a large number of concepts is a laborious task. Therefore, at the stage of formation of the map structure, it is necessary to ensure that the importance of each concept is assessed.

The concept of part *A* affects *B* if:

- in the non-directivity of the network, degree *A* of greater degree *B*;
- in sentences *A* is before *B*;
- term *A* is included in term *B* (for example, “student” → “student of KPI”).

The value of each concept depends on the values of the connected concepts with the corresponding weights and according to its previous value (Fig. 1).

When a map is built, it can be used to simulate and verify the operation of the system. We construct a matrix (Table 1). In order to better understand the nature of the mutual influence, we represent each card in the form of a square adjacency matrix. In this matrix, the variables acting as potential transmitters (affect other variables) on the vertical axis and the same set of variables that act as receivers (are influenced by other variables).

On the horizontal axis, we form a list of all individual variables compiled during the entire process and additional variables.

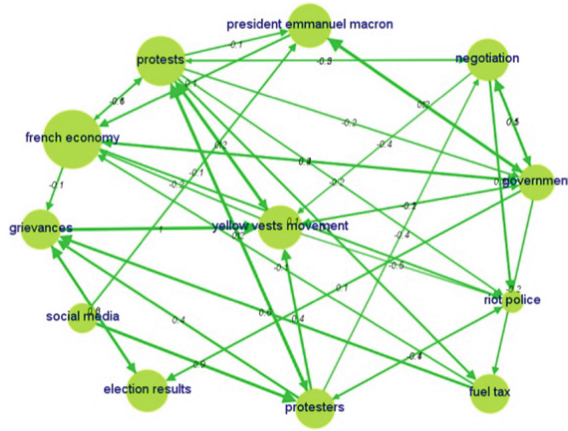


Fig. 1. The Cognitive map

Table 1. Matrix of mutual influences

Concept	Wight	president Emmanuel macron	election results	yellow vests movement	fuel tax	riot police	French economy	protests	negotiation	grievances	protesters	government	social media
president Emmanuel macron	1231	0	0	0	0	0	0,1	0	0	0	0	0,2	0
election results	1156	0	0	0	0	0	0	0	0	0,6	0	0	0
yellow vests movement	1218	0	0	0	0	-0,5	-0,2	1	0	0	0	-0,2	0,5
fuel tax	997	0	0	0	0	0	-0,1	0	0	0,6	0	0	0
riot police	208	0	0	0	0	0	0	0	0	0	-0,4	0	0
French economy	1940	0	0	0	0	0,1	0	-0,1	0	-0,1	0	0,4	0
protests	1500	-0,1	0	0,2	0,2	-0,2	-0,6	0	-0,2	0	0,2	-0,2	0,8
negotiation	1143	0	0	-0,4	0	0,3	0	-0,5	0	0	0	0,1	0
grievances	1033	0	0,2	1	0	0	0	0	0	0	0	0	0,8
protesters	1028	0	0	0,4	0	-0,1	0	1	-0,4	0,4	0	0	0,3
government	849	1	0,1	-0,1	-0,2	0	0,2	0	0,5	0	0	0	0
social media	565	-0,1	0	0	0	0	0	0	0	0	0,9	0	0

A qualitative analysis of the cognitive map (the content of its constituent blocks, target and control factors, analysis of paths and cycles, relationships between elements does not reveal the entire depth of phenomena and the process that takes place in a real system.

The next stage of the study is modeling the impulse process of the propagation of disturbances in the cognitive map, which leads to the transition of the system from one

state to another. Such a process is a possible scenario for the development of the system.

Optimization methods are suitable for fuzzy cognitive maps in which concepts can take values from a range of real numbers [0, 1]. The term “fuzzy” only means that the mutual relations between factors can take not only the values 0 or 1, but also belong to the range of real numbers, which allows modeling to more accurately express the mutual influence of factors [6].

To determine the value of the scenarios, it is proposed to use the ant algorithm (application of ant colony optimization algorithm [7], analysis of such algorithms [8], meta-heuristic approach in the task of optimizing an ant colony [9]), one of the effective polynomial algorithms for finding approximate solutions to the problems of finding routes on graphs. Using a modification of the ant algorithm, which provides a reduction in the time of formation of the script. The algorithm is based on the behavior of an ant colony – marking successful routes with a large amount of pheromone.

$$\begin{cases} P_{ij,k}(t) = \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}]^\beta}{\sum_{i \in J_{ik}} [\tau_{il}(t)]^\alpha \cdot [\eta_{il}]^\beta}, j \in J_{i,k}; \\ P_{ij,k}(t) = 0, j \notin J_{i,k}, \end{cases} \tag{2}$$

where

- $\tau_{ij}(t)$ is the number of pheromones in the transition from i to j ,
- η_{ij} is the attractiveness of the transition to j ,
- T_k – passing route.

Pheromone vapor is calculated in the following way:

$$\Delta\tau_{ij,k}(t) = \begin{cases} \frac{Q}{L_k(t)}, (i,j) \in T_k(t); [\\ 0, (i,j) \notin T_k(t), \end{cases} \tag{3}$$

where

- T_k – passing route,
- L_k is the length of the k -th route,
- J – number of unvisited nodes.

Unlike the traditional ant colony algorithm, where the edge weights are used together with the pheromone value to decide on the next vertex, the modified heuristic algorithm uses the dispersion importance of concepts – a term that is corroded to measure the advantage of some concepts over others (scripting based on analysis of terminological networks [10], creation of terminological ontologies [11], weight of concepts in integration of an ontological model with relational data [12]):

$$g_A = \frac{\sqrt{\langle \Delta A^2 \rangle - \langle \Delta A \rangle^2}}{\langle \Delta A \rangle}, \tag{4}$$

where

$\langle \Delta A \rangle$ – the average value of the sequence $\Delta A_1, \Delta A_2 \dots \Delta A_K$,
 $\langle \Delta A^2 \rangle$ – sequences $\Delta A_1^2, \Delta A_2^2 \dots \Delta A_K^2$.

The script execution program can be described in the form of an algorithm and in turn includes the execution of certain elementary operations.

The scenario structure is defined by a set of triples, which are put in accordance with the actions [10]:

$$AM_r = \langle \{X_i\}_{i=1}^{n_r}, A_r, X_{j_r}, \rangle, \tag{5}$$

where A_r is the action; $\{X_i\}_{i=1}^{n_r}$ – the set of input states AM_r for actions; X_{j_r} – the initial state of AM_r , that is, the state, after the successful completion of the action; AM is the set of all actions.

A complex scenario is described as follows:

$$SM_r = \langle \{X_i\}_{i=1}^{n_r}, S_r, X_{j_r}, \rangle, \tag{6}$$

where SM_r is the script, is the set of input states of SM_r , $\{X_i\}_{i=1}^{n_r}, X_{j_r}, X_{j_r}$ is the initial state of the script, that is, the state after the successful execution of the action; AM is the set of all actions.

$SM_r \in \{SM_k\} = SM$ – the set of all possible scenarios for the evolution of a situation.

We define $SV = AM \cup SM$, that is, the set of all models for the evolution of a situation and the corresponding scenarios.

As can be seen from Fig. 2, after the simulation, select 2 scenarios.

Scenario 1: No additional growth effects.

Scenario 2: Government – reduces the impact on fuel tax, instead increases the impact on negotiation and social media.

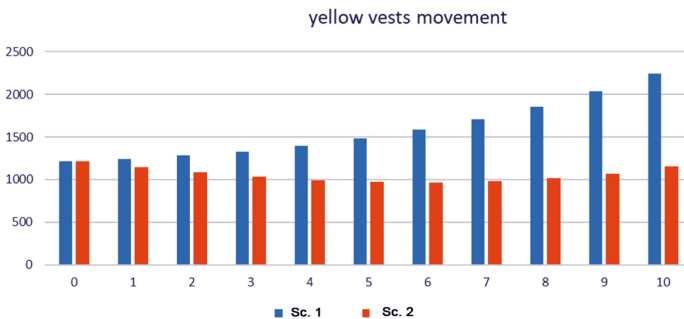


Fig. 2. Scenarios for the situation

5 Building an Ontology

Ontologies in computer science, this is the representation of knowledge about a specific subject area using a conceptual scheme (application of ontologies in influence models [13], fuzzy cognitive map based ontologies [14], ontology for applications development [15], web ontology language as standard [16], influences in ontological structures [17]). Typically, such a scheme consists of a data structure containing all relevant classes of objects, their relationships, as well as certain rules and restrictions inherent in a particular subject area (the modern direction in the field of artificial intelligence). In general terms, ontologies are defined as a knowledge base of a special kind, or as a “specification of conceptualization” of a subject area—conceptual graphs.

The general algorithm for transferring data from the semantic network to OWL is presented below (Fig. 3).

The Protégé 5 editor was chosen as the basic platform for developing ontologies in the design of knowledge-based analytical systems.

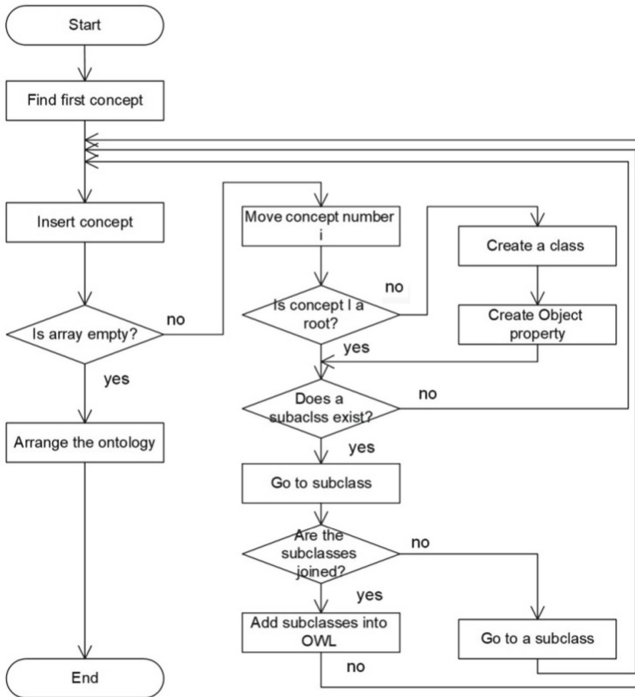


Fig. 3. An algorithm for transferring concepts to ontology

To integrate ontological and relational data models in the Protégé 5 editor, the Ontop platform is used, which is a concrete implementation of the data integration method based on the Ontology-Based Data Access approach (Protégé as ontology

editor and framework for systems building [18], information retrieval with ontologies [19], ontology and semantic management [20]).

To build the ontology, we apply the following steps:

- definition of terms;
- definition of synonyms;
- definition of concepts;
- building a concept taxonomy;
- definition of relationships;
- definition of rules (Fig. 4).

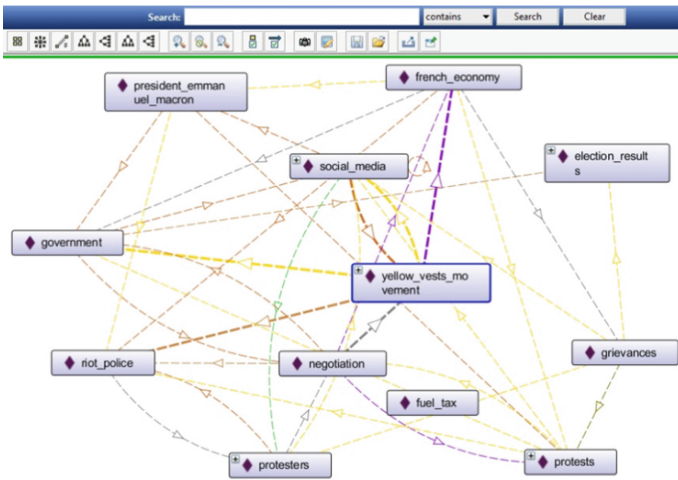


Fig. 4. Modeling of ontology with Protégé

The obtained information models and algorithms for the study of semantic networks are quite universal, due to which they can be used for control systems in various subject areas. In particular, when solving the problem of generating scenarios of analytical activity, a model was formed based on the graph shown in Fig. 5.

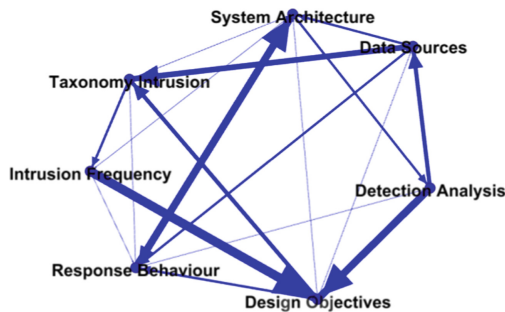


Fig. 5. Generalized graph of analytical activity factors

6 Conclusion

The article proposes an improved information technology for the automated formation of a semantic network by monitoring the information space and extracting concepts from it with the greatest frequency and taking into account their interaction. The essence of the improvement is to use an algorithm for building a network of concepts, which is based on simulating the interaction of ants – a set of dynamic mechanisms by which the program reaches a global goal as a result of the interaction of elements using only local data in combination with procedures for extracting textual information on a computer network and procedures domain research using ontologies.

The construction of cognitive maps allows you to reflect the main factors and possible reciprocal flanking between them, and is the basis for building more detailed computer scripts to develop the situation. Thus, the use of cognitive maps in the implementation of the scenario approach can significantly increase the effectiveness of analytical activities. The approach considered allows us to structure the problem, identify the most significant concepts (factors), take into account the connections between the concepts and the nature (strength) of these connections, and also choose the best combination of methods and thereby increase the validity of decisions.

The research results, in particular, information models, data analysis and visualization algorithms, are used in scenario generation tools in several security and defense decision-making support systems.

Acknowledgement. This research was supported by CyRADARS project (SPS G5286 “Cyber Rapid Analysis for Defense Awareness of Real-time Situation”) in the frame of the NATO Science for Peace and Security program.

References

1. Axelrod, R.: Structure of Decision: The Cognitive Maps of Political Elites. Princeton University Press, Princeton (2015)
2. Buneman, P.: Semistructured data. In: Proceedings of the Sixteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems. Mathematical Optimization and Economic Theory, vol. 7, pp. 117–121 (1997)
3. Calvanese, D.: Ontology-Based Data Access: From Theory to Practice, Université Clermont Auvergne on Vimeo (2012). <http://www.inf.unibz.it/~calvanese/presentations/BDA-2012-obda-calvanese.pdf>. Accessed 26 Oct 2012
4. Kayser, V., Shala, E.: Generating futures from text: scenario development using text mining. In: Daim, T., Chiavetta, D., Porter, A., Saritas, O. (eds.) Anticipating Future Innovation Pathways Through Large Data Analysis. Innovation, Technology, and Knowledge Management, pp. 229–245. Springer, Cham (2014)
5. Kok, K.: The potential of fuzzy cognitive maps for semiquantitative scenario development, with an example from Brazil. *Glob. Environ. Change* **2009**(19), 122–133 (2009)
6. Kontchakov, R., Rodríguez-Muro, M., Zakharyashev, M.: Ontology-based data access with databases: a short course. In: Reasoning Web International Summer School. Reasoning Web 2013: Reasoning Web. Semantic Technologies for Intelligent Data. Lecture Notes in Computer Science, vol 8067, pp. 194–229. Springer, Heidelberg (2013)

7. Liu, X., Yi, H., Ni, Z.: Application of ant colony optimization algorithm in process planning optimization. *J. Intell. Manuf.* **24**(1), 1–13 (2013)
8. López-Ibáñez, M., Stützle, T., Dorigo, M.: Ant colony optimization: a component-wise overview. *Handbook of Heuristics* 371–407 (2018)
9. Dorigo, M., Di Caro, G.: The ant colony optimization meta-heuristic. In: Corne, D., Dorigo, M., Glover, F. (eds.) *New Ideas in Optimization*, pp. 11–32. McGraw-Hill (1999)
10. Dodonov, O.G., Lande, D.V., Boychenko, A.V.: Scenario approach to the study of the dynamics of information flows on the Internet. In: *Proceedings of Open Semantic Technologies for Intelligent Systems (OSTIS-2015): Materials V International Scientific and Technical Conference*, pp. 225–230 (2015)
11. Lande, D., Snarsky, A.: An approach to the creation of terminological ontologies. *Ontol. Des.* **2**(12), 83–91 (2014)
12. Senchenko, V., Boychenko, A.: Investigation of methods and technologies of integration of an ontological model with relational data. *Regist. Storage Process. Data* **20**(3), 91–101 (2017)
13. Grabisch, M., Rusinowska, A.: Determining models of influence. *Oper. Res. Decis.* **26**, 69–85 (2016)
14. Zaghdoud, M.: Contextual fuzzy cognitive map for intrusion response system. *Int. J. Comput. Inf. Technol.* **02**(03), 471–477 (2013)
15. Novogrudska, R.L., Globa, L.S., Koval, O.V., Senchenko, V.R.: Ontology for applications development. In: Ciza, T. (eds.) *Ontology in Information Science* (2018). <http://doi.org/10.5772/intechopen.74042>
16. OWL 2 Web Ontology Language Document Overview (Second Edition) W3C Recommendation 11 December 2012 (2012). <http://www.w3.org/TR/owl2-overview/>. Accessed 17 May 2020
17. Basaras, P., Katsaros, D., Tassioulas, L.: Detecting influential spreaders in complex, dynamic networks. *Computer* **46**(4), 24–29 (2013)
18. Protégé – Free, open-source ontology editor and framework for building intelligent systems (2020). <http://protege.stanford.edu>. Accessed 17 May 2020
19. Styltsvig, H.B.: *Ontology-based Information Retrieval*, Ph.D. thesis, Computer Science Section, Roskilde University, Denmark (2006)
20. Rehman, Z., Kifor, C.V.: An ontology to support semantic management of FMEA knowledge. *Int. J. Comput. Commun. Control* **11**(4), 507–521 (2016)



Performance of the Reverse Load Balancer Method in Cluster and Cloud Infrastructures

Oleksandr Khoshaba¹(✉) , Vitalii Lytvynov² , Viktor Grechaninov¹ ,
and Kostiantyn Zaverailo² 

¹ Institute of Mathematical Machines and Systems Problems of the Ukraine National, Academy of Science, 42 Academician Glushkov Avenue, Kyiv 03187, Ukraine
Oleksandr.Khoshaba@gmail.com

² Chernihiv National University of Technology, 95 Shevchenka Street,
14035 Chernihiv, Ukraine

Abstract. The article proposes a new load balancing method based on the reverse direction of data flows that provide traffic redirection modules. This is due to the use of the synchronization mechanism of the internal modules with the load balancer pool. Using this method allows replacing traffic redistribution algorithms with a mechanism for synchronizing load balancer modules. Also, the use of single-threaded modules allows eliminating queues to nodes, which ensures more efficient consumption of remote resources. The article describes the problems of existing load balancing methods in cluster and cloud infrastructures and the advantages of the proposed method. Based on the general principles of operation of existing load balancing algorithms, a problem is formulated and a mathematical model is proposed that describes the developed method. The mathematical model allows you to set the loaded and unloaded state of the load balancer and the saturation point using certain conditions and implications. Also, the performance of the benchmark and the load balancer module was evaluated using the proposed method. In conclusion, the most relevant areas for further research and conclusions are presented.

Keywords: Load balancing method · Benchmark and load balancer performance · Load balancer states

1 Introduction. Existing Load Balancing Techniques for Cluster and Cloud Infrastructures

For two decades, there have been a number of load balancing methods in cluster and cloud infrastructures. These methods mainly concerned the development of adaptive algorithms. Also, such methods were aimed at the efficient use of information resources, increasing the bandwidth of a cluster or cloud infrastructure,

reducing the response time and reducing the cost of computing operations. The papers [2–5, 12] are devoted to these issues, where the fundamental principles of load distribution on information resources, the creation of a mathematical apparatus, models and methods for distributing data flows are investigated.

Usually, to achieve the goal of efficient load balancing in cluster and cloud infrastructures, algorithms were developed to search for optimal solutions for distributing traffic between the nodes on which the applications are located. Mostly, such algorithms should provide a proportional load of information resources.

Therefore, such algorithms were based on the OSI standard network model and were accordingly divided into content-dependent (seventh level of the OSI model) and content-independent (fourth level of the OSI model) groups [9, 11]. After that, all the algorithms within these groups were divided [1, 6–8, 10] to ensure existing load balancing methods in cluster and cloud infrastructures.

2 Problem Formulation

Analysis of existing methods and algorithms for load balancing in cluster and cloud infrastructures allows us to conclude that their work is not efficient enough and has drawbacks. First of all, a number of important factors are not taken into account. These include:

- the overall performance of all nodes located in cluster and cloud infrastructures is not the same;
- the performance of all nodes that are in a cluster or cloud infrastructure is not constant, i.e. the performance of all nodes varies over time;
- in all nodes in a cluster or cloud infrastructure, the amount of free RAM is dynamically changing, which affects application performance;
- due to the dynamic change in performance on nodes, the cost of computing and resource consumption dynamically changes, which must be analyzed when implementing load balancing;
- a number of other factors.

There is also a problem, which is the appearance of queues for applications in the presence of a large and uneven redistribution of the flow of requests. Then situations may arise when one node has processed traffic and is idle, while the other is still executing a request queue. In this case, there is no sufficiently effective mechanism for redistributing queues between nodes in cluster and cloud infrastructures.

The disparity in the processing of requests by different nodes, specialists try to solve by using remote agents. However, such agents also consume resources, contribute to lower node performance in cluster and cloud infrastructures, and generate new traffic from the node to the load balancer.

Another problem is the correct determination of the upper load (T_{low}) and lower overload (T_{high}) of the node when using the LARD algorithm on load balancers. The complexity of such studies lies in the correct determination of the input parameters of the LARD algorithm for various nodes that do not consume resources equally and have large changes in performance over short

periods of time. In this case, additional studies of node performance metrics in cluster and cloud infrastructures are needed. Only on the basis of such studies, decisions are made on the effective direction of requests for a short period of time.

3 Load Balancing Algorithms for Cluster and Cloud Infrastructures

The considered groups of load balancing algorithms in the previous section have both advantages and disadvantages in the distribution of request flows to the nodes on which the applications are located. Let’s consider some of them.

One of the most common load balancing algorithms is the Round Robin (RR) static algorithm. He belongs to the fourth level group of the OSI model. Its functions include sequentially sending requests to each application of the remote nodes (Fig. 1).

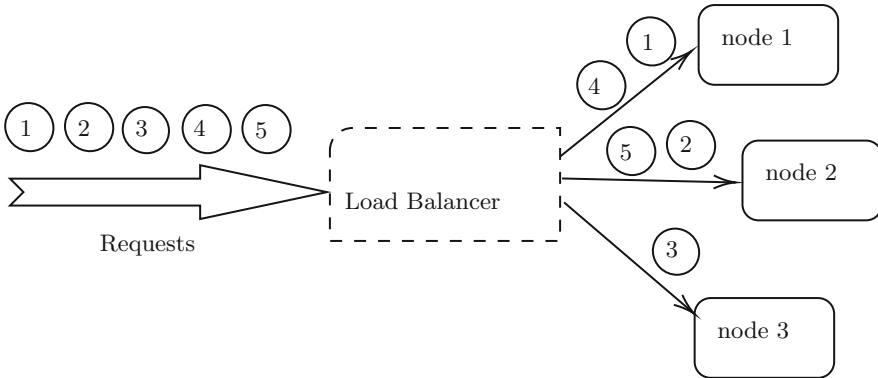


Fig. 1. Distribution of directions of incoming requests to node applications during the operation of the RR algorithm

This example (Fig. 1) shows the process of redistributing requests between nodes according to the strict order of their placement in a cluster or cloud infrastructure. In this case, the total number of requests is distributed equally among all nodes of the cluster or cloud infrastructure, regardless of their congestion. Often, this distribution of requests leads to queues on some nodes and downtime on others.

A more sophisticated load balancing algorithm is the Weighted Round Robin algorithm (WRR). It is the algorithm for working at the fourth level of the OSI model. When this algorithm works, the load balancer after a certain time performs periodic polling of nodes for their load. Using the information received, the balancer sends a certain number of requests, taking into account the workload of the nodes.

The above methods based on resource requests belong to the first group of classification of load balancing algorithms (i.e., the fourth level of the OSI model). The most common algorithm of the second group (seventh level of the OSI model) is Locality Aware Request Distribution (LARD).

The features in the operation of this algorithm include its ability to determine the characteristics of traffic and send a request to a less loaded node (Fig. 2). However, for this it is necessary to determine for the algorithm two input parameters for decision making. These parameters include traffic distribution: along the upper load (T_{low}) and lower overload (T_{high}) boundaries.

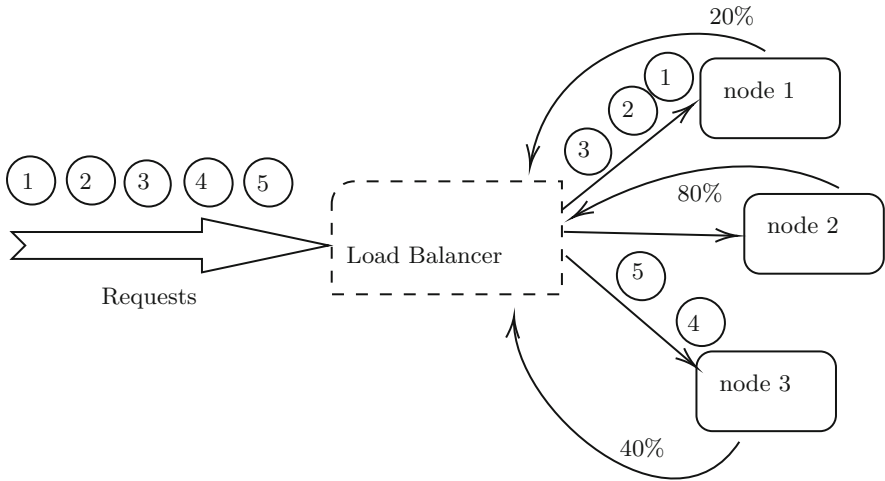


Fig. 2. Distribution of directions of incoming requests to node applications during the operation of the LARD algorithm

This example (Fig. 2) shows the process of redistributing requests to the less loaded first node with respect to the first group. The rest of the traffic will be redirected to the third node. Requests will not be sent to a more loaded second node, which currently has 80% of the load.

4 Proposed Load Balancing Method in Cluster and Cloud Infrastructures

The paper proposes another method of balancing load effects, which is based on some principles of data processing in Service Oriented Architectures (SOA). This method improves the efficiency of the load balancer by replacing the load balancing algorithms with the synchronization mechanism of the internal modules.

Traditional methods for redistributing requests in cluster and cloud infrastructures are based on the transfer of requests from load balancers to applications of remote nodes.

The proposed reverse method is based on the following principle of operation (Fig. 3). As usual, requests from users or benchmarks (when performing load testing) come to the load balancer. Such requests are placed in the order of arrival on the pool inside the load balancer. Modules that are also inside the load balancer are synchronized with the request pool. There can be several such modules and their number can be regulated by load testing engineers. When a request arrives in the balancer pool, one of several modules detects this event, after which it sends a request to the application of the remote node. Next, the request is processed on the remote site application. The result of processing the request is transmitted to the module that sent it. From the module, the response to the request goes to the user or benchmark. After that, the freed module is synchronized with the pool of the load balancer and the request processing cycle on the applications of the remote nodes is repeated.

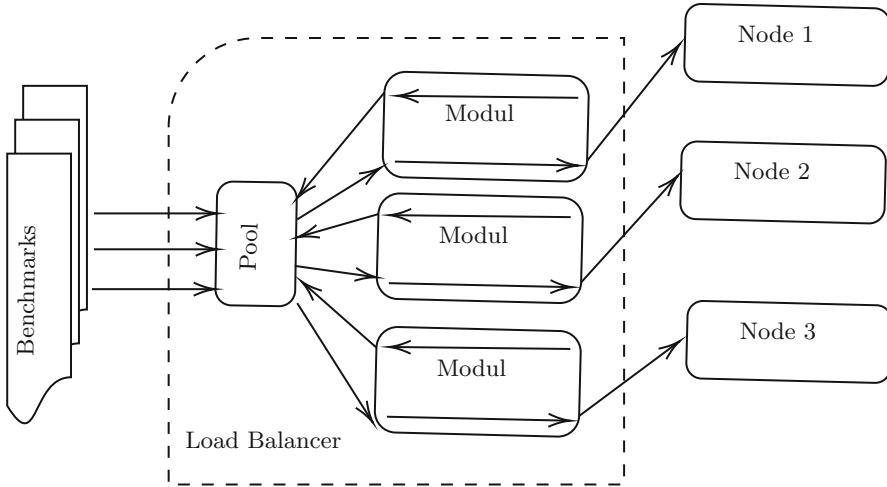


Fig. 3. Demonstration of the reverse load balancing method

This ensures a “reverse”, that is, the reverse direction of requests - from remote applications to the load balancer. Using internal modules that are synchronized with the pool and transmit the request, the effect of the application of remote nodes to the load balancer is created.

5 Deterministic Load Balancer Performance Model

Let us describe the deterministic model of the performance of the load balancer in the form of representing the analytical dependence of the data processing that comes from the benchmark.

In this model, the concepts are used: unloaded and loaded states of the load balancer, saturation points of the load balancer.

Definition 1. *The condition for the presence of intervals characterizes the time interval between the processing of requests at which the subsequent (or next) request is sent after receiving a response to the previous one. Otherwise, when the balancer receives a response to the previous request from the node before sending the next one, there is no interval.*

We define the implication of the condition for the presence or absence of an interval as:

$$In := \{\text{interval exists}\} \text{ and } \overline{In} := \{\text{no interval}\}$$

Definition 2. *The condition of synchronism of flows characterizes the sequence of requests for which there are single-threaded and multi-threaded connections. It is believed that single-threaded connections use the standby mode for processing application requests on nodes of a cluster or cloud infrastructure.*

We define the implication of the condition for synchronization of flows as: $Sn := \{\text{single-threaded connections}\}$ and $As := \{\text{multi-threaded connections}\}$

To determine the saturation point of the load balancer, we will use the following conditions: the presence of intervals, synchronization of flows, and the presence of a queue.

Definition 3. *An unloaded state ($S0$) is characterized by the presence of an interval between sending and receiving requests (In), synchronous (Sn) or asynchronous (As) connections, and the absence of a queue (Qu), where:*

$$q \in S0(In \wedge Sn \wedge \overline{Qu}). \quad (1)$$

Definition 4. *The load balancer saturation point (Sp) is characterized by a zero value of the interval between sending and receiving requests (In), synchronous (Sn) or asynchronous (As) connections and the absence of a queue (Qu), where:*

$$q \in Sp(\overline{In} = 0) \wedge Sn \wedge \overline{Qu}. \quad (2)$$

Definition 5. *The loaded state ($S1$) is characterized by the presence or absence of an interval between sending and receiving requests (In), synchronous (Sn) or asynchronous (As) connections, and the presence of a queue (Qu), where:*

$$q \in S1(\overline{In} < 0) \wedge As \wedge Qu. \quad (3)$$

Proof. Suppose, from the benchmark to the load balancer received N requests (R_i). The total number of requests will be equal to:

$$N = \sum_{i=0}^{n-1} R_i. \quad (4)$$

Let time t be determined for each of the requests that are sent as follows. For the first request (R_0), the time t_0 corresponds. The initial sending time for all requests will be equal to $R_i = t_i$. The interval between sending requests (In) is defined as $t_{i+1} - t_i$.

Suppose that the system determines the time for processing and sending requests (Response Time, RT) to remote nodes of a cluster or cloud infrastructure. The total number (TRT) of RT_i measurements will be equal to:

$$TRT = \sum_{i=0}^{n-1} RT_i. \quad (5)$$

Then, from the following ratio, we determine the state of the performance of the load balancer:

$$State = \begin{cases} unloaded_state(S0) & \text{if } RT_i \leq t_{i+1} - t_i; \\ loaded_state(S1) & \text{if } RT_i > t_{i+1} - t_i. \end{cases} \quad (6)$$

Also, we determine the saturation point if we have the condition for (Sp): if $RT_i = t_{i+1} - t_i$.

Based on formulas 1–3, 6, we determine the necessary and sufficient conditions for the forms of implications of the states $S0$, $S1$ and the point of saturation (Sp). For state $S0$ we define implications:

$$\boxed{S0 \Leftrightarrow In}; \quad S0 \Rightarrow Sn; \quad S0 \Rightarrow \overline{Qu};$$

For state $S1$ we define implications:

$$\boxed{S1 \Leftrightarrow \overline{In} < 0}; \quad S1 \Rightarrow As; \quad \boxed{S1 \Leftrightarrow Qu};$$

For saturation point (Sp) we define implications:

$$\boxed{Sp \Leftrightarrow (\overline{In} = 0)}; \quad Sp \Rightarrow Sn; \quad Sp \Leftrightarrow \overline{Qu};$$

Notes and Comments. Equivalent implicators are shown in a square.

Thus, the description of the deterministic model of the performance of the load balancer in the form of a representation of the analytic dependence on the basis of implicators made it possible to obtain four simultaneous necessary and sufficient conditions (equivalent implicators) for the forms of the consequences of states $S0$, $S1$ and the saturation point (Sp).

6 Experimental Investigations and Their Analysis

The study of the performance of the load balancer in the first state is to determine the saturation process. The load balancer saturation process was performed using two conditions:

- sequential generation of requests to the load balancer;
- the absence of a time interval between the generation of each individual request.

The basis of the performance of the load balancer is single-threaded and multi-threaded connections. In Fig. 4 shows the performance of single-threaded connections between the benchmark and the load balancer module.

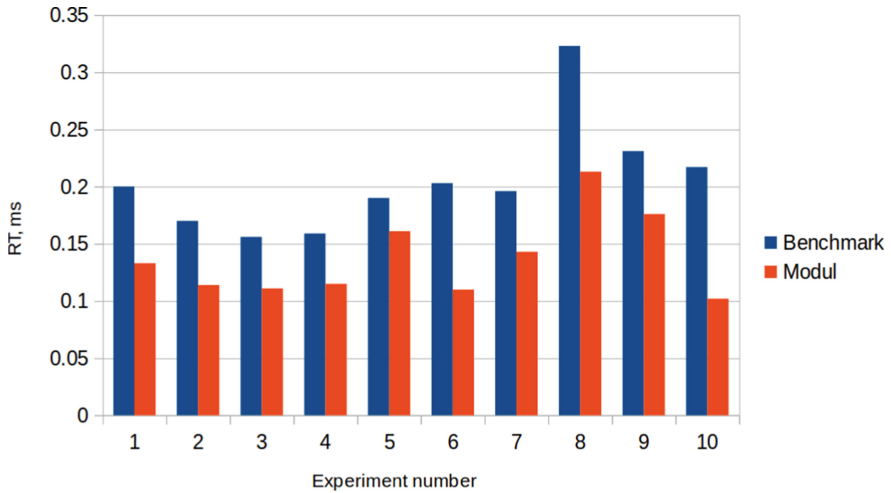


Fig. 4. Comparison of performance for benchmark and one load balancer module

The results of the study allow us to conclude that the load balancer module has a higher performance than the benchmark.

The modules used in the load balancer are single-threaded. One of their drawbacks is not high enough efficiency, which is based on high response times (RT) as a result of idle (Id). These idle result from delays in transporting requests:

$$RT = ST + Id, \quad (7)$$

where, ST is the service time of one request at the remote node.

As a result of this, the remote host waits for subsequent requests. To increase the efficiency of the load balancer, it is possible to use several modules at the same time to service one node.

In Fig. 6 shows the changes in the values of the performance indicator (Response Time, RT) depending on the use of modules in an amount of one to ten per remote node.

The dynamics of the values of changes in performance indicators of the load balancer modules is shown in Fig. 6.

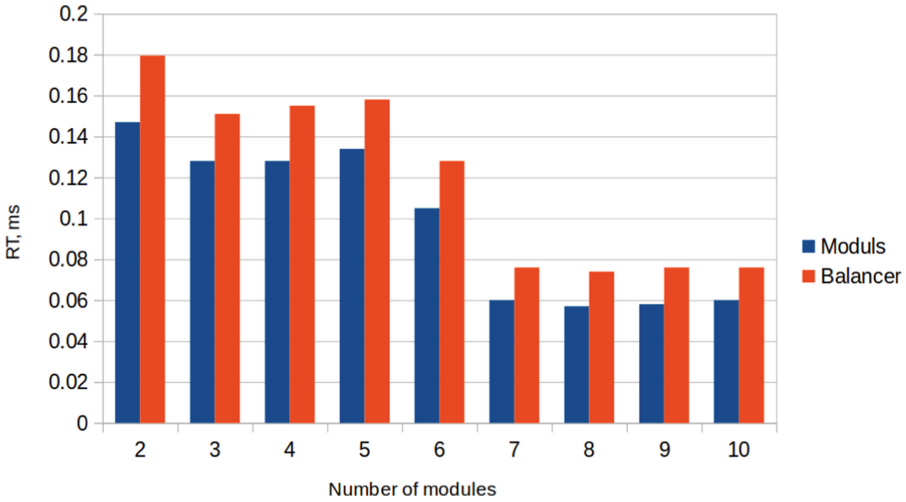


Fig. 5. Comparison of performance for benchmark and several load balancer modules

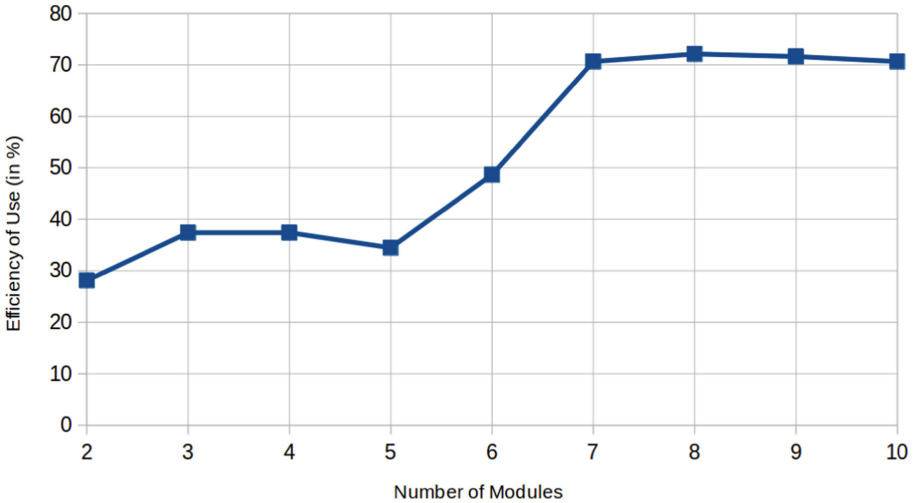


Fig. 6. The dynamics of the values of changes in performance indicators depending on the various number of load balancer modules

The Fig. 6 shows that the efficiency of the modules increases with their number. However, in this case, starting from the seventh module, performance indicators increase slightly. The presence of a different software and hardware environment or load changes can get a different number of modules (Fig. 5).

Also, the study measured the performance of testing modules and loaded balancing in the range from one to ten. Performance parameters, such as response time and response speed, were measured.

7 Most Relevant Areas for Further Research

For load balancer performance:

- study of the performance of the formation and processing of requests by the load balancer;

For load balancer model:

- the study of points of degradation and failure with an increase in the speed of receipt of requests to the load balancer;
- formalization of conditions and the creation of criteria for degradation and failure of the load balancer.

8 Conclusions

The article proposes a new method of work of the load balancer in cluster and cloud infrastructures. This method improves the efficiency of the load balancer by replacing the load balancing algorithms with the synchronization mechanism of the internal modules. Improving the efficiency of the load balancer is the lack of need:

- use of load balancing algorithms;
- creating queues by applications of remote nodes;
- the use of remote agents and the transfer of the results of congestion nodes of a cluster or cloud infrastructure;
- determination of load states of remote nodes.

Also, the article conducted a study of the performance of the load balancer.





References

1. Bo, Z., Ji, G., Jieqing, A.: Cloud Loading Balance algorithm, pp. 5001–5004 (2010)
2. Cardellini, V., Casalicchio, E., Colajanni, M., Yu, Ph.S.: The state of the art in locally distributed web-server systems. *ACM Comput. Surv. (CSUR)* **34**, 263–311 (2002)
3. Casalicchio, E., Colajanni, M.: A client aware dispatching algorithm for web clusters providing multiple services. In: *Proceedings of the 10th International Conference on WWW*, pp. 535–544 (2001)

4. Hong, Y.S., No, J.H., Kim, S.Y.: DNS-based load-balancing in distributed web-server systems. In: Fourth IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems (WCCIA 2006), p. 4 (2006)
5. Kamra, A., Misra, V., Nahum, E.M., Yaksha: A self-tuning controller for managing the performance of 3-tiered web sites. In: 12th IEEE International Workshop on Quality of Service, IWQOS 2004, pp. 47–56 (2004)
6. Keyvanpour M.R., Mansourifar H., Bagherzade B.: A novel classification of load balancing algorithms in distributed systems. In: Thaug K. (ed.) Advanced Information Technology in Education. Advances in Intelligent and Soft Computing, vol. 126, pp. 313–320. Springer, Heidelberg (2012)
7. Mallikarjuna, B., Doddi, A.: The role of load balancing algorithms in next generation of cloud computing. *J. Adv. Res. Dyn. Control Syst.* **11**, 1715–1733 (2019)
8. Mcheick, H., Alardawy, Z., Lakiss, A.: Evaluation of load balance algorithms. In: Proceedings - 2011 9th International Conference on Software Engineering Research, Management and Applications, SERA 2011, pp. 104–109 (2011)
9. Palpandi, B., Palpandi B., Geetharamani, G., Pandian, J.A.: Performance enhancement in OSI network model using fuzzy queue. *Ind. J. Appl. Res.* **4**, 377–386 (2011). <https://doi.org/10.15373/2249555X/APR2014/116>
10. Samunnisa, K., Sunil G., Kumar, V., Madhavi, K.: A circumscribed research of load balancing techniques in cloud computing. *Int. J. Innov. Technol. Exploring Eng. (IJITEE)* 8(6S4), 336–339 (2019). ISSN 2278-3075
11. Xiao-ling, F., Ying, X.: A load balance algorithm for hybrid P2P network model. In: Proceedings - ISECS International Colloquium on Computing, Communication, Control, and Management, CCCM 2008, vol. 1, pp. 236–239 (2008). <https://doi.org/10.1109/CCCM.2008.325>
12. Schroeder, T., Goddard, S., Ramamurthy, B.: Scalable web server clustering technologies. *IEEE Netw.* **14**, 38–45 (2000)



Building of Directed Weighted Networks of Terms for Decision-Making Support During Information Operations Recognition

Dmytro Lande^{1,2} , Oleh Dmytrenko¹ ,
Oleh Andriichuk^{1,2} , Vitaliy Tsyganok^{1,2} ,
and Yaroslava Porplenko¹

¹ Institute for Information Recording, National Academy of Sciences of Ukraine,
Kiev, Ukraine

dwlande@gmail.com, dmytrenko.o@gmail.com,
{andriichuk, tsyganok}@ipri.kiev.ua, daliss@ukr.net

² National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic
Institute”, Kiev, Ukraine

Abstract. The study is devoted to automated processing of documents of a subject domain for formation of text corpus, construction of networks of terms and using of these networks for building of knowledge bases of decision support systems during information operations recognition. In this work, for building the undirected networks of terms the Horizontal Visibility Graph algorithm is used.

A new approach for determining the directions and the weights of links in the network of terms that correspond to certain concepts of the considered subject domain is proposed. By its application, an ontological model of the subject domain related to the information operations recognition was built.

A new approach for building the knowledge bases of decision support systems during the information operations recognition using the directed weighted networks of terms of subject domain is proposed. Using it for building of knowledge bases allows saving time and financial resources by reducing the use of expert information and make it possible to detect gaps in the knowledge bases of decision support systems.

Keywords: Information space · Text corpus · Directed weighted networks of terms · Horizontal visibility graph · Decision support system · Knowledge base · Information operation recognition

1 Introduction

The modern information space is characterized by the fast-moving development of dynamic information arrays and flows distributed in webspace.

In the information space such as the Internet, the main results of human communication activity are concentrated.

Today the totality of information, which is accessible online rapidly grows. No doubt, the impact of this information on people is also constantly growing. So, there is a problem of informational influence on people’s minds and their decisions. Usually,

this influence is an intentional (premeditated) and better known as an «information Operation».

Information operations (IO), also known as influence operations, are used in Information Warfare. IO are actions taken to affect adversary information and information systems while defending one's information and information systems. This type of modern warfare is very popular and used by many countries all over the world.

So, in the conditions of the current rapid development of information technologies and their comprehensive and deep penetration into all spheres of human life the task to recognize any manifestations of information operations is an urgent issue.

The consequence of IO is the formation of an information environment that somehow influences society, social groups and individuals within it [1]. IO make it possible to form a certain opinion and attitude in the target audience to some issue, topic, object [2, 3]. It can even provoke some negative social tension [4]. Many experimental studies confirm the presence of the effects of misinformation and gossip, and, as a consequence, the formation of belief in it inside a society [5, 6]. During information operations recognition, information contained in open sources can be used by content monitoring systems that allow building the appropriate directed weighted networks of terms based on this information [7].

On the other hand, IO belong to semi-structured subject domains [8], so it is advisable to use decision support systems (DSS) to recognize IO [9, 10]. In this case, the DSS tools build appropriate knowledge bases (KB) that describe the specifics of the object of the IO and the information environment in which it is located. For the building of the KB along with directed weighted networks of terms expert knowledge and objective information are used. The use of experts requires considerable time and financial costs, as well as the issues solution related to errors (due to human psychophysiological characteristics) and subjectivity of expert assessments. Thus, the reduction of the number of address to experts during the building of the DSS knowledge base is an urgent problem now.

2 Formation of the Corpus of Text Documents

For forming the text corpus, a tool for professionals in the field of information analysis and information warfare – the content monitoring system «Info stream» [11] was used. As a result, 135 publications that thematically related to Brexit were downloaded.

During the formation of the text corpus, it was an open question to define enough texts for complete, reliable and representative coverage of the subject area related to the subjects of the research object. In general, for this purpose, when building models of natural language within the framework of computational linguistics, the certain patterns based on the following effects are used: the appearance of new unique words with appearance new texts in the corpus [12], constancy of the ratio of the number of profile thematic publications to the number of partially profile and a total number of publications [13], constancy of the ratio of word frequency to its rank [14, 15]. The disadvantage of using these regularities is that a sufficient amount of corpus documents will be achieved with a sufficiently large number of processed texts. It, in turn, leads to significant computational costs informing and calculating of this corpus and bulding an

appropriate directed weighted network of terms. In addition, a sufficient number of texts may not be achieved due to the limited presentation of some targeted themes in the media. Since the directed weighted network of terms will be used for the further processing within the framework of the building of the DSS KB during the IO recognition, then an advisable condition of a stop is a stabilization of the values of the weights of the corresponding terms during expanding the corpus of the target thematic texts.

Table 1 presents the dynamics of variation of the weight of terms when building the text corpus.

Table 1. Top 28 key words for the text corpus that thematically related to Brexit for different sizes of the corpus.

60 documents		80 documents		105 documents		121 documents		135 documents	
Label	weight	Label	weight	Label	weight	Label	weight	Label	weight
eu	0.03027	eu	0.043649	eu	0.041886	eu	0,042131	eu	0.041618
uk	0.016056	uk	0.037739	uk	0.036625	uk	0,036394	uk	0.035999
brexit	0.011355	leav	0.013201	leav	0.013856	leav	0,014264	leav	0.013976
leav	0.009713	brexit	0.012087	brexit	0.012665	brexit	0,012906	brexit	0.012657
britain	0.008863	vote	0.007993	vote	0.008576	vote	0,009134	vote	0.009089
european	0.008523	countri	0.006734	trade	0.006392	remain	0,006149	trade	0.006483
trade	0.006909	trade	0.006492	countri	0.006273	countri	0,006078	remain	0.005978
vote	0.006881	econom	0.006007	remain	0.006114	trade	0,006078	countri	0.005864
countri	0.006456	remain	0.005983	econom	0.006035	econom	0,005720	econom	0.005489
remain	0.005833	union	0.005159	referendum	0.0053	referendum	0,005684	referendum	0.005473
british	0.00555	referendum	0.005062	peopl	0.004963	campaign	0,005059	campaign	0.005115
econom	0.005324	peopl	0.004723	union	0.004764	peopl	0,004951	peopl	0.004952
europ	0.005324	polit	0.004699	campaign	0.004625	union	0,004808	union	0.004675
referendum	0.005154	market	0.00453	polit	0.004467	polit	0,004522	market	0.004528
union	0.004955	campaign	0.004384	market	0.004367	year	0,004379	year	0.004365
market	0.004701	year	0.004142	year	0.004288	market	0,004290	polit	0.004317
peopl	0.004361	immigr	0.004045	support	0.003831	support	0,003843	govern	0.003893
support	0.004191	support	0.003997	govern	0.003811	govern	0,003718	support	0.003698
member	0.004163	govern	0.003948	immigr	0.003533	immigr	0,003414	member	0.003502
polit	0.004134	member	0.003924	member	0.003474	member	0,003361	immigr	0.003258
immigr	0.004106	nation	0.003754	parti	0.003395	parti	0,003325	nation	0.003225
year	0.004049	parti	0.003488	nation	0.003275	nation	0,003182	parti	0.00303
campaign	0.004021	membership	0.003222	membership	0.003037	membership	0,003057	membership	0.003013
govern	0.003766	free	0.003004	time	0.002918	time	0,002914	time	0.002883

3 Text Corpus Processing

In this work, the tokenization, lemmatization, stop-words removal, stemming process and terms weighting are made.

The tokenization allows separating the text into the set of tokens. The lemmatization process returns the lemmas – the dictionary form of a word [16].

After the pre-processing of the textual documents in this work it is proposed to remove stop-words that informationally unimportant ones. For example, such words as ‘the, a, as, are, be’ etc. are commonly used in the English language and have no

semantic strength. In this work the stop dictionaries [17, 18] were used. Also, we used the stop dictionary that was formed by experts within the considered subject domain.

After the stages described above, the process of stemming was made. Stemming makes it possible to combine the words with a common root into a single word. The Porter's algorithm [19, 20], that was used in this work, is the most common and empirically very effective algorithm for stemming English.

The next step of the text corpus processing is an extraction of key terms. For this goal the terms weighting is made. GTF (Global Term Frequency) [21] – a modification of classic statistical weight indicator TF-IDF [22] – is used as a weight of terms to reflect the term to number.

4 Building a Directed Network of Terms

To build the undirected network of terms in this work a common Visibility Graph algorithm [23] that maps a time series into a network – the Horizontal Visibility Graph (HVG) algorithm is used. The process of building the Horizontal Visibility Graph consists of two steps and presented in detail in the work [24]. In the first step, the sequences of nodes are marked on a horizontal axis (x -axis) in the order it appears in the text. On a vertical axis (y -axis) the weight indicator GTF is marked. In the second step a classical Horizontal Visibility Graph is built. Two nodes t_i and t_j are connected in the HVG if and only if these nodes are in the «direct visibility». «Direct visibility» between nodes means that they can be connected with the horizontal line which does not intersect any vertical line in the resulting plot.

Thus, the HVG algorithm allows building an undirected network of terms in case, when the numerical values are assigned to separated words or phrases of a thematic text corpus.

In the work [25] the approach for determining the directions of links in undirected networks of terms that correspond to certain concepts of the considered subject domain is proposed. We will use this approach to build a directed network of terms. It is supposed that a causal link exists in the direction from the node t_i to the node t_j in the undirected network if, within the sentence, the term to which the node t_i corresponds precedes the term to which the node t_j corresponds.

As a result of applying a described above approaches, the directed network of terms of a thematic text corpus is built.

5 Determining the Weights of Links

This work proposes a new approach for determining the weights of links between nodes in the directed network of terms based on a thematic text corpus. Using the method described above the directed network of terms is built.

At the graph level, the general principle is described as follows: the nodes corresponding to the same terms of the directed network built at the previous stage are merged. Since any graph is defined by an adjacency matrix, the task of determining the weighted values of the links is reduced to the concatenation of columns and

corresponding rows. In other words, it is a weighted compactification of the Horizontal Visibility graph [24].

More formally, the process of determining the weights of links in the network of terms is as follows. Let D be the directed network of terms that built according to the described above rule: $D := (V, E)$ where V is the set of nodes, E is the set of the ordered pairs of nodes from the set V that correspond to the causal links between the nodes. Let A is a square matrix of size n in which the value of the element a_{ij} is 1 if there is an edge (arc) from vertex i to vertex j else a_{ij} is 0. Let $T = \{t_1, \dots, t_m\}$ is the set of nodes that corresponds to the same terms of text (where $1 \leq m \leq n$). It is obvious that each node t_k ($1 \leq k \leq m$) in the set T has a column a_{ik} and row a_{kj} in the matrix A . Therefore, the corresponding column (row) elements of all same nodes are summarized and written to a new column (row, respectively) – w_{ik} (w_{kj} respectively), and a new matrix W is formed. As a result of the concatenation process described above, in the resulting matrix W the value of the element w_{ij} is equal to the number of edges from vertex i to vertex j .

Also, the above-mentioned process of the concatenation of columns (rows) uses the so-called hash table of similar or synonymous terms that correspond to the same concepts. This table is formed by experts within the considered subject domain. This table makes it possible to further merge nodes that correspond to the same terms in the text.

The resulting matrix W defines an oriented weighted graph formed of nodes that correspond to the unique terms within the considered text. The weight of the edge connecting the vertex i to the vertex j is defined by the number of appearances of the term to which the node t_i corresponds before the term to which the node t_j corresponds (the number of occurrences of an element t_i of the time series before the element t_j) in the considered text.

6 Building of Knowledge Bases of Decision Support Systems During Information Operations Recognition

The building of the KB of DSS is carried out within the framework of the method of hierarchical decomposition and complex target-oriented dynamic evaluation of alternatives [26]. At first, decomposition of the main goal into elementary sub-goals (factors, criteria) that are its constituents and directly influenced by it is performed. Then there is a decomposition of each component into respective sub-goals and so on. The decomposition process of goals stops when direct actions (projects) are received as sub-goals within the next decomposition [10]. In addition to expert knowledge, decomposition can also use objective information (for example, the values of certain target indicators). After the decomposition process of goals is completed, each decomposition calculates the partial impact coefficients (PIC) of the sub-goals on the respective goal. To do this, the expert estimation with the software tools of DSS is used. In the area of IO recognition, we have as projects specific themes of publications related to the object of IO. Thus, using DSS tools it is possible to calculate recommendations in the form of ranking of information impact of publications themes on the object of IO.

It is supposed that there is a sufficiently high quality directed weighted network of terms (of sufficient volume, representative, without errors, without redundancy and with a sufficient level of stability of weights) that is built using the content monitoring system of the IO subject domain. It is the initial network of terms, which will be used later.

The process of building the DSS KB using the directed weighted networks of terms during IO recognition is proposed:

1. The preliminary building of the DSS KB is carried out with the use the content monitoring and expert information [27, 28], a number of decompositions are defined, within which the appropriate PIC must be determined.
2. An analysis of the initial directed weighted network of terms in order to define the completeness of coverage of the subject domain is performed. Also, the pre-built DSS KB and the selected decomposition is taken into account. A situation might arise, for example, when not all decompositions are sufficiently covered by the network of terms. For such decompositions, other available approaches to determining the PIC [28, 29] or even consult experts are used.
3. For each “covered” decomposition, the necessary level of abstraction and stratification and corresponding terms of the initial network that correspond to each of the objects of the selected decomposition are determined.
4. A new network of terms is formed by merging certain nodes of the initial network according to each of the objects of the selected decomposition (goals and sub-goals).
5. The values of the impact of sub-goals on the target by the method of the optimal impact are determined [7]. The obtained values of the impacts are normalized and entered in the DSS KB as the appropriate PCI.
6. Next move on to step 3 until we go over all “covered” decompositions.

The advantages of the proposed approach are as follows: saving time and financial resources by reducing the use of expert information; the possibility of detecting gaps in the DSS KB during the analysis of the initial directed weighted network of terms; objectification of the PCI definition.

The disadvantages of the proposed approach are the complexity and, sometimes, the ambiguity of finding the conformity between some rather complex and broad goals and the terms of the network; the lack of the ability to apply the approach to other areas other than IO recognition.

7 Practical Example

Described above approaches are illustrated by the example of Brexit. Currently, Brexit is a topical issue that is widely researched by the scientific community [30]. Applying the stages of the computerized processing of text corpus described above, the key links

between the nodes were extracted. The Python programming language and its module NLTK (Natural Language Toolkit open-source library) [31] are used to build the software realization of the proposed and considered approaches and methods.

Table 2 shows the list of the most influential and significant links between the corresponding nodes in the network of terms that, in turn, correspond to certain concepts of the considered subject domain.

Table 2. Top 17 significant links for the corpora «Brexit».

№	Source	Target	Weight
1	uk	eu	889
2	leav	eu	404
3	eu	uk	319
4	brexit	uk	213
5	vote	leav	211
6	eu	union	174
7	uk	leav	145
8	remain	eu	120
9	brexit	eu	102
10	leav	uk	95
11	trade	eu	88
12	eu	leav	77
13	eu	countri	73
14	eu	membership	72
15	uk	vote	69
16	leav	campaign	66

Using the software for modeling and visualization of graphs – Gephi [32, 33], the directed weighted network of terms was built and visualized (Fig. 1). Figure 1 depicts the key words of the considered subject domain.

Also, using the Gephi software tools, the following parameters of the built network were obtained: the number of nodes is 28; the number of links is 640; the network density is 0.847; the number of connected components is 1; the average path length is 1,153; the average clustering coefficient is 0.856.

After analyzing the obtained results, it was established that the most significant links between the corresponding nodes in the network of terms built for the subject domain Brexit are: «uk → eu», «leav → eu», «eu → uk», «brexit → uk» and «vote → leav» (Fig. 2).

8 Conclusion

In this work, the approach for determining the weights of links in the network of terms was proposed. The directed weighted network of terms was built for the thematically targeted subject domain.

The approach was suggested for building of knowledge bases of decision support systems during information operations recognition which allows to provide the decomposition of the topics of the information operations and assess rating of the effectiveness of these topics.

Acknowledgment. This study is funded by the NATO SPS Project CyRADARS (Cyber Rapid Analysis for Defense Awareness of Real-time Situation), Project SPS G5286.

References




1. Chertov, O., Rudnyk, T., Palchenko, O.: Search of phony accounts on Facebook: Ukrainian case. In: International Conference on Military Communications and Information Systems ICMCIS, pp. 22–23 (2018)
2. U.S. Department of Defense.: Information operations roadmap. Washington, D.C. GPO (2003)
3. Military Information Support Operations. Joint Publication 3-13.2. 07 January 2010 Incorporating Change 1 20 December 2011. p. 10. <http://fas.org/irp/doddir/dod/jp3-13-2.pdf>
4. Shchoholiev, M., Tretynyk, V.: The System of Operative Determination of the Level of Tension in Society Based on Data from Social Networks. *Inf. Secur. Int. J.* **43**(3), 375–382 (2019). <https://doi.org/10.11610/isij.4328>
5. Lewandowsky, S., Ecker, U.K.H., Seifert, C.M., Schwarz, N., Cook, J.: Misinformation and its correction continued influence and successful debiasing. *Psychol. Sci. Publ. Interest* **13**(3), 106–131 (2012)
6. Berinsky, A.J.: Rumors and health care reform: experiments in political misinformation. *Br. J. Polit. Sci.* **47**(2), 241–262 (2017)
7. Dmytrenko, O., Lande, D., Andriichuk, O.: Method for searching of an optimal scenario of impact in cognitive maps during information operations recognition. In: Mathematical Modeling and Simulation of Systems (Selected Papers of 14th International Scientific-Practical Conference, MODS, Advances in Intelligent Systems and Computing Series, 24–26 June 2019, Chernihiv, Ukraine), vol. 1019, pp. 182–193 (2019). https://doi.org/10.1007/978-3-030-25741-5_19. ISSN 978-3-030-25740-8
8. Kadenko, S.V.: Prospects and potential of expert decision-making support techniques implementation in information security area. In: CEUR Workshop Proceedings (ceur-ws.org). Selected Papers of the XVI International Scientific and Practical Conference “Information Technologies and Security” (ITS 2016), Kyiv, Ukraine, 1 December 2016, vol. 1813, pp. 8–14 (2016). <http://ceur-ws.org/Vol-1813/paper2.pdf>. Vol-1813 urn:nbn:de:0074-1813-0
9. Lee, D.T.: Expert decision-support systems for decision-making. *J. Inf. Technol.* **3**(2), 85–94 (1988)

10. Tsyganok, V.V.: Decision-making support during planning of measures for counteracting informational operations. In: CEUR Workshop Proceedings (ceur-ws.org). Selected Papers of the XVI International Scientific and Practical Conference “Information Technologies and Security” (ITS 2016) Kyiv, Ukraine, 1 December 2016, vol. 1813, pp. 24–31 (2016). <http://ceur-ws.org/Vol-1813/paper4.pdf>. Vol-1813 urn:nbn:de:0074-1813-0
11. InfoStream. <http://infostream.ua/>. Accessed 10 Apr 2020
12. Heaps, H.S.: Information Retrieval — Computational and Theoretical Aspects. Academic Press, Cambridge (1978)
13. Bradford, S.C.: Sources of information on specific subjects. *Engineering*. **137**, 85–86 (1934). *An Illustrated Weekly Journal*, London
14. Manning, C.D., Schütze, H.: Foundations of Statistical Natural Language Processing. The MIT Press, Cambridge (1999)
15. Salton, G., Wong, A., Yang, C.A.: Vector space model for automatic indexing. *Commun. ACM* **18**(11), 613–620 (1975)
16. Korenius, T., Laurikkala, J., Järvelin, K., Juhola, M.: Stemming and lemmatization in the clustering of finnish text documents. In: Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management, pp. 625–633 (2004)
17. Google Code Archive: Stop-words. <https://code.google.com/archive/p/stop-words/downloads/>. Accessed 10 Apr 2020
18. Text Fixer: Common English Words List. <http://www.textfixer.com/tutorials/common-english-words.php>. Accessed 10 Apr 2020
19. Porter, M.F.: An algorithm for suffix stripping. *Program* **14**(3), 130–137 (1980). <https://doi.org/10.1108/eb046814>
20. Willett, P.: The Porter stemming algorithm: then and now. *Program* **40**(3), 219–223 (2006). <https://doi.org/10.1108/00330330610681295>
21. Lande, D.V., Dmytrenko, O.O., Snarskii A.A.: Transformation texts into complex network with applying visibility graphs algorithms. In: CEUR Workshop Proceedings (ceur-ws.org). Selected Papers of the XVIII International Scientific and Practical Conference on Information Technologies and Security (ITS 2018), vol. 2318, pp. 95–106. (2018). <http://ceur-ws.org/Vol-2318/paper8.pdf>. Vol-2318 urn:nbn:de:0074-2318-4
22. Salton, G., Buckley, C.: Term-weighting approaches in automatic text retrieval. *Inf. Process. Manag.* **24**(5), 513–523 (1988). [https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)
23. Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuno, J.C.: From time series to complex networks: the visibility graph. *Proc. Nat. Acad. Sci.* **105**(13), 4972–4975 (2008). <https://doi.org/10.1073/pnas.0709247105>
24. Lande, D.V., Snarskii, A.A., Yagunova, E.V., Pronoza, E.V.: The use of horizontal visibility graphs to identify the words that define the informational structure of a text. In: 2013 12th Mexican International Conference on Artificial Intelligence, pp. 209–215 (2013). <https://doi.org/10.1109/micai.2013.33>
25. Lande, D., Dmytrenko, O., Radziivska, O.: Determining the directions of links in undirected networks of terms. In: CEUR Workshop Proceedings (ceur-ws.org). Selected Papers of the XIX International Scientific and Practical Conference “Information Technologies and Security” (ITS 2019), vol. 2577, pp. 132–145. (2019). <http://ceur-ws.org/Vol-2577/paper11.pdf>. Vol-2577 urn:nbn:de:0074-2318-4. ISSN 1613-0073
26. Totsenko, V.G.: One approach to the decision making support in R&D planning. part 2. the method of goal dynamic estimating of alternatives. *J. Autom. Inf. Sci.* **33**(4), 82–90 (2001)

27. Andriichuk, O.V., Kachanov, P.T.: A methodology for application of expert data-based decision support tools while identifying informational operations. In: CEUR Workshop Proceedings (ceur-ws.org). Selected Papers of the XVI International Scientific and Practical Conference “Information Technologies and Security” (ITS 2016) Kyiv, Ukraine, 1 December 2016. vol. 1813, pp. 40–47 (2016). <http://ceur-ws.org/Vol-1813/paper6.pdf>. Vol-1813urn:nbn:de:0074-1813-0
28. Andriichuk, O., Lande, D., Hraivoronska, A.: Usage of decision support systems in information operations recognition. In: Proceedings of the XVIII International Conference on Data Science and Intelligent Analysis of Information, Recent Developments in Data Science and Intelligent Analysis of Information, Advances in Intelligent Systems and Computing, 4–7 June 2018, Kyiv, Ukraine. vol. 836, pp. 227–237. Springer Nature (2018). ISSN 2194-5357
29. Lande, D.V., Andriichuk, O.V., Hraivoronska, A.M., Guliakina, N.A.: Application of decision-making support, nonlinear dynamics, and computational linguistics methods during detection of information operations. In: CEUR Workshop Proceedings (ceur-ws.org). Selected Papers of the XVII International Scientific and Practical Conference on Information Technologies and Security (ITS 2017), vol. 2067, pp. 76–85 (2017). <http://ceur-ws.org/Vol-2067/paper11.pdf>. Vol-2067urn:nbn:de:0074-2067-8
30. Bachmann, V., James, D.: Sidaway Brexit geopolitics. *Geoforum* **77**, 47–50 (2016)
31. Bird, S., Klein, E., Loper, E.: *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O’Reilly Media Inc, Sebastopol (2009)
32. Cherven, K.: *Network Graph Analysis and Visualization with Gephi*. Packt Publishing Ltd (2013)
33. Gephi. <https://gephi.org>. Accessed. 10 Apr 2020



Personalization of E-Learning Process Using AI-Powered Chatbot Integration

John N. Davies¹ , Mariya Verovko² , Oleksandr Verovko²,
and Iryna Solomakha³ 

¹ CARDS Glyndŵr University, Wrexham, UK
johnndavies@gmail.com

² AgileVision sp. z o.o., Krakow, Poland

miya.tevkun@gmail.com, averovko@gmail.com

³ Chernihiv National University of Technology, Chernihiv, Ukraine
iveria60@gmail.com

Abstract. Personalized learning path is considered as one of the most effective way to archive the best outcomes from the eLearning process. The introduction of artificial intelligence (AI) and machine learning (ML) into the educational process is one of the most important trends of modern eLearning since it can provide flexibility of learning and can performing adaptation of the learning process based on the personal needs of each user. AI technologies are positioned as the best solution for the analysis of each learner's personal requirements and for generating a corresponding learning path with personalized educational content. One of the possible outcomes of introducing AI and ML into the eLearning process is the usage of AI-powered chatbots. Chatbots could significantly simplify the learning process by providing learners with the resources they are looking for which have previously been tailored for their learning personal characteristics. The integration of chatbots provides answers corresponding to the personal requirements of each learner into the eLearning system and could cover the gap in real-time consulting, which is not available with offline courses. Development of a chatbot is a trivial task for modern software engineering. However the determination of a chatbot behavior and its source dictionary is a complex task that requires development and implementation of additional analytic algorithms and models. This paper describes an approach for the design and development of a chatbot for an eLearning system with the personalization of consulting materials in-line with the learner's personal needs.

Keywords: Personalized learning · E-Learning · Adoptive learning · Artificial intelligence (AI) · Machine learning (ML) · Chatbots · Amazon services · Amazon Lex

1 Introduction

1.1 Adaptive Learning as an ELearning Constant Trend

The initial adaptive learning concepts were proposed back in the 1970s [1]. Today in 2020, almost 50 year later, adaptive learning is still specified as one of the main trends of modern eLearning.

A traditional educational scenario involves the same learning process with the same modules and tasks for all students. Such an approach takes no consideration of many important factors particularly those associate with the uniqueness of each individual learner. Additionally the different knowledge basis, knowledge consumptions and attention spans, are ignored by traditional learning techniques and this creates difficulties in deriving positive results for the learning process.

Adaptive learning is the approach that involves the customization of the learning process based on personal requirements and characteristics of each learner. The main goal of adaptive learning is to provide the most effective learning path for each student in-line with their specific needs [2].

Adaptive learning includes the development of specific learning activities that can be implemented in either of two modes [3]. The first mode is algorithmic adaptivity. This is the mode that current suppliers of educational modules and content provide, they correspond to the personal requirements of each learner, which are evaluated from the learner's personal feedbacks. This mode of adaptive learning considers that the most effective training process is developed using algorithms to analyze the student's learning abilities and habits. The second mode is designed adaptivity. This learning process is based on differentiation of learning paths, which is performed based on the learner's unique needs and preferences.

This approach is used to make learning easier, clear and more enjoyable, that leads to better results of education. On the other hand adaptive learning provides outcomes, optimized for learner purposes.

1.2 Usage of AI-Powered Chatbots in ELearning Trainings

Artificial intelligence (AI) is a trend not only in education, but in other aspects of our life [4]. The majority of modern eLearning innovations are connected with AI. AI powered models are widely used to perform customization of learning path by analyzing user data, collected during previous training.

According to multiple references, the key AI outcomes in 2020 are expected to be AI-powered chatbots. In general a chatbot is software developed using AI and ML to perform human interaction. [5] The meaning of the term Chatbot is derived from "chat robot" - provide engaging voice and text conversational experiences, which can be customized and used via web browsers, mobile devices, and on different chat platforms e.g. Slack or Facebook Messenger. [6] With the development of deep learning technologies (text-to-speech, automatic speech recognition and natural language processing) usage of voice-based chatbots in different areas of human life also becomes possible. The simulation of human conversation via a chatbot is widely used in call centres and

customer services. The most well-known chatbot examples are Apple's Siri or Amazon's Alexa.

In the context of eLearning chatbots could be used to provide an interactive educational process. AI-powered chatbots are positioned to perform a role of virtual instructors. This software can not only simplify searching for required content by answering some general questions. Usage of ML algorithms allows the implementation of a personalized chatbot behavior specific to the students's personal needs. Learning path with personalized chatbot settings will not only optimize time resources of each knowledge seeker, but also makes it easier to understand the learning content, since all explanations will take into account the learner's abilities and previous experience [7].

2 Chatbot Development Process

2.1 General Approach to Chatbot Development

Chatbot development is a complex task and the software designed for this purpose is directly dependent on the requirements of the business objectives. However there is a set of general steps that should be performed during the chatbot development process [8].

First step is the definition of the problem that should be solved using developed chatbot. Understanding the purpose of the chatbot is the key point for a definition of the conversation flow and determination of chatbot type.

Second step is the creation of a conversation flow. This step is the most challenging part of development, since it requires the definition of an ideal model for a chatbot potential user. Based on a developed user model, a conversation strategy, a conversation dictionary and a conversation flow are built. This task should be implemented by someone who is most involved in the business process that the chatbot is developed for.

Third step includes selection of chatbot platform. The most important decision is to choose between usage of existing platforms, such as Chatfuel, Botsify, Bottr, ChattyPeople etc., or developing a new chatbot from scratch.

Fourth step is the integration of a developed chatbot into the system it has been developed for and a fifth step is testing of the developed chatbot. Chatbot testing is a responsible task, which should not been restricted to only code testing. Chatbot operation should be evaluated only in the case of the complex testing of all system with the chatbot. Chatbot testing with different levels of load should also be performed to prevent possible failures [9].

Sixth step includes conversational analysis, which means that all conversations performed by chatbot should be analyzed to identify that the developed software archives its goals, define possible weaknesses and gaps. Performing corresponding improvements to the chatbot operation is the seventh step.

2.2 Development of a Chatbot for Personalized ELearning Process

According to the this flow of a chatbot development the most challenging task is the design of the conversation flow. In the case of personalized learning there is no possibility of identifying a single model of a perfect user. The goal of such an approach is

to cover as many different user models as possible. The operation of a developed chatbot should be different dependent on the user characteristics. This means that there are multiple variants of conversation flows [7]. This requirement makes this task more complex.

To solve this task a single correspondence between question and answer are generally used for chatbots dictionaries, should be replaced by a function:

$$x = f \left(q, \begin{bmatrix} a_1 & \dots & a_n \\ b_1 & \dots & b_n \\ c_1 & \dots & c_n \end{bmatrix} \right) \quad (1)$$

The function shown in Eq. 1 depends on two parameters: q which is a request from learner and the second parameter is a matrix of learner characteristics. The matrix can have different sizes and the content should be determined for each course before it begins. Considering this fact the determination of a response for a learner can be identified by a system of functions shown in Eq. 2:

$$\left\{ \begin{array}{l} x = f_1(q, [y_1 \dots y_m]) \\ [y_1 \dots y_m] = f_2 \left(\begin{bmatrix} a_1 & \dots & a_n \\ b_1 & \dots & b_n \\ c_1 & \dots & c_n \end{bmatrix} \right) \end{array} \right\}. \quad (2)$$

This approach considers the implementation of two functions – first is the required characteristics values from a set of all characteristics and the second is the determination of a personalized response for a provided request using previously identified characteristics.

This particular functionality is not supported by default by existing chatbot platforms and requires implementation of additional algorithms and business logic. Implementation of a chatbot has been performed using a combination of Amazon services.

3 AI-Powered Chatbot Implementation Using AWS Services

3.1 Development of Chatbot Using AWS Lex

Implementation of the chatbot described in Sect. 2 has been performed by using AWS provided services. Amazon Lex service is used as a platform for developing the chatbot. It is an easy-to-use secured auto-scalable fully managed service that provides tools to solve deep learning problems [6]. Amazon Lex is integrated with AWS Lambda, which can perform processing, and business logic tasks, and Amazon CloudWatch, which can significantly simplify the process of conversational analysis.

The chatbot has been designed using the following Amazon services:

- Amazon Cognito, which performs user authentication and user management;
- Amazon DynamoDB, used as additional storage for personal user characteristics required for personalization of learning content;

- Amazon Lex, which is directly used for developing chatbot that will perform the interaction learning process and recognize provided data;
- AWS Lambda, that performs additional analysis of requested information considering personal user characteristics, obtained from DynamoDB, and returns corresponding data to Amazon Lex;
- Amazon CloudWatch to monitoring chatbot operation;
- Amazon S3 to store learning materials, which are tagged in accordance with the level of learner characteristics, valuable for their understanding;

The general architecture of the developed chatbot can be seen in Fig. 1.

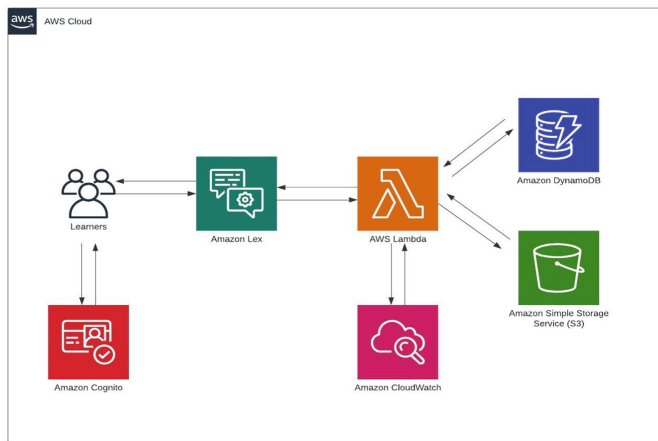


Fig. 1. Chatbot architecture

To implement this flow course, the trainer should specify characteristics, which are valuable for the course understanding, and provide a few versions of the learning content group based on the characteristics values. The set of characteristics, specified for the course, are stored in DynamoDB, and the educational materials are placed in Amazon S3. The search of the required content is implemented using tagging. Each file or record in S3 receives the set of tags which corresponds to the characteristics specified for the course where the materials belong. Each tag can have a single value or range of values. Materials, stored in S3 can contain a single answer, dictionary of answers or a set of tasks, which should be implemented by the learner with specific parameters. This is a high flexibility approach provided in the settings of the chatbot and expands its functionality. Flow course trainers should specify characteristics, which are valuable for the course understanding, and provide a few versions of the learning content group based on the characteristic values. The set of characteristics, specified for the course, are stored in DynamoDB, and the educational materials are placed in Amazon S3. The search of the required content is implemented using tagging. Each file or record in S3 receives the set of tags, which corresponds to the characteristics, specified for the course, where the materials belong. Each tag can have a single value or

range of values. Materials stored in S3 can contain a single answer, dictionary of the answers or set of tasks, which should be undertaken by the learners with specific parameters. This approach provides a high flexibility in the settings of the chatbot and expands its functionality.

Each learner of the system has a record in the Amazon Cognito user pool and corresponding account in DynamoDB. Amazon Cognito service performs secure authentication and DynamoDB is used to save the personal characteristics of each user. DynamoDB is noSQL, which allows storing unique characteristics set for each user. At the start of each course, the learner will have generated a set of the evaluated characteristics, which are specified as options for this course. The characteristics value could be obtained from the previous courses, e.g. based on the results of the course in related disciplines, or completed manually by the user or trainer.

When a particular learner sends a request to the chatbot, it is processed first by Amazon Lex to define the question according to its own dictionary. Then it is sent to AWS Lambda. AWS Lambda obtains the user personal information and course settings from DynamoDB and requests data for response from the S3 bucket, using tags as search parameters. The received data is then sent back to Amazon Lex to provide the user with a chatbot.

3.2 Experimental Usage of AI-Powered Chatbot for Learning Courses

The AI-powered chatbot, developed using the approach described in Sect. 3.1 has been integrated with the simple eLearning system used in the English language courses. Chatbot has been used by students as a tool to perform as home exercises. It has provided not only answers to the learner questions, but also the sequence of tasks for each learner. Experiments have been conducted with a small group. These groups are generally organized from the people with a similar level of previously obtained knowledge (pre-intermediate, intermediate, upper-intermediate etc.). It means that dividing material according to background level has not been implemented. The determined set of characteristics includes learner position (manager, developer, designer etc.), professional or educational background (computer science, math or finances) and personal characteristics, obtained via psychological testing (e.g. open for communication, leadership, requirement of stability or variety in working process, flexibility etc.).

The division of the learning materials has been performed according to the following rules:

1. All examples have been provided according to the topics related to the learner's professional background to make them more understandable.
2. Different psychological types obtain different tasks via chatbot as a task source. People with stability preference received the set of similar tasks, learners with that require a variety have been provided with different types of tasks. Learners with communication problems obtain more reading tasks, with communication preferences – more interview tasks etc.

The results of the conducted experiment were evaluated compared to the results of another group, that didn't work with a chatbot for home exercises. The comparison can be seen in Table 1. All results are presented as ratios in percentage terms and have been calculated as a summary of the complete group.

Table 1. Comparison of the learning outcomes for group with and without chatbot support

Outcome	Using chatbot	Without chatbot
Total number of completed tasks	87%	64%
Number of the tasks, completed successfully (60–80% of correctness)	51%	39%
Number of the tasks, completed successfully (>80% of correctness)	37%	22%
Time, spent on the tasks completion (compared between groups)	86%	100%
Successful passed assessments	83%	71%

As can be seen from the table, the value of each evaluation is higher, than for the group without chatbot that performs learning path personalization. However there are drawbacks to such an approach, for example people who didn't perform speaking tasks didn't get speaking practice however the general task of improvement of English level has been implemented successfully.

4 Conclusions

The main benefit of adaptive learning is the ability to deliver more effective training. Introduction of AI and ML into eLearning leads to the appearance of a new modification of adaptive learning, which is called personalized learning. Personalized learning is based on the personalized learning paths and customized learning context. AI-powered chatbots are the popular eLearning direction from previous years. Chatbots could provide customized guidance for each learner dependent on their personal needs. Such chatbots could be a solution to the problem of the lack of communication as in the case of online learning system usage. This paper describes the development of a chatbot for an eLearning system, which can provide customized material based on different user parameters. Integration of this type of chatbot can simplify the educational process, but only in the case of high-quality content, provided by real instructor. Chatbot could perform the role of online assistant, but there are no reasons to make the decision that it could replace real instructors. Chatbot is just a tool to simplify their work and make their courses more available. The proposed chatbot has been constructed using Amazon services. It is based on Amazon Lex and processes only text message, however Amazon Lex also provides the ability of using automatic speech recognition (ASR), which can be used for further improvements.

References

1. Carbonell, J.R.: AI in CAI: an artificial intelligence approach to computer aided instruction. *IEEE Trans. Man-Mach. Syst. MMS* **11**, 190–202 (1970)
2. Brusilovsky, P.: Methods and techniques of adaptive hypermedia. In: *User Modeling and User-Adapted Interaction*, pp. 87–129. Kluwer Academic Publishers (1996)
3. Webb, G.I., Pazzani, M.J., Billsus, D.: Machine learning for user modeling. *User Model. User Adapt. Interact.* **11**(1), 19–29 (2001)
4. Crutzen, R., Peters, G.J.Y., Portugal, S.D., et al.: An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and alcohol: an exploratory study. *J. Adolesc Health* **48**(5), 514–519 (2011)
5. Bostrom, N., Yudkowsky, E.: The ethics of artificial intelligence. In: Frankish, K. and Ramsey, W.M. (eds.) *The Cambridge Handbook of Artificial Intelligence*, pp. 316–334. Cambridge University Press (2014)
6. AWS Documentation. <https://docs.aws.amazon.com>
7. Dumais, S.T., Platt, J., Heckerman, D., Sahami, M.: Inductive learning algorithms and representations for text categorization. In: *Proceedings of the 7th International Conference on Information and Knowledge Management* (1998)
8. Tripathy, A.K., Carvalho, R., Pawaskar, K., et al.: Mobile based healthcare management using artificial intelligence. In: *2015 International Conference on Technologies for Sustainable Development (ICTSD)*, pp. 1–6. IEEE (2015)
9. Davies, J.N., Verovko, M.V., Posadska, A.S., Solomakha, I.V.: Usage of simulation for QA in real-time network environments. In: *Математичні машини і системи*, no. 4, pp. 117–125 (2019)
10. Gaudioso, E., Boticario, J.G.: User data management and usage model acquisition in an adaptive educational collaborative environment. In: *Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*, pp. 143–152. Springer (2002)



Methods for Determining the Group Ranking of Alternatives for Incomplete Expert Rankings

Hrygorii Hnatiienko^(✉) , Nataliia Tmienova ,
and Alexander Kruglov 

Faculty of Information Technology, Taras Shevchenko National University
of Kyiv, Kiev, Ukraine
g.gna5@ukr.net, tmyenovox@gmail.com,
awahrhaft@gmail.com

Abstract. Much of the practical situations can be formalized in a class of object group ranking problems. Since it is impossible to solve the object group ranking problem by exact algorithms in large dimensions, it is promising to apply genetic and heuristic algorithms. This paper analyzes the algorithms for determining the resulting ranking based on incomplete expert rankings. We identified the advantages and disadvantages of the proposed approaches. Among the main results, we would like to highlight the following. First of all, the problem of determining the group ranking of alternatives based on incomplete expert rankings was formalized. The methods of aggregation of expert data, taking into account the peculiarities of incomplete information received from experts, were considered. Besides, we developed the algorithms for solving the problems of large-scale collective ranking calculations. The proposed approaches and algorithms can solve the following problems: development of the system of integrated cybersecurity measures; search for the ranking of the most popular cryptocurrencies on the exchanges and determining the integral ordering of the cryptocurrency popularity; building a sequence of lectures in the problems of curriculum development.

Keywords: Object group ranking · Incomplete expert ranking · Metric · Distance · Median

1 Introduction

Much of the practical situations can be formalized in a class of object group ranking problems. In this case, the number of objects that need to be ordered can be estimated in the tens and hundreds, and the number of experts whose preference systems should be considered can reach several thousand. Since the object group ranking problem is NP-hard, it is impossible to solve it by exact algorithms in large dimensions. Therefore, it is promising to apply genetic and heuristic algorithms to solve object collective ranking problems.

At the same time, in practice, there are often situations where experts are unable to accurately rank the objects from the entire set that characterize the decision-making situation. The advantage of the alternative ranking procedure is the ease of

implementation of this method of information structuring. However, as the number of alternatives approaches twenty, it becomes difficult for experts to arrange them, as this procedure requires that all the alternatives are interconnected. The ability to operate large amounts of information is limited by the psychological characteristics of humans, so in such cases, experts can make significant mistakes. To require them to accomplish this task is to knowingly create precarious initial data.

When formalizing decision support problems using expert technologies, a mathematical apparatus, which has proven itself well in many practical problems such as selection problems, binary relations, and criterion selection [1], is used. In addition, to solve problems arising from the use of expert technologies, heuristics are used. In some cases, experts are invited to participate in the formation of the problem adequate model, in others, to choose the formula of information aggregation, to determine the preference relation on the individual elements of the mathematical model, to determine the event probability, or to choose metrics and criteria for determining the best problem solution, etc. Moreover, in many practical situations, expert estimation is the only possible way to solve real problems. The successful use of expert estimates in this process is largely driven by the excellence of the mathematical apparatus for analyzing and processing expert information [2].

Methods of expert information processing are divided into three main groups, such as statistical methods, scaling methods and algebraic methods [1]. In this paper, we will discuss the problem formulation and the algorithms, which are based on an algebraic approach. In addition, we will use a combination of qualitative and quantitative analysis of object ranking problems. The essence of algebraic methods is that the distance sets on the acceptable estimates set and the resulting estimate is defined in such way: the distance from the estimate to the expert estimates is minimal for the selected criterion. Based on the algebraic approach, we will determine the resulting rankings closest to the incomplete expert rankings.

Some issues of using the incomplete ranking of alternatives were considered in [3]. But in the monograph [3], the algebraic methods of calculating the median were not applied to the problems of determining the resulting object ranking. At present, many scientific works are devoted to the problem of incomplete expert assessments, in particular [4, 5].

2 The Purpose of the Article

The purpose of the article is to develop approaches to solving the group ranking problem based on the incomplete expert ranking of alternatives. Based on the proposed approaches, it is necessary to develop the algorithms for calculating the resulting ranking of alternatives, which is consistent with the given incomplete expert rankings, taking into account different metrics and different criteria. To create a toolkit for comparing the incomplete ranking of alternatives, a mathematical model, which will include an approach that allows determining the distances between incomplete rankings, should be developed.

3 The Methods of Research

The main methods used in this article are the methods of decision theory [1], expert technologies [2], genetic algorithm [2] and the heuristic algorithm [2]. Methods of decision theory [1] originate from the theory of operations research; they stimulated the development of artificial intelligence [1]. Expert technologies are widely used in various fields of human life and have been strongly developing in recent decades. One of the key concepts of expert technologies is heuristics [2], which can be axioms, postulates, assumptions, presumptions, paradigms, hypotheses, additions, propositions, etc. Heuristics is an empirical methodological rule that can help to find a solution and to define the incorrectly assigned problems. Heuristic algorithms help to generate solutions that are close to optimal but finding the optimal solution of the task is not guaranteed. Genetic algorithms are evolutionary algorithms used, in particular, to solve optimization problems, but they also do not guarantee global optimality of the problem.

4 The Main Results

For the further presentation of the material, we introduce the definition of incomplete object ranking R^H of the set A . R^H is a binary relation defined on a subset of objects A' , $A' \subset A$, which satisfies the properties of completeness, reflexivity, antisymmetry, transitivity, but only on the subset A' , $A' \subset A$. Let the team of experts sets k incomplete ranking of alternatives R^{iH} , $i = 1, \dots, k$. It needs to find some resulting (aggregated, collective, consensus, group) ranking of n alternatives (tasks or cybersecurity activities) $R^* = (a_{i_1}, \dots, a_{i_n})$, $i_j \in I = \{1, \dots, n\}$, $j \in I$, which is built on the logic that characterizes the processes of functioning of an organizational system. In other words, ranking R^* is based on the individual ordering of problems performed by k system elements (experts) $R^{iH} = (a_{1_i}, \dots, a_{n_i})$, $i \in J = \{1, \dots, k\}$, where n_i is number of tasks in individual expert ranking $i \in J$.

The problem is to determine the resulting ranking R^* , which relatively to some criterion is “the closest” to all individual rankings given by k experts. This ranking should also be Pareto optimal and satisfy the unanimity axioms [1]. The problem of finding order on the set of objects is a complex combinatorial problem, NP-hard in the strong sense [2]. Therefore, local optimization algorithms, heuristic algorithms, sequential analysis schemes, or branch and boundary algorithms are used to construct the resulting ranking R^* . It should be noted that, in the general case, the coefficients of relative competence of experts or other sources of information that specify incomplete ranking of alternatives may be known, calculated or given [6–8]. According to the conventional approach [1, 4], when applying the algebraic approach, metrics are introduced to measure distances. The most common metrics for such problems are the Cook metric of the ranks (places, positions) mismatch of alternatives and the Hemming metric. The criteria most often used in such cases are additive and minimax. The most common method of finding the resulting ranking of objects is to calculate the median of the given rankings. This group of methods of generating expert information is the most reliable and mathematically sound. Problem solutions, which are determined by

applying different metrics and different criteria, are the medians of the experts' linear orders, whose formal descriptions and names will be discussed below.

Let Ω^R be the set of all possible rankings of n objects; Ω^B be the set of pairwise comparison matrices (PCM) corresponding to the all possible rankings of n objects, R^A be the set of experts' rankings and corresponding binary relations; R^B be the set of corresponding rankings of the PCM and binary relations between alternatives. Problems associated with incomplete PCM were explored, for example, in [9].

For the case considered in this paper, the cardinals of R^A and R^B are the same. It is clear that $R^A \subset \Omega^R, R^B \subset \Omega^B$. In the general case, the cardinal of set is $|\Omega^B| = |\Omega^P| = 2^{n(n-1)/2}$. But for the method described in this paper, we will consider only their subsets, marking these subsets in the same way $|\Omega^R| = |\Omega^B| = n!$, since we are not interested in non-transitive elements of the solution space Ω^B .

4.1 Union of Subsets of Alternatives Used in Expert Ranking

Union of the all experts' alternatives $a_i \in A, i = 1, \dots, n$, into a single set to determine the resulting ranking R^* is the initial stage of the problem solving described in this paper. Based on the k experts' incomplete rankings $R^{iH}, i \in I = \{1, \dots, k\}$, on subsets $A_i \subset A, i \in I$, we define a set of alternatives: $A = \bigcup_{i=1}^k A_i$.

4.2 Using the Cook Metric for Incomplete Ranking of Alternatives

In the first case, to apply the algebraic approach on the set of rankings to measure the distances between the rankings of alternatives, the Cook metric of the ranks (places, positions) mismatch of the alternatives in the individual rankings is used [1]:

$$d(R^j, R^l) = \sum_{i \in I} |r_i^j - r_i^l|, \tag{1}$$

where r_i^l is the rank of the i -th alternative in the ranking of the l -th expert, $R^l, l \in L, 1 \leq r_i^l \leq n$.

For the Cook metric (1) using the utilitarian criterion, the following characteristics are calculated:

- Cook-Sayford median [1]:

$$R^{CS} \in \Omega^{CS} = \text{Arg min}_{R \in \Omega^R} \sum_{l \in L} d^r(R, R^l) \tag{2}$$

- modified Cook-Sayford median [6]:

$$R^{MCS} \in \Omega^{MCS} = \text{Arg min}_{R \in R^A} \sum_{l \in L} d^r(R, R^l). \tag{3}$$

When using the egalitarian criterion, the following characteristics are calculated:

- GV-median (compromise) [1]:

$$R^{GV} \in \Omega^{GV} = \text{Arg min}_{R \in \Omega^R} \max_{l \in L} d^r(R, R^l) \tag{4}$$

- modified GV-median:

$$R^{MGV} \in \Omega^{MGV} = \text{Arg min}_{R \in R^A} \max_{l \in L} d^r(R, R^l). \tag{5}$$

The Cook metric is popular in alternative ranking problems [1]. To use it to solve the problem of the analysis of incomplete rankings, we will introduce heuristics and determine the distances from the experts' rankings to the reference ranking, using the entered heuristics.

Heuristics E1. The distance from any ranking $R^{*(0)}$ to each expert's ranking R^{iH} , $i = 1, \dots, k$, is equal to the sum of the probabilistic and determined part.

Heuristics E2. The alternative not specified by expert generates an unknown relationship between all other alternatives and does not participate in the ranking, that is, this alternative is not represented in the incomplete ranking. In other words, when setting incomplete rankings for each expert, we have a number of alternatives: n_i – alternatives specified by expert in ranking R^{iH} , $i = 1, \dots, k$, which will form a determined part of the distances;

$(n - n_i) = v_i$ – alternatives not specified by expert in ranking R^{iH} , $i = 1, \dots, k$, which will form a probabilistic part of the distances.

4.3 The Transition from the Space of Ranges to the Space of Pairwise Comparisons of Alternatives

Individual incomplete benefits given by each expert on subsets of alternatives R^{iH} , $l \in L$, can also be represented as an incomplete pairwise comparisons matrix

$$B^{iH} = (b_{ij}^{iH}), j \in I, l \in L, \tag{6}$$

where $b_{ij}^l = 1, i, j \in I, l \in L$, if and only if the i -th alternative outweighs the j -th alternative according to the l -th expert. If the l -th expert considers that $a_i \prec a_j$, then $b_{ij}^l = -1$, when $a_i \succ a_j$, then $b_{ij}^l = 1, i, j \in I, l \in L$. In the case when the l -th expert does not specify the preference relation between the alternatives a_i and a_j , then this fact will be denoted by $b_{ij}^{iH} = *, j \in I, l \in L$.

The Hemming metric is used to determine distances between relations (6): $d^h(B^j, B^l) = 0, 5 \sum_{i \in I} \sum_{s \in I} b_{is}^j - b_{is}^l$. Since the matrices of the relations B^H and R^H of the foThen the distance betweenrm (1) are skew-symmetric, we will use vectors based on them: $c_t = b_{ij}, x_t = r_{ij}, t = (i - 1)n + j - (i + 1)i/2, 1 \leq i < j \leq n$.

Let $N = n(n - 1)/2$ be the elements number of vectors c_t , and $H = \{1, \dots, N\}$ be the indices set of elements of these vectors. Then the distance between the relations B and R will be written in the form

$$d^h(B^j, B^l) = \sum_{t \in H} |c_t^j - c_t^l|, \quad t \in H. \tag{7}$$

4.4 Using of Hemming Metric for Incomplete Ranking of Alternatives

Heuristics E3. The mathematical expectation of indefinite distances between alternatives in the ranking is equal to 1. That is, the distance between the elements of the PCM at least one of which is undefined should be equal to 1, assuming that distance values “-1” or “1” are equally probable.

The probabilistic part from an expert-specified ranking $R^{ih}, i = 1, \dots, k$, to any other ranking for the Hemming metric is always equal to $v_i \cdot (v_i - 1)/2, i = 1, \dots, k$.

In order to apply the heuristic algorithm, we can take into account the interrelation between the ordinal and the cardinal ways of preference setting. Partial-order aggregation can be accomplished by accumulating preferences between all pairs of alternatives. After that, we need to make the transition from the ordinal scale to the cardinal scale and calculate the integral value of the preference relations in the cardinal scale. After calculating the resulting cardinal relation, we need to translate it into the ordinal scale and obtain the resulting ranking of alternatives.

For the Hemming metric (7) using the utilitarian criterion, the following characteristics are calculated:

- median of Kemeni-Snell:

$$R^{KS} \in \Omega^{KS} = \text{Arg min}_{B \in \Omega^B} \sum_{l \in L} d^h(B, B^l), \tag{8}$$

- modified median [2] of Kemeni-Snell:

$$R^{MKS} \in \Omega^{MKS} = \text{Arg min}_{B \in R^B} \sum_{l \in L} d^h(B, B^l). \tag{9}$$

When using the egalitarian criterion, the following characteristics are calculated:

- VG-median (compromise) [3]:

$$R^{VG} \in \Omega^{VG} = \text{Arg min}_{B \in \Omega^B} \max_{l \in L} d^h(B, B^l). \tag{10}$$

- modified VG-median:

$$R^{MVG} \in \Omega^{MVG} = \text{Arg min}_{B \in R^B} \max_{l \in L} d^h(B, B^l). \tag{11}$$

4.5 Genetic Algorithm for Determining the Median of Expert Ranking of Alternatives

Mutations and crossovers are used in genetic algorithms to generate new generations. However, for rankings, classic crossover techniques do not work because strict ranking $R \in \Omega^R$ should consist of n non-repeating elements.

In the case of a single mutation, we will rearrange two random elements $r_i, r_j \in R, i \neq j, i, j \in \{1, \dots, n\}$, in the resulting ranking R^* concerning to their initial position. The mutation function $f(R)$ will be:

$$f(R) = \begin{cases} \exists r_i, r_j \in R, i \neq j : (r_{i-1} > r_i > r_{i+1}) \vee (r_{j-1} > r_j > r_{j+1}), \\ \exists r_i^*, r_j^* \in R^* : (r_{i-1} > r_j^* > r_{i+1}) \vee (r_{j-1} > r_i^* > r_{j+1}), r_i = r_i^*, r_j = r_j^*, \end{cases}$$

where R^* is the ranking resulting from the mutation. This transformation is repeated m times, so, we have $f^m(R), m \in \{0, 1, 2, 3, 4\}$.

Ranking R^* will be the crossover of pair R^1, R^2 . We define the function of the crossover in such way:

$$g(R^1, R^2, i, j) : R^* \cap R^1 \forall r_k \in R^1, i \geq k \geq j, R^* \cap R^2 \forall r_k \notin R^* \cap R^1.$$

Thus, some of the elements of the resulting ranking R^* will be ordered as in R^1 , and all other elements will be ordered as in R^2 .

Each distance from the current median to expert rankings $R^{iH} = (a_1, \dots, a_{n_i}), i \in J = \{1, \dots, k\}$, is calculated according to the rules for incomplete rankings described in Sect. 4.2 for the space of alternatives ranks or in Sect. 4.4 for the space of pairwise comparisons between alternatives.

4.6 The Results of the Application of Genetic Algorithm for Determining the Median of Expert Ranking of Alternatives

Numerous computational experiments performed using the genetic algorithm prove the effectiveness of this approach. For randomly generated rankings of 100 objects, the program in the space of all possible rankings $R \in \Omega^R$ consistently calculates the medians in forms (2), (4), (8), (10). They are closer to the median given by experts $R^l, l \in L$, than the corresponding modified medians in forms (3), (5), (7), (11) by an average of 8–10%. That is, the medians selected from expert rankings $R^{iH}, i \in I = \{1, \dots, k\}$, are consistently dominated by the medians calculated using the genetic algorithm. Moreover, the effectiveness of the method is fixed for all the medians described above.

4.7 Method of Heuristic Determining the Medians

Let describe a proposed heuristic algorithm for determining the resulting ranking R^* of the given set of incomplete rankings $R^{iH}, i \in I$, in the form of the Kemeni-Snell.

1. Constructing a deviation matrix $P = (p_{ij}), i = 1, \dots, n, j = 1, \dots, N$, based on an analysis of all possible relations between the triple of alternatives $p_{lj} = m_{lt}, j = (l - 1) * n + t - l * (l + 1) / 2, 1 \leq l \leq t \leq n$,

$$p_{ij} = \begin{cases} m_{ls} / m_{ts}, l \langle s, s \rangle t, \\ m_{ls} * m_{ts}, l < s, s < t, \\ m_{st} / m_{sl}, l > s, s > t, \end{cases}$$

where $m_{lt} = \sum_{\substack{i \in I, \\ c_{ij} \geq 0}} c_{ij} / \sum_{\substack{i \in I, \\ c_{ij} < 0}} |c_{ij}|, 1 \leq l \leq t \leq n, j \in J$, are the elements of a metrized matrix formed by calculating the number of preferences between alternatives given in incomplete expert rankings $R^{iH}, i \in I; s = 1, \dots, n, s \neq l, s \neq t, i = l + 1, l = 1, \dots, n - 1, t = l + 1, \dots, n, j = (l - 1) * n + t - l * (l + 1) / 2$.

2. For each index $i = 2, \dots, n$, replacing one of the elements of the matrix $M = (m_{lt}), l, t \in I$, by the element of the matrix $P = (p_{ij}), i = 2, \dots, n, j = 1, \dots, N$, if the elements do not match: $m_{lt} \neq p_{ij}$, for $j = (l - 1) * n + t - l * (l + 1) / 2$.
3. Determining the weighting factors of each of the n alternatives by the method of row or column sums, which is the standard of computational ease of determining “weight”, for each next generated matrix.
4. Ordering the elements of the calculated vector by decreasing values (for row sums) or increasing values (for column sums). Indices of such ordering vector will be considered as indices of alternatives in the resulting ranking.
5. Determining of the sum of distances from the ranking obtained above to ranking specified by experts according to the rules described for incomplete pairwise comparison matrices.

4.8 The Results of the Application of Heuristic Algorithm for Determining the Median of Expert Ranking of Alternatives

The authors conducted experiments for matrices that have dimensions from 6×6 to 10×10 , that is, when the rankings of 6-10 objects are given. Experiments using the described method, verified by exact methods, gave the following results:

- if the matrices specified by experts are poorly matched, the set of effective solutions can be very large, that is, the number of Kemeni-Snell medians is up to 15% of the total number of object rankings on the set of objects, i.e. up to $0,15 \cdot n!$;
- between solutions found using the algorithm described above, when applying the row sums method, up to 50% of the object rankings are the Kemeni-Snell medians, when applying the column sums method, up to 83% of the object rankings are the Kemeni-Snell medians;
- the application of the described algorithm in some cases allows to find up to 39% of rankings belonging to the set of Kemeni-Snell medians;
- among the compromise rankings found using the described algorithm, up to 22% are simultaneously VG-medians. Thus, the described algorithm is a convenient

heuristic algorithm for determining the Kemeni-Snell median sets and the VG-median sets. Although all the set of effective solutions can not be determined, some of the medians are guaranteed to be calculated. The conducted research confirms the interconnection between the ordinal and cardinal models of assigning expert preferences, as well as the prospect of application of heuristic algorithms in the problems of expert evaluation.

5 Conclusions

This paper analyzes the algorithms for determining the resulting ranking based on incomplete expert rankings. We identified the advantages and disadvantages of the proposed approaches. Among the main results, we would like to highlight the following. First of all, the problem of determining the group ranking of alternatives based on incomplete expert rankings was formalized. The methods of aggregation of expert data, taking into account the peculiarities of incomplete information received from experts, were considered. Besides, we developed the algorithms for solving the problems of large-scale collective ranking calculations.

The results of the computational experiments showed the obvious advantages of the developed algorithms, their increased accuracy in comparison with those used so far.

The described algorithms for calculating the resulting ranking of the given incomplete expert ranking of alternatives can be applied to solve various problems in different subject areas:

- development of the system of integrated cybersecurity measures;
- search for the ranking of the most popular cryptocurrencies on the exchanges and determining the integral ordering of the cryptocurrency popularity;
- building a sequence of lectures in the problems of curriculum development.



References

1. Voloshin, O., Mashchenko, S.: Decision-Making Theory: Tutorial (Ukrainian). Kyiv University, Kyiv (2006)
2. Snytyuk, V.Y., Suprun, O.: Evolutionary techniques for complex objects clustering. In: 2017 IEEE 4th International Conference on Actual Problems of Unmanned Aerial Vehicles Developments, APUAVD – Proceedings, Kyiv, Ukraine (2017)
3. Dodonov, A., Lande, D., Tsyganok, V., Andriichuk, O., Kadenko, S., Graivoronskaya, A.: Information Operations Recognition. From Nonlinear Analysis to Decision-Making. Lambert Academic Publishing (2019)
4. Samokhvalov, Yu Ya.: Development of the prediction graph method under incomplete and inaccurate expert estimates. *Cybern. Syst. Anal.* **54**(1), 75–82 (2018). <https://doi.org/10.1007/s10559-018-0008-1>

5. Tsyganok, V., Kadenko, S., Andriichuk, O., Roik, P.: Combinatorial method for aggregation of incomplete group judgments. In: Proceedings of 2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC), Kyiv, Ukraine, pp. 25–30. Igor Sikorsky Kyiv Polytechnic Institute (2018)
6. Tsyganok, V., Kadenko, S., Andriichuk, O.: Considering importance of information sources during aggregation of alternative rankings. In: CEUR Workshop Proceedings, Selected Papers of the XVII International Scientific and Practical Conference on Information Technologies and Security (ITS 2017), Kyiv, Ukraine, vol. 2067, pp. 132–141 (2017)
7. Hnatiienko, H., Snytyuk, V.: A posteriori determination of expert competence under uncertainty. In: CEUR Workshop Proceedings, Selected Papers of the XIX International Scientific and Practical Conference “Information Technologies and Security” (ITS 2019), Kyiv, Ukraine, vol. 2577, pp. 82–99 (2019)
8. Tsyganok, V., Kadenko, S., Andriichuk, O.: Significance of expert competence consideration in group decision making using AHP. *Int. J. Prod. Res.* **50**(17), 4785–4792 (2012)
9. Bozóki, S., Tsyganok, V.: The (logarithmic) least squares optimality of the arithmetic (geometric) mean of weight vectors calculated from all spanning trees for incomplete additive (multiplicative) pairwise comparison matrices. *Int. J. Gen Syst* **48**(4), 362–381 (2019)
10. Samokhvalov, Y., Naumenko, E.: Expert Evaluation. Methodical Aspect: Monograph (Ukrainian). DUKT, Kyiv (2007)
11. Orlov, A.: Management Decision-Making Methods: Textbook (Russian). KNORUS, Moscow (2018)



Profiling of Clusters in Information Technologies of Intellectual Monitoring

Serhii Holub^(✉)  and Svitlana Kunytska^(✉) 

Cherkasy State Technological University, Cherkasy, Ukraine
s.holub@chdtu.edu.ua, kunytskaya33@gmail.com

Abstract. The article represents the results of studies of the processes of formation (profiling) of agents-clusterer in the monitoring intellectual systems. The method of clustering according to the results of modeling with simultaneous construction of clusters with heterogeneous properties was investigated. Clustering profiling is done by parametric optimization of model synthesis process and agent structural optimization. Optimization was performed by changing the base model, the criterion for model selection and the number of models that go to the higher row of selection. The agent efficiency criterion was used to compare the results. An agent effective indicator is suggested. The weight of the modeling variables was used to describe the agent profile. The distance between the points is a characteristic of the results of their modeling by profiled agents. Experimental confirmation of the homogeneity of observation points in clusters grouped by cluster agents with different profiles was obtained. The obtained results make it possible to expand the capabilities of intelligent monitoring systems.

Keywords: Clustering · Profiling · Intellectual monitoring · Agent approach

1 Introduction

Intelligent Monitoring Information Technology (IMIT) implements an agentic approach to providing knowledge of decision-making processes. IMIT Contains Methods and Tools for Using Specialized Intelligent Agents capable of Solving Global Monitoring Objectives (GMO). An external contract specifies the formulation of a GMO. The GMO then agrees with the customer.

IMIT builds one or more Monitoring Intelligent Agents (MIA). If necessary, MIA is combined in Monitoring Agent Functionalities. For this purpose, a subsystem of multi-agent monitoring environments of the intelligent monitoring system is being adapted, which adapts each of the stages of intellectual monitoring to the tasks agreed with the customer. Depending on the global monitoring task, a list of controlled features and observation points is formed, the frequency of observations is determined, and a hierarchy of data processing and information form transformation tasks is formed. From the numerical arrays of observation results, information is obtained about the properties of the monitoring object and knowledge about trends and patterns of change in these properties.

Subsystem multi-agent environments are being formed to implement one of the stages of smart monitoring technology. Converter-agents are elements of a multiagent environment that are part of the MIA model knowledge base. Intelligent agents of various purposes are built and used to process the results of observations. The identification of the object state and the assessment of the influence of environmental factors and internal states are ensured by grouping the observation points in the input data array (IDA).

This paper examines the processes of constructing and profiling clusterer-agents that detect the presence or absence of properties specified by the researcher in the monitoring object. They also identify monitoring objects with properties of interest to the researcher.

2 Results of Informative Search of Analogues

Grouping observation points in IDAs with predefined properties is done by solving the classification task. If it is recognized that the observation object belongs to a class or cluster, it is considered established that the properties of the class or cluster and the state of the monitoring object at the time of obtaining the numerical characteristics of its features are the same.

Classes are formed from observation points whose properties are predefined. The direct application of classification features makes it easier to identify similar points. If the direct application of classification features is ineffective, then decisive rules are built. These rules formalize the process of using classification features to identify similar observation points.

There are monitoring tasks for which experts cannot find points with known properties. And since such a task is performed for the first time, it is not possible to use previous experience to describe the properties of observation points. Therefore, to group the observation points in the IDA, it is necessary to solve the problem of clustering [1]. To combine observation points into groups according to traditional clustering methods [2] it is necessary to choose the criteria and metrics of their similarity, to build and analyze a similarity matrix. Distance, correlation, and others are used as similarity metrics [3]. One cluster will include points whose distance from the center of the cluster is less than their distance to the center of the IDA. The distance of Euclid, Mahalanobis, Hemming is used as a measure of similarity between points [2].

Today, a common measure of distance is the standard deviation of characteristic vectors from the centers of clusters, which are defined in one way or another.

Analysis of similarity matrix is performed to find the centers of clusters around which clusters of observation points are located [2]. To solve the problems of similarity matrix analysis, probabilistic modeling is used, plausible heuristics are used, in particular, the method of k-means [3], neural networks [5], the Group Method of Data Handling (GMDH) [6].

These traditional clustering methods are universal. They do not take into account the purpose of clustering. That is, for clustering observation points with tasks that differ from each other, the number of clusters and their composition will not differ.

In [7] describes a new approach to determining the degree of similarity of observation points, which allows to develop a new method of clustering based on simulation results. It is proposed to use the characteristics of models constructed from a given array of observation results (modeling error, standard deviation, etc.) as a measure of the distance between points. Models can be built according to one of the algorithms of GMDH or other evolutionary methods, can take the form of trained neural networks. In our opinion, this method, after improvement, will solve the problem of profiling clusters.

3 Formulation of the Problem and Approach to Its Solution

The universality of existing clustering methods on the one hand, and the various tasks of grouping observation points on the other, lead to contradictions. The large number of observation point grouping tasks involves identifying properties that display relationships between simulation variables that are more complex than additive. And having made the observation point the center of the cluster, it is not possible to detect similar points by existing clustering methods.

Profiling of clusterizers eliminates this contradiction and finds observation points that describe the processes of interest to the researcher. The purpose of this work was to create a method of profiling clusterer-agents that would allow grouping of observation points by predefined properties.

4 Research and Their Results

The observation points are called X_i feature vectors that describe the properties of the monitoring object and are fixed at a particular point in time or at a specific value of one of the status indicators of that object, $y_j \in Y$.

$$X_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,k}\}, i = \overline{1, n} \quad (1)$$

where n – is the number of observations (rows in the input array), k – is the number of characters (columns) whose numerical characteristics are determined.

The array of input data is formed by feature vectors formed by observation results [7]. They allow you to build a model of the monitoring object:

$$Y_m = f(X_i) \quad (2)$$

It was hypothesized that a) the cluster profile is described by the values of the influence characteristics of the modeling variables; b) forming a profile of the cluster model can be achieved by changing the parameters of their construction; c) clusterer profiling allows you to group homogeneous points with predefined properties.

The values of the weights of the variables describe the properties of the observation object at a given monitoring point.

The variable clusterer-model weights are calculated by the expression:

$$W_{x_j} = \frac{y'_{x_j}}{\sum_{j=1}^k y'_{x_j}} 100\%, j = \overline{1, k} \quad (3)$$

where y'_{x_j} – is the partial derivative of model (2) by variable x_j .

An experiment was conducted to test this hypothesis.

IDA is generated by converting data obtained from sources [8–10].

Table 1 provides a list of features that characterize the state of the information economy of the world states during 2000–2013.

Sample Heading (Forth Level). The contribution should contain no more than four levels of headings. The following Table 1 gives a summary of all heading levels.

Table 1. List of modeling variables

#	Characteristic	Variable
1	Gross domestic product per capita, mln. \$	y
2	Place of observation	x ₁
3	Time, year	x ₂
4	Government Debt, mln. \$	x ₃
5	External debt, mln. \$	x ₄
6	High-Tech Exports, mln. \$	x ₅
7	Exports of goods created by information and telecommunication technologies, mln. \$	x ₆
8	Imports of goods created by information and telecommunication technologies, mln. \$	x ₇
9	Gross domestic expenditures on research and development work, mln. \$	x ₈
10	Export of services provided by information and telecommunication technologies, mln. \$	x ₉
11	Imports of ICT services, mln. \$	x ₁₀

Using a variety of agents containing inductive, neural network, and hybrid model synthesizers, models that solve the problem (2) of identifying the functional dependence of GDP per capita y on the information economy $x_1 - x_{10}$ presented in Table 1 have been constructed. To investigate the properties of these agents, the influence of the features was estimated by calculating the weight coefficients of their variables by expression (3).

The clusterer agent profile was formed by selecting a model, the value of the weighting factors (3), which described the desired properties of the monitoring object. The choice was made according to the results of the tournament among models built on the algorithms of GMDH, neural networks, genetic and hybrid algorithms implemented in the synthesizer of models of monitoring intellectual system. Thereafter, the structural

optimization of the agent was performed to ensure the extreme value of its affectivity index E .

$$E = \sum_{i=1}^n \frac{1}{\Delta_i k} \quad (4)$$

where Δ_i – is absolute error in modeling an object's state indicator at the i -th observation point; n – the number of observation points IDA; k – number of metrics IDA, included in the structure of the model of the agent-cluster and the value of the weight coefficient $Wx_i > 0$.

In this work, for example, the results of structural optimization of one of the agents-cluster (agent 1) are represented. The optimization process is organized according to the scheme of small-factorial experiment. The effect on the agent's effectiveness of the following factors was investigated: the form of the reference model; type of model selection criterion (absolute error and standard deviation of simulation results from true values of status indicator); number of models moving to the next selection row. The conditions of the studies are presented in Table 2. 24 trials were conducted. One of the parameters was changed. All others were fixed.

Figure 1 shows graphically the effectiveness of the agent for each test described in Table 2.

Research has shown that the effectiveness of agent 1 will be optimal if the formation of its structure will use the parameters used in test # 10 according to Table 2.

Thus, the structural and parametric optimization of each of the clusterer-agents was carried out. Table 3 for these agents represents the impact coefficients to the numerical characteristics of the traits.

Each of the agents described in Table 3 reflects the presence of certain processes in the economy of the state, the signs of which are the correlation of values of weight coefficients. The interpretation of these ratios allows us to understand the deep processes that determined the GDP per capita of a particular country during the reporting year.

Agent 1 is able to combine observation points that describe the economic relations of countries with significant factors related to the impact of changing government debt on PPPs and the processes of exporting High-Tech goods.

Agent 2 describes the processes in which exports and imports of ICT services were of the highest importance.

Agent 3 is able to group states in the information economies that have processes in place that have been influenced by the export of ICT services and have significant regional characteristics.

The cluster formed by agent 4 contains processes that determine the regional characteristics of the country. They depend on factors that are not in the list of features and are not significantly affected by processes related to other factors, the numerical characteristics of which are presented in Table 1.

Agent 5 is able to locate observation points where GDP per capita is heavily dependent on exports of information and telecommunications services. In doing so, it is necessary to further investigate what exactly caused this impact to reduce or increase these costs.

Table 2. Trial parameters.

Test number	Form of reference model	Selection Model criterion	Number of models moving to the next selection row
1	$y = a_0 + a_1y_1 + a_2y_2$	Error	10
2	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2$	Error	10
3	$y = a_0 + a_1y_1 + a_2y_2 + a_4y_1^2 + a_5y_2^2$	Error	10
4	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2 + a_4y_1^2 + a_5y_2^2$	Error	10
5	$y = a_0 + a_1y_1 + a_2y_2$	Standard deviation	10
6	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2$	Standard deviation	10
7	$y = a_0 + a_1y_1 + a_2y_2 + a_4y_1^2 + a_5y_2^2$	Standard deviation	10
8	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2 + a_4y_1^2 + a_5y_2^2$	Standard deviation	10
9	$y = a_0 + a_1y_1 + a_2y_2$	Error	6
10	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2$	Error	6
11	$y = a_0 + a_1y_1 + a_2y_2 + a_4y_1^2 + a_5y_2^2$	Error	6
12	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2 + a_4y_1^2 + a_5y_2^2$	Error	6
13	$y = a_0 + a_1y_1 + a_2y_2$	Standard deviation	6
14	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2$	Standard deviation	6
15	$y = a_0 + a_1y_1 + a_2y_2 + a_4y_1^2 + a_5y_2^2$	Standard deviation	6
16	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2 + a_4y_1^2 + a_5y_2^2$	Standard deviation	6
17	$y = a_0 + a_1y_1 + a_2y_2$	Error	3
18	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2$	Error	3
19	$y = a_0 + a_1y_1 + a_2y_2 + a_4y_1^2 + a_5y_2^2$	Error	3
20	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2 + a_4y_1^2 + a_5y_2^2$	Error	3
21	$y = a_0 + a_1y_1 + a_2y_2$	Standard deviation	3
22	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2$	Standard deviation	3
23	$y = a_0 + a_1y_1 + a_2y_2 + a_4y_1^2 + a_5y_2^2$	Standard deviation	3
24	$y = a_0 + a_1y_1 + a_2y_2 + a_3y_1y_2 + a_4y_1^2 + a_5y_2^2$	Standard deviation	3

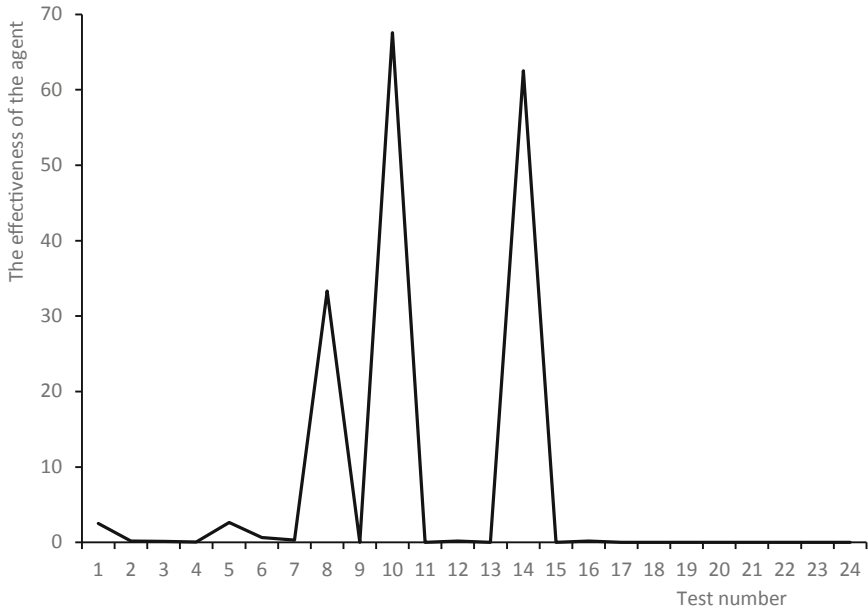


Fig. 1. The effectiveness of the agent.

Unlike Agent 2, which describes the impact of ICT services, Agent 6 can be used to find countries whose economies during the reporting year had significant processes associated with the import of ICT products and were dependent on gross domestic expenditures for research and development work.

Agent 7 can be used to identify processes that are poorly related to the components of a state's information economy, and significant factors that are not included in the time-table characteristics of Table 1 and are affected by changes in external debt.

Cluster of observation points was based on one of the characteristics of the agents - the error in predicting the simulated indicator - GDP per capita. The clustering results are represented in the Table 4.

The results of the grouping were evaluated by the homogeneity of the observation points, i.e. characteristics of the informative economy of the state during the reporting year. The higher the homogeneity of the observation points, the smaller the modeling error of the GDP per capita by the agent. The simulation results are presented in the Table 4.

According to the data in Table 5, it can be stated that in each of the 7 clusters homogeneous points of the IDA are grouped.

The results obtained provide additional information for analyzing the processes occurring in the different economies of the respective states.

Table 3. The impact of variable.

Agent	Variable	Weighting factor $Wx_i, \%$	Agent	Variable	Weighting factor $Wx_i, \%$	Agent	Variable	Weighting factor $Wx_i, \%$
1	x_1	0,90	4	x_1	91,94%	7	x_{10}	0,65
	x_2	0,00		x_2	8,06		x_1	0,00
	x_3	71,32		x_3	0,00		x_2	63,96
	x_4	0,00		x_4	0,00		x_3	0,00
	x_5	23,98		x_5	0,00		x_4	36,04
	x_6	1,50		x_6	0,00		x_5	0,00
	x_7	0,00		x_7	0,00		x_6	0,00
	x_8	0,00		x_8	0,00		x_7	0,00
	x_9	2,29		x_9	0,00		x_8	0,00
	x_{10}	0,00		x_{10}	0,00		x_9	0,00
2	x_1	19,11	5	x_1	0,20		x_{10}	0,00
	x_2	0,00		x_2	12,83			
	x_3	0,00		x_3	0,00			
	x_4	5,23		x_4	7,75			
	x_5	0,00		x_5	0,06			
	x_6	5,61		x_6	0,00			
	x_7	14,16		x_7	0,18			
	x_8	0,00		x_8	0,00			
	x_9	22,45		x_9	76,61			
3	x_1	48,81	6	x_{10}	2,36			
	x_2	0,00		x_1	8,58			
	x_3	0,00		x_2	0,00			
	x_4	0,00		x_3	0,00			
	x_5	2,57		x_4	0,00			
	x_6	0,00		x_5	4,50			
	x_7	0,00		x_6	0,00			
	x_8	0,00		x_7	26,64			
	x_9	48,62		x_8	58,89			
	x_{10}	0,00		x_9	0,74			

Table 4. Cluster structure.

Cluster	Country	Reporting year	Cluster	Country	Reporting year
1	USA	2000	5	Ukraine	2009
	Finland	2005		Canada	2010
	Singapore	2011		Netherlands	2012
	Singapore	2012		USA	2010
	Canada	2000		USA	2011
	Ukraine	2000		Canada	2011
2	Singapore	2011		Netherlands	2009
	Finland	2000		Ukraine	2012
	Netherlands	2012		Ukraine	2011
	Sweden	2009		Canada	2009
	Finland	2005		USA	2009
	Ukraine	2005		Ukraine	2010
3	Norway	2000	6	Netherlands	2009
	Sweden	2008		USA	2011
	Netherlands	2011		Netherlands	2010
	Canada	2009		Сінгапур	2005
	Netherlands	2009		USA	2010
	USA	2009		Ukraine	2011
	Ukraine	2008		7	Japan
4	Norway	2005	Japan		2011
	New Zealand	2008	Netherlands		2012
	Sweden	2011	USA		2009
	Canada	2012	Netherlands		2010
	Canada	2013	USA		2008
	Japan	2009	Ukraine		2012

Table 5. Simulation results.

#	Cluster	Average cluster point modeling error, %
1	Without clustering	26,19
2	1	0,36
3	2	0,58
4	3	1,2
5	4	0,62
6	5	0,02
7	6	0,96
8	7	0,98

5 Conclusions






Profiling clusterer agents allows grouping of observation points with predefined properties without first forming their classes. The modeling accuracy of such points is increasing. The efficiency of profiling of clusterer agents by the values of the weight coefficients of the variables of their models was experimentally confirmed. To confirm the homogeneity of the points grouped into clusters, the characteristic of the modeling results of these points was used.

References

1. Oldenderfer, M.S., Bleshfield, R.K.: *Klasternyy analiz*. In: *Faktornyy, diskriminantnyy i klasternyy analiz*: Translated from English by the editor Yenyukova I. S. *Finansy i statistika* Moscow (1989)
2. Mandel, I.D.: *Klasternyy analiz*. *Finansy i statistika*, Moscow (1988)
3. Gorban, A.N., Zinovyev, A.Yu.: Method of elastic maps and its applications in data visualization and data modeling. *J. Comput. Anticip. Syst.* **12**, 353–369 (2002)
4. Ayvazyan, S.A., Bukhshtaber, V.M., Yenyukov, I.S., Meshalkin, L.D.: *Prikladnaya statistika. Klassifikatsiya i snizheniye razmernosti*. *Finansy i statistika*, Moscow (1989)
5. Litinskiy, L.B., Romanov, D.Ye.: *Neyrosetevoy podkhod k zadache klasterizatsii ob'yektov*. In: *Neyroinformatika. Part 2*. MIFI, Moscow (2006)
6. Osipenko, V.V.: Induktivnyy algoritm klaster-analiza v instrumentarii sistemnykh informatsionno-analicheskikh issledovaniy. *Upravlyayushchiye sistemy i mashiny* (2), 26–32 (2013)
7. Golub, S.V., Burlyay, Í.V.: Pidvishchennya yefektivností klasterizatsií za rezul'tatami modelyuvannya v ínformatsiyniy tekhnologiií operativnogo pozhezhnogo monitoringu. *Sistemi obrobki ínformatsií*. **2**(118), 253–257(2014)
8. <http://unctadstat.unctad.org>. Accessed 10 Mar 2020
9. <http://ru.tradingeconomics.com>. Accessed 10 Mar 2020
10. www.ukrstat.gov.ua. Accessed 10 Mar 2020



Proof of Stake for Blockchain Based Distributed Intrusion Detecting System

Ivan Burmaka¹(✉) , Nikolai Stoianov² , Vitalii Lytvynov¹ ,
Mariia Dorosh¹ , and Svitlana Lytvyn¹ 

¹ Chernihiv National University of Technology, Chernihiv, Ukraine
Ivan.Bourmaka@stu.cn.ua, v.v.lytvynov.dept@gmail.com,
mariyaya5536@gmail.com, chdtu.fld@gmail.com

² Bulgarian Defence Institute Prof. Tsvetan Lazarov, Sofia, Bulgaria
n.stoianov@di.mod.bg

Abstract. One of the most important component of each corporate network is an intrusion detecting system. But it is a problem to find a mechanism for setting a trust between the nodes in a big distributed system. A blockchain can be used as such kind of mechanism, but most of working blockchains use cases related to cryptocurrencies. So there is no suitable consensus protocol for a permissionless blockchain, which can be used as a part of IDS. In this paper we propose a few modifications to proof of stake consensus protocol to adopt it for using with a blockchain based IDS. Also an agent based model was built to simulate a staking process for our modified consensus protocol. Modeling result will help us to evaluate performance of a blockchain based IDS in a normal state and in some critical situations and to find the most suitable consensus protocol parameters depending on a network size.

Keywords: Intrusion detecting system · Distributed IDS · Blockchain · Consensus protocol · Proof of Stake · Chain-based PoS

1 Introduction

Cybersecurity specialists are registering new more advanced and more complicated cyberattacks every year. So most of companies widely implement intrusion detecting systems to detect network intrusions in time and prevent their network infrastructure damage. Old intrusion detection systems, even the network, were based on a single machine, but in our time the amounts of traffic is very high, so to protect a big network the distributed and collaborative network intrusion detection systems are used. In big networks it is difficult to setup a trust between huge amounts of nodes, because some of the nodes can be modified and can sent some untrusted and irrelevant data, as a result all distributed nodes may work incorrectly. As a solution of this problem researchers proposed to use a blockchain technology to set up trusted data exchange between nodes [1]. Most of works

propose to use a permissioned blockchain, which means that a blockchain based system still needs trusted authority to manage blockchain network participants. But this makes the system hard to scale because of permissioned blockchain features. So the main our goal is to find a permissionless consensus protocol which can be modified for usage in distributed intrusion detecting system.

The paper is structured as follows. Section 1 is Introduction. Section 2 discusses the related works. Section 3 outlines the general characteristics of a proof-of-stake consensus protocol, it's pros and cons. Section 4 describes the proposed blockchain consensus protocol and experimental results for modeling a staking stability. Conclusions are provided in Sect. 5.

2 Related Work

A separate intrusion detection system usually does not have information about the network which it tries to protect. This leaves the opportunity for attackers to bypass IDS examination. In this case, there is a great need for a collaborative system or IDS network to leverage the detection performance of a single IDS [2]. Li et al. [3] figured out the fact, that the main problem of collaborative IDS is that most of distributed intrusion detection systems rely on the centralized data processing and a distributed part includes only some set of sensors. So they propose CIDS based on the emerging decentralized location and routing infrastructure. But this mechanism requires that all peers in the network must be trusted. One of the ways to solve a trust problem in decentralized system is to use a blockchain.

The name 'blockchain' stems from its technical structure—a chain of blocks. Each block is linked to the previous block with a cryptographic hash. A block is a data structure which allows storing a list of transactions. Transactions are created and exchanged by peers of the blockchain network and modify the state of the blockchain. As such, the transactions can exchange monetary amounts, but are not restricted to financial transactions only and for example allow executing an arbitrary code within so-called smart contracts [4]. How to apply blockchains in the field of intrusion detection is an interesting and important topic. Many studies have started researching in this area. Alexopoulos et al. [5] describe a framework of a blockchain-based CIDS, where they consider a set of raw alarms produced by each IDS as transactions in a blockchain. Then, all collaborating nodes employ a consensus protocol to ensure the validity of the transactions before delivering them in a block. This can guarantee the stored alerts are tamper resistant in the blockchain, but they did not implement and evaluate their method in practice. Li et al. also proposed a blockchain based solution for IoT environment [1]. The idea of this solution is to use a blockchain to setup data exchange between the signature based intrusion detecting nodes. But that solution is oriented on small IoT networks and has a few performance issues.

3 Proof of Stake as Consensus Protocol for IDS

If we take a look at the list of available blockchain consensus protocols for permissionless blockchains, we can see two major protocols—proof of work (PoW) and proof of stake (PoS). A main consideration regarding the operation of blockchain protocols which are based on PoW—such as bitcoin [6]—is the energy required for their execution. For example, generating a single block on the bitcoin blockchain requires a number of hashing operations exceeding 2^{60} , which results in striking energy demands. Indeed, early calculations indicated that the energy requirements of the protocol were comparable to that of a small country [7].

PoS attempts to solve the energy expenditure problem created by PoW. To do so, PoS replaces PoW's competition by randomly selecting stake-holders to append to the blockchain. So, rather than miners investing computational resources in order to participate in the leader election process, they instead run a process that randomly selects one of them proportionally to the stake that each possesses according to the current blockchain ledger. The simplest implementation of PoS, known as the Follow-The-Satoshi (FTS) algorithm, involves each blockchain branch selecting uniformly and randomly from the universe of native coins [8].

The owner of the selected coin receives the opportunity to append to the branch that selected his coin and simultaneously collect a block reward. There are four classes of PoS protocols: a chain-based PoS, a committee-based PoS, a BFT-based PoS, and a delegated PoS (DPoS) [9].

3.1 PoS Types

Chain-Based PoS. A chain-based PoS is an early PoS scheme proposed as an alternative block generation mechanism to PoW. It is within the framework of Nakamoto consensus in that the gossiping-style message passing, a block validation rule, the longest-chain rule, and probabilistic finality are preserved. A minter can solve the hashing puzzle only once for a clock tick. Since the hashing puzzle difficulty decreases with the minter's stake value, the expected number of hashing attempts for a minter to solve the puzzle can be significantly reduced if his stake value is high. The examples of the chain based PoS are Peercoin [10] and Nxt [11] cryptocurrencies. Their major difference is in the way the stake is valued. The stake value is initially proportional to the the stake quantity. Peercoin uses the coin age metric for stake valuation, which lets the value of a stake appreciate linearly with time since the deposit. At the end of a block cycle, the value of the winner's stake returns to its base value. The stake appreciation continues only for 90 days and stays flat since then. As a result, the chances of small stakeholders to generate a block are supplemented with time value. In comparison, Nxt does not appreciate the stake value continuously across block cycles. This is because the latter's coin age metric may lower the attack cost—attackers can just invest a small amount of stakes and keep attempting to generate blocks

until they succeed. Instead, Nxt requires the stake value appreciate only within one block cycle and be reset to the base value once the block cycle ends.

A chain-based PoS can tolerate up to 50% of all stakes being maliciously controlled (50% of all tokens in the network). If attackers control more than 50% of stakes, they can grow their malicious chain faster than the others and carry out a double-spending attack. As staking is recorded in the form of transaction scripts, the blockchain users can retrieve the staking records from which the consensus protocol can legally issue punishment to violators, such as nullifying stakes and disbaring the violators from participating in the future staking process.

Committee-Based PoS. As an alternative to a Chain-based PoS hashing puzzle mechanism, a committee-based PoS proposes to determine a committee of stakeholders based on their stakes and allowing the committee to generate blocks in turns. A secure multiparty computation (MPC) scheme is often used to derive such a committee in the distributed network. MPC is a genre of distributed computing in which multiple parties beginning with individual inputs will output the same result [12]. The MPC process in the committee-based PoS essentially realizes the functionality that takes in the current blockchain state which includes the stake values from all stakeholders, and outputs a pseudo-random sequence of stakeholders (we call it the leader sequence) which will subsequently populate the block-proposing committee. This leader sequence should be the same for all stakeholders and those with higher stake values may take up more spots in the sequence.

In spite of having an orderly block proposing scheme, committee-based PoS still adheres to the longest chain rule for probabilistic finality. So long as fewer than 50% stakes are held by the malicious party, the honest parties can safely maintain the longest chain. Meanwhile, the expansion of the committee may lead to a significant drop in the protocol performance and desynchronization of block proposal. The round-based committee election process with a predetermined round duration faces scalability problems, as large committee sizes may lead to never ending consensus cycles. This problem can be solved by extending the duration of a communication round. This however leads to a longer transaction conformation latency and lower throughput. A more straightforward approach is limiting the committee size by imposing a minimum stake requirement [9].

BFT-Based PoS. A chain-based PoS and committee-based PoS largely follow the Nakamoto consensus framework in which the longest-chainrule is still used to provide probabilistic finality of blocks. In comparison, a BFT-based PoS (or hybrid PoS-BFT) incorporates an extra layer of BFT consensus that provides a fast and deterministic block finalization. Algorithm 6 shows the general procedure of a BFT-based PoS for every participant. Block proposing can be done by any PoS mechanism (round-robin, committee-based, etc.) as long as it injects a stable flow of new blocks into the BFT consensus layer. As a result, the longest-chain rule can be safely replaced by the most-recent-stable-checkpoint rule for

determining the stable main chain. Popular BFT-based PoS blockchain protocols include Tendermint [13], Algorand [14], and Casper FFG. DPoS protocols such as EOSIO also use BFT consensus for a block finalization within delegates.

Security analysis BFT-based PoS's consensus fault tolerance varies among the three above-mentioned implementations. In Tendermint, although block proposers are determined based on PoS, all validators have the equal weight in the consensus process. Therefore Tendermint tolerates up to 1/3 of Byzantine validators. Since a typical BFT consensus protocol can incorporate a checkpointing mechanism to ensure deterministic finality of blocks, costless simulation attacks can be naturally avoided.

Delegated PoS (DPoS). A DPoS can be seen as a democratic form of committee-based PoS as the committee (consensus group) is chosen via public stake delegation. It is currently used by EOSIO [15] and Cosmos [16]. DPoS was designed to control the size of the consensus group so that the messaging overhead of the consensus protocol remains manageable. The members of the consensus group are also called delegates. The election of delegates is called the delegation process. In the actual case, the delegation process and the soliciting of votes may involve outside incentives. Generally, an aspiring delegate needs to attract enough votes from normal token holders. This is often accomplished by offering a popular application and building up the reputation through propaganda campaigns. By casting a vote to a delegate via a blockchain transaction, a token holder entrusts the delegate his own stake [15]. As a result, the delegate harvests the stake voting power from his voters. The illustration of the delegation process in DPoS: DPoS's incentive mechanism is designed to encourage honest delegation and consensus participation and to enforce transaction validation and consensus safety. Each delegate receives a daily vote reward proportionally to the votes he has. Once ascending to the consensus group, delegates also receive block rewards for validation work.

An assuming BFT is used by the consensus group for a block finalization, which is recommended since the group size is limited, a DPoS can tolerate 1/3 of delegates being malicious. For example, EOSIO can tolerate at most 6 out of 21 delegates being malicious. In the real world they may not wish to misbehave or collude at all, since all delegates have revealed their identities to voters and would be scrutinized for any misconduct.

3.2 Vulnerabilities of PoS

Although heralded as the most promising mechanism to replace a PoW, a PoS still faces several vulnerabilities.

- 1) Costless simulation: Costless simulation is a major vulnerability of non-BFT-based PoS schemes, especially a chain-based PoS in which a PoS is used to simulate the would be PoW process. Costless simulation literally means that any player can simulate any segment of a blockchain history at the cost of

no real work but speculation, as a PoS does not incur intensive computation while the blockchain records all staking history. This may give attackers shortcuts to fabricate an alternative blockchain. The four subsequent vulnerabilities, namely *nothing-at-stake*, *posterior corruption attack*, *long-range attack*, and *stake-grinding attack* are all based on costless simulation [9].

- 2) Nothing-at-stake: Nothing-at-stake is the first identified costless simulation problem that affects a chain-based PoS. It is also known as a “multi-bet” or “rational forking” problem. Unlike a PoW miner, a PoS minter needs a little extra effort to validate transactions and generate blocks on multiple competing chains simultaneously. Consequently if a significant fraction of nodes performs the “multi-bet” strategy, an attacker holding far less than 50% of tokens can mount a successful double spending attack [17].
- 3) Posterior corruption: Dubbed by Bentov [18] as a “bribing attack”. The key enabler of posterior corruption is the public availability of staking history on the blockchain, which includes stakeholders’ addresses and staking amounts. An attacker can attempt to corrupt the stakeholders who once possessed substantial stakes, and if he corrupt stakeholders with a significant portion of tokens, double spending will be possible.
- 4) Long-range attack: Coined by the Ethereum founder Vitalik Buterin [19]. It foresees that a small group of colluding attackers can regrow a longer valid chain that starts not long after the genesis block. Because if there are only a few stakeholders, the attackers can grow the malicious chain very fast and redo all the PoS blocks. BFT protocols use a checkpoint mechanism to ensure the finality of system agreements and safely discard older records.
- 5) Centralization risk: a PoS faces a similar wealth centralization risk as a PoW. In a PoS the minters can lawfully reinvest their profits into staking perpetually, which allows the one with a large sum of unused tokens become wealthier and eventually reach a monopoly status [9].

4 Proof of Stake Modification for IDS

So as we can see, the Proof of work consensus algorithm is not suitable for using in a distributed intrusion detecting system because of a high energy consumption. But the proof of stake can be used as a consensus protocol for a distributed intrusion detecting system. The simplest type of the proof of stake, chain-based, will represent an open IDS model, where any compatible node can join the network without any restrictions. The chain based PoS consensus is used only in a permissionless blockchain, so an IDS which is based on this type of consensus protocol will be fully decentralized. This will make the system more robust and reliable.

A classical chain-based PoS algorithm is not fully suitable for our purpose. As we can see in most of the cryptocurrencies the PoS consensus protocol works based on staking some amount of cryptocurrency. But this way is not suitable for us, because our transaction is not related to any currency. So the best alternative value to stake in our case will be the time, because it is easy to measure on the bases of block timestamps.

A block structure for our algorithm stays mostly common. A block has a header with a few common fields. First of all it is a hash of a previous block, which helps to link separate blocks into a chain. Another important field is a timestamp which is used to check a correctness of the block hash. In our case a timestamp is also used by nodes to measure their staked time. For simplification we do not need to store staked time in the blockchain, because we can just measure it for each node from “join the network” transaction and the blocks which were created by this node. One more field is the block difficulty which will be used to generate the next block. This value is calculated based on the current blockchain growing speed and in our case also based on current blocks queue in a memory pool. The most important field is a current block signature.

First of all a validator checks an eligibility to create a block (proof of eligibility) by using a hash function and comparing it with a target for a current validator. So the input values for our hash function will be *PublicKey* and *current block time*. The condition of generating a new block is shown in Eq. 1 which we get by modifying the standard PoS equation [20]:

$$\text{Hash}(\text{PubKey}_v, t_{\text{block}}) < T_b * d_{\text{current}} * t_{\text{stake}}, \quad (1)$$

where *PubKey_v* is a current validator node public key and *t_{block}* is a current block time, *T_b* is a base target which is constant, but can be modified by a block difficulty *d_{current}* and it is multiplied by a current validator stake *t_{stake}*. So, the higher stake value is—the higher probability that Eq. 1 will be true.

Forks in the chain for our algorithm will be solved in a common for a chain-based proof of stake way. Nodes think that the best chain from a few forks is the chain where the sum of targets for all blocks is the lowest (equals to bigger amount of work in a PoW consensus).

But as we can see above, simple time staking is a bit dangerous, because some malicious nodes can say offline to stake more time (without participating in a consensus process) and then use it to create an alternative chain. So we need some way to limit the time staking. The simplest solution here is to make staked time value constant after some value, as in Peercoin, but this will limit the stake value and malicious nodes will still have the highest possible value. Another solution is to use modular arithmetic—then the staked value will be limited by a module, and will be reset periodically during the time. This will be unfair in case of cryptocurrencies, where the winner node gets reward for creating a new block, because a module based stake resetting can lead to a skipping reward. But in case of IDS nodes do not get any reward for block creating, so this will only motivate the nodes to stay online as long as possible to use their stake more effective. Also limiting the highest stake value this way will help to decrease a number of stakeholders with a high stake value and to avoid a lot of forked blocks. Forked chains are normal for PoS based blockchain.

To evaluate a blockchain state with such kind of stake modification we built an agent based model, which simulates a work of blockchain nodes which use the time staking. Our simulation is a little simplified and does not use any transactions, because they are not so important for our current model. The

main goal of this model is to determine will the blockchain based network with such kind of stake limitation that has enough stake to validate all blocks in time.

So our model consists of the agents some of which were added to the network on network creation, another agents joined the network in the next few blocks. On every block creation one agent (validator) fully lost his stake after a block was created. In case of fork few agents can lose the stake (by resetting staked time), but when the fork is resolved and other blocks are discarded, other agents will get their stakes back. The agent can also lose a stake value if it becomes bigger than a limiting module.

The first simulation result is shown in Fig. 1. Here we chose a limit time in 500 s and block time in 20 s for a faster simulation (values in a real system will depend on the amounts of data and the nodes quantity). In this case a limit value is a little low, so that some nodes do not have possibility to use their stake before reaching the limiting module.

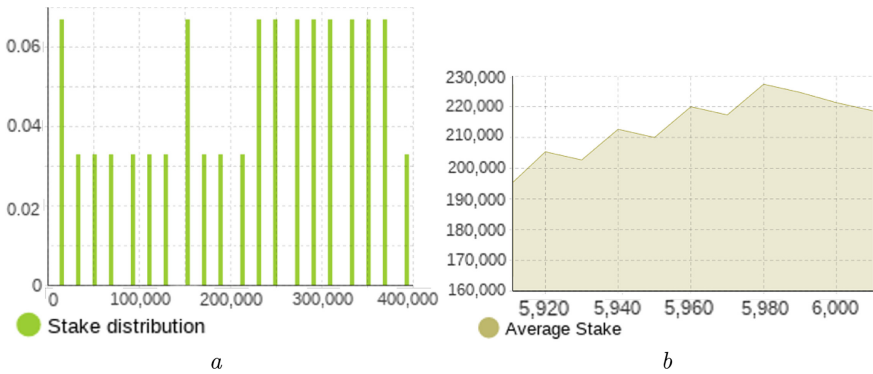


Fig. 1. Staking simulation with a low limit value: *a*—stake value distribution, *b*—average stake value

As we can see on the plot, a low stake limit can cause a higher range of average stake fluctuation and a distribution shows that the values are not distributed uniformly, but all ranges still have a good amount of values. This means that system will be still usable even if a lot of nodes will skip their staked values because of reaching a module limit.

In another simulation we increased a module limit to 1000 s, all other parameters were left unchanged. The result of the second simulation is shown in Fig. 2.

Here we can see that the fluctuations of the average stake value between the agents are much lower. The distribution in this experiment shows that the stake values in a network are distributed almost uniformly, which means that any time there are some nodes which have enough stake to validate the block. Also we can see here that the nodes do not reach the stake limit because all theirs stake is used to validate blocks.

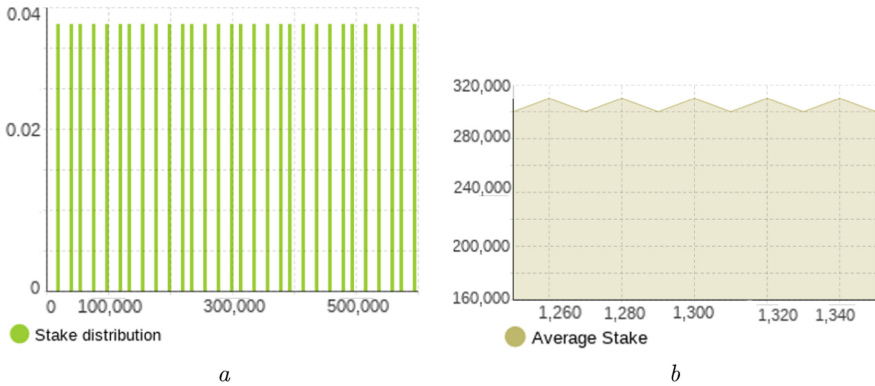


Fig. 2. Staking simulation with a high limit value: *a*—stake value distribution, *b*—average stake value

So, limiting a maximum stake value by a module gives us a good stake distribution which is suitable for using in a blockchain based IDS. Even if the limit is low and a lot of nodes are reset their stakes by reaching the limit, the system will still have enough stake resources to validate the blocks.

5 Conclusion and Future Work

This paper has proposed a modification of the standard proof of stake consensus algorithm for building a distributed and fully decentralized intrusion detecting system. A modified algorithm allows us to use a permissionless blockchain for building an IDS, which means that a blockchain network will not have any central node and will be easy to scale. An effective PoS usage here is achieved by using a time as a staked value. This will be especially important for using with a huge amount of nodes, for example big IoT networks which are hard to administrate.

There are a few directions for future work. First of all to design a transaction system for exchanging information between the nodes in a blockchain based intrusion detecting system, where the main problem is to achieve a small block size with the minimum information loss. Another major task is to test an intrusion detecting system with a PoS based blockchain in a real network and optimize possible bottlenecks to achieve a good performance.

Acknowledgement. This study is funded by the NATO SPS Project CyRADARS (Cyber Rapid Analysis for Defense Awareness of Real-time Situation), Project SPS G5286.



References

1. Li, W., Tug, S., Meng, W., Wang, Y.: Designing collaborative blockchained signature-based intrusion detection in IoT environments. *Future Gener. Comput. Syst.* **96**, 481–489 (2019). <http://www.sciencedirect.com/science/article/pii/S0167739X18327237>
2. Wu, Y.S., Foo, B., Mei, Y., Bagchi, S.: Collaborative intrusion detection system (CIDS): a framework for accurate and efficient IDS. In: *Proceedings of 19th Annual Computer Security Applications Conference*, pp. 234–244, December 2003
3. Li, Z., Chen, Y., Beach, A.: Towards scalable and robust distributed intrusion alert fusion with good load balancing. In: *Proceedings of the 2006 SIGCOMM Workshop on Large-Scale Attack Defense, LSAD 2006, Pisa, Italy*, pp. 115–122. ACM, New York (2006). <https://doi.org/10.1145/1162666.1162669>
4. Wüst, K., Gervais, A.: Do you need a blockchain? In: *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, pp. 45–54, June 2018
5. Alexopoulos, N., Vasilomanolakis, E., Ivanko, N.R., Muhlhauser, M.: Towards blockchain-based collaborative intrusion detection systems. In: D’Agostino, G., Scala, A. (eds.) *Critical Information Infrastructures Security, Lecture Notes in Computer Science*, pp. 107–118. Springer International Publishing, Cham (2018)
6. Bitcoin Developer Reference - Bitcoin. <https://bitcoin.org/en/developer-reference#remote-procedure-calls-rpcs>
7. O’Dwyer, K.J., Malone, D.: Bitcoin mining and its energy footprint (2014)
8. Xiao, Y., Zhang, N., Lou, W., Hou, Y.T.: A survey of distributed consensus protocols for blockchain networks. *IEEE Commun. Surv. Tutor.* **22**(2), 1432–1465 (2020)
9. Xiao, Y., Zhang, N., Lou, W., Hou, Y.T.: A Survey of distributed consensus protocols for blockchain networks. *IEEE Commun. Surv. Tutor.* 1 (2020). [arXiv: 1904.04098](https://arxiv.org/abs/1904.04098)
10. King, S., Nadal, S.: PPCoin: peer-to-peer crypto-currency with proof-of-stake. Self-Published Paper, 19 August 2012
11. Nxt Community: Nxt whitepaper revision 4 (2014). <https://www.dropbox.com/s/cbuwrorf672c0yy/NxtWhitepaperv122rev4.pdf>
12. Rabin, T., Ben-Or, M.: Verifiable secret sharing and multiparty protocols with honest majority. In: *Proceedings of the Twenty-First Annual ACM Symposium on Theory of computing, STOC 1989, Seattle, Washington, USA*, pp. 73–85. Association for Computing Machinery, February 1989. <https://doi.org/10.1145/73007.73014>
13. Kwon, J.: TenderMint: consensus without mining. Draft v. 0.6, Fall 1(11) (2014)
14. Gilad, Y., Hemo, R., Micali, S., Vlachos, G., Zeldovich, N.: Algorand: scaling byzantine agreements for cryptocurrencies. In: *Proceedings of the 26th Symposium on Operating Systems Principles*, pp. 51–68 (2017)
15. Larimer, D.: DPOS—pipelined byzantine fault tolerance, May 2018. <https://medium.com/eosio/dpos-bft-pipelined-byzantine-fault-tolerance-8a0634a270ba>, [libraryCatalog:medium.com](https://libRARYCatalog.medium.com)
16. Cosmos SDK documentation cosmos SDK. <https://docs.cosmos.network/master/cosmos-hub/validators/overview.html>
17. Martinez, J.: Understanding proof of stake: the nothing at stake theory (2018)
18. Mizrahi, I.B.C.L.A., Rosenfeld, M.: Proof of activity: extending bitcoin’s proof of work via proof of stake (2014)

19. Buterin, V.: Long-range attacks: the serious problem with adaptive proof of work (2018). Accessed 26 Dec 2014
20. Li, W., Andreina, S., Bohli, J.M., Karame, G.: Securing proof-of-stake blockchain protocols. In: Garcia-Alfaro, J., Navarro-Arribas, G., Hartenstein, H., Herrera-Joancomartí, J. (eds.) *Data Privacy Management, Cryptocurrencies and Blockchain Technology*. Lecture Notes in Computer Science, pp. 297–315. Springer International Publishing, Cham (2017)



Perfect Reconstruction Condition for Rational Wavelet Transform with Reducible Rational Dilation Factor

Oleg Chertov  and Volodymyr Malchykov  ^(✉)

Igor Sikorsky Kyiv Polytechnic Institute, Kiev, Ukraine
mavr2k@gmail.com

Abstract. Wavelet analysis is a modern and very powerful technique for effective analysis of data of different nature (time series, various kinds of multimedia data such as images, video and audio, and so on). Dyadic wavelet transforms are commonly used due to their simple and effective software implementation. Applying of non-dyadic wavelet transforms gives more precise separation of data features. The most general approach of non-dyadic wavelet transforms is the rational multiresolution analysis. Usually irreducible fraction is used as a value of dilation factor for rational wavelet transforms. But using of reducible dilation factor can improve detection of singularities. Property of perfect reconstruction of source data is required for any wavelet transform. It was proved that this condition is satisfied in the case of irreducible dilation factor. In this paper we will show that perfect reconstruction condition is also satisfied for the reducible dilation factor by the example of its value of $6/4$.

Keywords: Non-dyadic wavelet · Dilation factor · Perfect reconstruction

1 Introduction

1.1 Wavelet Transform as Linear Transformation of Signals

In engineering practice, different classes of transformations are used to study various signals of both natural and artificial origin: Fourier, Laplace, Walsh, Radon, etc. They are based on the classical idea of decomposing a complex object into such components that represent its essential features.

Wavelet transformation (WT) is a special type of linear transformation of signals and data which are the results of the processes and properties of both natural or artificial environments and objects. WT performs a time-frequency analysis, the results of which contain both the general frequency response of the signal (the distribution of the signal energy over the frequency components), as well as information about certain local coordinates at which a rapid change of the corresponding frequency occurs. This is achieved by the fact that the WT uses soliton-like functions as a basis, whose graph is in the form of small waves (wavelets), unlike the functions sine and cosine defined on the entire numerical axis, which are used in the Fourier transform [1, 2].

Wavelet transforms are used extensively in mathematical modeling of infectious diseases dynamics [3] like SARS or COVID-19, for detection of some click-fraud

attacks in Internet marketing [4], for ensuring individual and group data anonymity in area of information security [5], in modeling of human emotions [6], and in more classic applications to approximate features and compress large amounts of information with negligible losses (JPEG2000 format, some MPEG-4 video codecs) [7] or numerical simulation of PDEs [8].

A number of recent surveys provide insight into the widespread use of wavelet transformations in civil engineering (particularly for denoising, discontinuity detection, feature extraction, frequency identification, system modeling, and data compression) [9], for big data analysis [10], for non-stationary time-series analysis [11], in fusion of medical images [12, 13], or generally for image processing [14].

1.2 Dyadic and Non-dyadic Wavelet Transform

Let N be a real number and $\varphi(x)$ – square-integrable on real number line function; $\varphi(x)$ is called wavelet function if it satisfies the refinement equation

$$\varphi(x) = \sqrt{N} \cdot \sum_{n \in \mathbb{Z}} h_n \cdot \varphi(Nx - n). \tag{1}$$

Let $\psi(x)$ be a square-integrable on real number line function. It is called scaling function if it satisfies the equation

$$\psi(x) = \sqrt{N} \cdot \sum_{n \in \mathbb{Z}} g_n \cdot \varphi(Nx - n). \tag{2}$$

\mathbb{Z} indicates the set of integer numbers. Number N is called a dilation factor of WT. Dilation factor is a discrete wavelet transform characteristic. According to [15] we can take as a value of dilation factor any real number greater than one.

Discrete WT is called dyadic when dilation factor equals 2 and non-dyadic otherwise. Main advantage of dyadic WT is its simple and effective program implementation. But, non-dyadic transforms in contrast with dyadic WT allow to get more precise separation of analyzed data singularities. It was shown [16–19] that non-dyadic WT are more suitable for applying in some areas. At the present time there are various approaches to non-dyadic WT that were proposed, among others, by Bratteli and Jorgensen [20], Pollock and Cascio [19], Auscher [21], Baussard, Nicolier and Truchet [22], Feauveau [23]. Main properties and features of all mentioned approaches were shortly described by authors in their previous work [24]. Rational multiresolution analysis that was proposed in [21, 22] seems to be the most general approach for non-dyadic WT.

2 Problem Formulation

Perfect reconstruction is a very important property of WT. It guarantees the reversibility of the performed decomposition, i.e. an exact reconstruction of arbitrary input signal from the decomposition coefficients.

In all previously published articles devoted to the consideration of similar issues, the irreducibility of the fraction specifying the dilation factor was a prerequisite.

The purpose of this work is to show that this condition is optional, i.e. perfect reconstruction condition for the irreducible dilation factor will also be satisfied in the case of reducible dilation factor, e.g. for value 6/4.

3 Problem Solution

3.1 Filters Conditions and Functions

In the previous work [24] authors got the conditions for the coefficients of low-pass and high-pass filters for the reducible dilation factor 6/4:

$$\left\{ \begin{array}{lll} \sum_n h_n^0 \overline{h_{n-6l}^0} = \delta_{0l} & \sum_n h_n^1 \overline{h_{n-6l}^1} = \delta_{0l} & \sum_n h_n^1 \overline{h_{n-6l}^0} = 0 \\ \sum_n h_n^2 \overline{h_{n-6l}^2} = \delta_{0l} & \sum_n h_n^2 \overline{h_{n-6l}^0} = 0 & \sum_n h_n^2 \overline{h_{n-6l}^1} = 0 \\ \sum_n h_n^3 \overline{h_{n-6l}^3} = \delta_{0l} & \sum_n h_n^3 \overline{h_{n-6l}^0} = 0 & \sum_n h_n^3 \overline{h_{n-6l}^1} = 0 \\ \sum_n h_n^3 \overline{h_{n-6l}^2} = 0 & \sum_n g_n^1 \overline{g_{n-6l}^1} = \delta_{0l} & \sum_n g_n^2 \overline{g_{n-6l}^2} = \delta_{0l} \\ \sum_n g_n^1 \overline{h_{n-6l}^0} = 0 & \sum_n g_n^2 \overline{h_{n-6l}^0} = 0 & \sum_n g_n^1 \overline{h_{n-6l}^1} = 0 \\ \sum_n g_n^2 \overline{h_{n-6l}^1} = 0 & \sum_n g_n^1 \overline{h_{n-6l}^2} = 0 & \sum_n g_n^2 \overline{h_{n-6l}^2} = 0 \\ \sum_n g_n^1 \overline{h_{n-6l}^3} = 0 & \sum_n g_n^2 \overline{h_{n-6l}^3} = 0 & \sum_n g_n^2 \overline{g_{n-6l}^1} = 0 \end{array} \right. \quad (3)$$

where $\{h_n^0\}$, $\{h_n^1\}$, $\{h_n^2\}$, $\{h_n^3\}$ are the coefficients of the corresponding low-pass filters, and $\{g_n^1\}$, $\{g_n^2\}$ are the coefficients of the corresponding high-pass filters.

Additionally, set of auxiliary functions

$$m_0^0(\omega), \quad m_1^0(\omega), \quad m_2^0(\omega), \quad m_0^3(\omega), \quad m_1(\omega), \quad m_2(\omega)$$

was introduced. These functions were defined by the next expressions

$$\left\{ \begin{array}{l} m_o^k(\omega) = \sqrt{\frac{4}{6}} \sum_n h_n^k e^{-in\omega}, \quad k = 0, 1, 2, 3 \\ m_l(\omega) = \sqrt{\frac{4}{6}} \sum_n g_n^l e^{-in\omega}, \quad l = 1, 2 \end{array} \right. \quad (4)$$

and they satisfy equations

$$\left\{ \begin{array}{l} \widehat{\varphi}(\omega) e^{-ik\omega} = m_0^k(\frac{4}{6}\omega) \cdot \widehat{\varphi}(\frac{4}{6}\omega), \quad k = 0, 1, 2, 3 \\ \widehat{\psi}_l(\omega) = m_l(\frac{4}{6}\omega) \cdot \widehat{\psi}_l(\frac{4}{6}\omega), \quad l = 1, 2 \end{array} \right. \quad (5)$$

where $\widehat{\varphi}(\omega)$ is the Fourier image of scaling function, $\widehat{\psi}_1(\omega)$ and $\widehat{\psi}_2(\omega)$ are the Fourier images of corresponding two wavelet functions.

3.2 Perfect Reconstruction

Li [25] proved that in the case of irreducible dilation factor p/q the necessary and sufficient condition for perfect reconstruction can be formulated as

$$\mathbf{M}^*(\omega) \cdot \mathbf{M}(\omega) = q \cdot \mathbf{I}_{p \times p} \quad (6)$$

where $\mathbf{I}_{p \times p}$ is a unit matrix of dimension p , matrix $\mathbf{M}(\omega)$ is defined according to

$$\mathbf{M}(\omega) = \begin{pmatrix} m_0^0(\omega) & m_0^0\left(\omega + \frac{\pi q}{p}\right) & \dots & m_0^0\left(\omega + \frac{\pi q(p-1)}{p}\right) \\ \vdots & \vdots & \vdots & \vdots \\ m_0^{q-1}(\omega) & m_0^{q-1}\left(\omega + \frac{\pi q}{p}\right) & \dots & m_0^{q-1}\left(\omega + \frac{\pi q(p-1)}{p}\right) \\ m_1(\omega) & m_1\left(\omega + \frac{\pi q}{p}\right) & \dots & m_1\left(\omega + \frac{\pi q(p-1)}{p}\right) \\ \vdots & \vdots & \vdots & \vdots \\ m_{p-q}(\omega) & m_{p-q}\left(\omega + \frac{\pi q}{p}\right) & \dots & m_{p-q}\left(\omega + \frac{\pi q(p-1)}{p}\right) \end{pmatrix}^T \quad (7)$$

and $\mathbf{M}^*(\omega)$ is the complex conjugate of the transpose of $\mathbf{M}(\omega)$.

So, now let's prove that condition (6) will also be satisfied in the case of irreducible dilation factor, particularly $6/4$. Matrix $\mathbf{M}(\omega)$ in this case will be

$$\begin{pmatrix} m_0^0(\omega) & m_0^1(\omega) & m_0^2(\omega) & m_0^3(\omega) & m_1(\omega) & m_2(\omega) \\ m_0^0\left(\omega + \frac{\pi}{3}\right) & m_0^1\left(\omega + \frac{\pi}{3}\right) & m_0^2\left(\omega + \frac{\pi}{3}\right) & m_0^3\left(\omega + \frac{\pi}{3}\right) & m_1\left(\omega + \frac{\pi}{3}\right) & m_2\left(\omega + \frac{\pi}{3}\right) \\ m_0^0\left(\omega + \frac{2\pi}{3}\right) & m_0^1\left(\omega + \frac{2\pi}{3}\right) & m_0^2\left(\omega + \frac{2\pi}{3}\right) & m_0^3\left(\omega + \frac{2\pi}{3}\right) & m_1\left(\omega + \frac{2\pi}{3}\right) & m_2\left(\omega + \frac{2\pi}{3}\right) \\ m_0^0\left(\omega + \frac{3\pi}{3}\right) & m_0^1\left(\omega + \frac{3\pi}{3}\right) & m_0^2\left(\omega + \frac{3\pi}{3}\right) & m_0^3\left(\omega + \frac{3\pi}{3}\right) & m_1\left(\omega + \frac{3\pi}{3}\right) & m_2\left(\omega + \frac{3\pi}{3}\right) \\ m_0^0\left(\omega + \frac{4\pi}{3}\right) & m_0^1\left(\omega + \frac{4\pi}{3}\right) & m_0^2\left(\omega + \frac{4\pi}{3}\right) & m_0^3\left(\omega + \frac{4\pi}{3}\right) & m_1\left(\omega + \frac{4\pi}{3}\right) & m_2\left(\omega + \frac{4\pi}{3}\right) \\ m_0^0\left(\omega + \frac{5\pi}{3}\right) & m_0^1\left(\omega + \frac{5\pi}{3}\right) & m_0^2\left(\omega + \frac{5\pi}{3}\right) & m_0^3\left(\omega + \frac{5\pi}{3}\right) & m_1\left(\omega + \frac{5\pi}{3}\right) & m_2\left(\omega + \frac{5\pi}{3}\right) \end{pmatrix} \quad (8)$$

where functions

$$m_0^0(\omega), \quad m_0^1(\omega), \quad m_0^2(\omega), \quad m_0^3(\omega), \quad m_1(\omega), \quad m_2(\omega)$$

are defined according to (4). Let's define matrix \mathbf{A} as

$$\mathbf{A} = \mathbf{M}^*(\omega) \cdot \mathbf{M}(\omega) \quad (9)$$

and calculate elements of this matrix. Value of the element a_{11} is

$$\begin{aligned} a_{11} = & \overline{m_0^0(\omega)} \cdot m_0^0(\omega) + \overline{m_0^0\left(\omega + \frac{\pi}{3}\right)} \cdot m_0^0\left(\omega + \frac{\pi}{3}\right) + \overline{m_0^0\left(\omega + \frac{2\pi}{3}\right)} \cdot m_0^0\left(\omega + \frac{2\pi}{3}\right) \\ & + \overline{m_0^0\left(\omega + \frac{3\pi}{3}\right)} \cdot m_0^0\left(\omega + \frac{3\pi}{3}\right) + \overline{m_0^0\left(\omega + \frac{4\pi}{3}\right)} \cdot m_0^0\left(\omega + \frac{4\pi}{3}\right) \\ & + \overline{m_0^0\left(\omega + \frac{5\pi}{3}\right)} \cdot m_0^0\left(\omega + \frac{5\pi}{3}\right). \end{aligned}$$

After substituting expressions (4) we get

$$\begin{aligned}
 a_{11} = & \sqrt{\frac{4}{6}} \sum_k \overline{h_k^0} e^{-ik\omega} \cdot \sqrt{\frac{4}{6}} \sum_n h_n^0 e^{-in\omega} + \sqrt{\frac{4}{6}} \sum_k \overline{h_k^0} e^{-ik(\omega + \frac{\pi}{3})} \cdot \sqrt{\frac{4}{6}} \sum_n h_n^0 e^{-in(\omega + \frac{\pi}{3})} \\
 & + \sqrt{\frac{4}{6}} \sum_k \overline{h_k^0} e^{-ik(\omega + \frac{2\pi}{3})} \cdot \sqrt{\frac{4}{6}} \sum_n h_n^0 e^{-in(\omega + \frac{2\pi}{3})} + \sqrt{\frac{4}{6}} \sum_k \overline{h_k^0} e^{-ik(\omega + \frac{3\pi}{3})} \\
 & \times \sqrt{\frac{4}{6}} \sum_n h_n^0 e^{-in(\omega + \frac{3\pi}{3})} + \sqrt{\frac{4}{6}} \sum_k \overline{h_k^0} e^{-ik(\omega + \frac{4\pi}{3})} \cdot \sqrt{\frac{4}{6}} \sum_n h_n^0 e^{-in(\omega + \frac{4\pi}{3})} \\
 & + \sqrt{\frac{4}{6}} \sum_k \overline{h_k^0} e^{-ik(\omega + \frac{5\pi}{3})} \cdot \sqrt{\frac{4}{6}} \sum_n h_n^0 e^{-in(\omega + \frac{5\pi}{3})}.
 \end{aligned}$$

By performing opening of parentheses and multiplying

$$\begin{aligned}
 a_{11} = & \frac{4}{6} \sum_k \sum_n \overline{h_k^0} h_n^0 e^{-i(n-k)\omega} + \frac{4}{6} \sum_k \sum_n \overline{h_k^0} h_n^0 e^{-i(n-k)(\omega + \frac{\pi}{3})} \\
 & + \frac{4}{6} \sum_k \sum_n \overline{h_k^0} h_n^0 e^{-i(n-k)(\omega + \frac{2\pi}{3})} + \frac{4}{6} \sum_k \sum_n \overline{h_k^0} h_n^0 e^{-i(n-k)(\omega + \frac{3\pi}{3})} \\
 & + \frac{4}{6} \sum_k \sum_n \overline{h_k^0} h_n^0 e^{-i(n-k)(\omega + \frac{4\pi}{3})} + \frac{4}{6} \sum_k \sum_n \overline{h_k^0} h_n^0 e^{-i(n-k)(\omega + \frac{5\pi}{3})}.
 \end{aligned}$$

or, finally

$$a_{11} = \frac{4}{6} \sum_k \sum_n \overline{h_k^0} h_n^0 e^{-i(n-k)\omega} \left(1 + e^{-\frac{i(n-k)\pi}{3}} + e^{-\frac{i(n-k)2\pi}{3}} + e^{-\frac{i(n-k)3\pi}{3}} + e^{-\frac{i(n-k)4\pi}{3}} + e^{-\frac{i(n-k)5\pi}{3}} \right).$$

In order to simplify this expression, we will consider separately its term in parentheses with n-k replaced by m

$$u(m) = 1 + e^{-\frac{i\pi m}{3}} + e^{-\frac{i2\pi m}{3}} + e^{-\frac{i3\pi m}{3}} + e^{-\frac{i4\pi m}{3}} + e^{-\frac{i5\pi m}{3}}$$

and evaluate it for the different values of m:

$$u(0) = 6 \quad u(1) = u(2) = u(3) = u(4) = u(5) = 0 \quad u(6) = 6 \quad u(7) = 0 \quad \dots$$

So, we find that

$$u(m) = \begin{cases} 6, & m = 6l, l \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases}$$

Now we can write a_{11} as

$$a_{11} = \frac{4}{6} \sum_n \sum_{k=n-6l} 6 \overline{h_k^0} h_n^0 e^{-i(n-k)\omega} = 4 \sum_l \sum_n \overline{h_{n-6l}^0} h_n^0 e^{-6il\omega}$$

that, after taking into account conditions (3), will give us $a_{11} = 4$.

It can be shown in a similar way that all other diagonal elements of matrix (9) are also equal to 4 and non-diagonal elements of this matrix are zeros. Thus, matrix **A** is

$$\mathbf{A} = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix} = 4 \cdot \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = 4 \cdot \mathbf{I}_{6 \times 6}$$

that corresponds to the condition (6). So, it is satisfied for the chosen value of reducible dilation factor.

4 Conclusions

In this article authors showed that for the reducible value $6/4$ of rational dilation factor perfect reconstruction conditions are satisfied. Further researches will deal with generalization of these results to the case of arbitrary reducible dilation factor p/q .

The result obtained is not only of purely theoretical significance. Now it becomes possible to use the dilation factors equivalent in terms of perfect reconstruction of the original signal, for example, $3/2$, $6/4$, $9/6$ or, say, $39/26$. In practice, the use of different dilation factors may make it possible to better isolate the desired features of the original signal (time series, images, etc.).

Acknowledgement. This research was supported by CyRADARS project (SPS G5286 “Cyber Rapid Analysis for Defense Awareness of Real-time Situation”) in the frame of the NATO Science for Peace and Security program.







References

1. Mallat, S.: A Wavelet Tour of Signal Processing, 3rd edn. Academic Press, Cambridge (2009)
2. Chui, K.: Wavelets: A Tutorial in Theory and Applications. Academic Press, Cambridge (1992)
3. Choisy, M., Guégan, J.-F., Rohani, P.: Mathematical modeling of infectious diseases dynamics. In: Tibayrenc, M. (ed.) Encyclopedia of Infectious Diseases, pp. 379–404 (2006)
4. Chertov, O., Malchykov, V., Pavlov, D.: Non-dyadic wavelets for detection of some click-fraud attacks. In: 2010 International Conference on Signals and Electronic Circuits (ICSES), pp. 401–404. (2010)
5. Chertov, O., Tavrov, D.: Providing group anonymity using wavelet transform. In: MacKinnon, L.M. (ed.) BNCOD 2010. LNCS, vol. 6121, pp. 25–36. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-25704-9_5
6. Islam, M., Ahmad, M., Yusuf, M.S.U., Ahmed, T.: Mathematical modeling of human emotions using sub-band coefficients of wavelet analysis. In: 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), pp. 1–6. Dhaka (2015)

7. Adams, M., Ward, R.: Wavelet transforms in the JPEG-2000 standard. In: 2001 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, vol. 1, pp. 160–163. IEEE, Canada (2001)
8. Yousefi, H., Noorzad, A.: Wavelet based simulation of elastic wave propagation. In: Zheng, Y. (ed.) *Wave Propagation Theories and Applications*. IntechOpen (2013)
9. Kim, B., Jeong, H., Kim, H., Han, B.: Exploring wavelet applications in civil engineering. *KSCE J. Civil Eng.* **21**(4), 1076–1086 (2017). <https://doi.org/10.1007/s12205-016-0933-3>
10. Zhou, W., Li, J., Shah, F.: Retrospect and prospect of wavelet and big data: a survey. In: 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), pp. 307–310. Chengdu (2016)
11. Rhif, M., Ben Abbes, A., Farah, I.R., Martinez, B., Sang, Y.: Wavelet transform application for/in non-stationary time-series analysis: a review. *Appl. Sci.* **9**(7), 1345–1366 (2019)
12. Srinivasan, K., Thyagarajan, K.K.: Survey on wavelet based image fusion techniques. *Int. J. Recent Innov. Trends Comput. Commun.* **4**(5), 466–473 (2016)
13. Yadav, S.P., Yadav, S.: Fusion of medical images using a wavelet methodology: a survey. *IEIE Trans. Smart Process. Comput.* **8**(4), 265–271 (2019)
14. Saha, M., Naskar, M.K., Chatterji, B.N.: Advanced wavelet transform for image processing —a survey. In: Mandal, J.K., Bhattacharya, K., Majumdar, I., Mandal, S. (eds.) *Information, Photonics and Communication. LNNS*, vol. 79, pp. 185–194. Springer, Singapore (2020). https://doi.org/10.1007/978-981-32-9453-0_19
15. Daubechies, I.: *Ten Lectures on Wavelets*. SIAM, Philadelphia (1992)
16. Hur, Y., Zheng, F.: Prime coset sum: a systematic method for designing multi-D wavelet filter banks with fast algorithms. *IEEE Trans. Inf. Theory* **62**(11), 6580–6593 (2016)
17. Xiong, R., Xu, J., Wu, F.: A lifting-based wavelet transform supporting non-dyadic spatial scalability. In: 2006 IEEE International Conference on Image Processing, pp. 1861–1864 (2006)
18. Gupta, C., Lakshminarayan, C., Wang, S., Mehta, A.: Non-dyadic haar wavelets for streaming and sensor data. In: 2010 IEEE 26th International Conference on Data Engineering, pp. 569–580 (2010)
19. Pollock, D.S.G., Cascio, I.L.: Non-dyadic wavelet analysis. In: Kontoghiorghes, E.J., Gatu, C. (eds.) *Optimization, Econometric and Financial Analysis: Advances In Computational Management Science*, vol. 9, pp. 167–204 (2007)
20. Bratteli, O., Jorgensen, P.E.T.: Isometrics, shifts, Cuntz algebras and multiresolution wavelet analysis of scale N . *Integr. Eqn. Oper. Theory* **28**(4), 382–443 (1997)
21. Auscher, P.: Wavelet bases for $L^2(\mathbb{R})$ with rational dilation factor. In: Ruskai, M.B., et al. (eds.) *Wavelets and their Applications*. Jones and Barlett (1992)
22. Baussard, A., Nicolier, F., Truchetet, F.: Rational multiresolution analysis and fast wavelet transform: application to wavelet shrinkage denoising. *Signal Process.* **84**(10), 1735–1747 (2004)
23. Feauveau, J.-C.: Analyse multiresolution avec un facteur de résolution $\sqrt{2}$. *J. Traitement du Signal* **7**(2), 117–128 (1990)
24. Chertov, O., Malchykov, V.: Rational wavelet transform with reducible dilation factor. In: *Selected Papers of the XIX International Scientific and Practical Conference “Information Technologies and Security” (ITS-2019)*, pp. 146–158 (2020)
25. Li, Z.: Orthonormal wavelet bases with rational dilation factor based on MRA. In: 7th International Congress on Image and Signal Processing, pp. 1146–1150. IEEE, China (2014)



Principles of Adaptive Corporate Network Security Management

Vitalii Lytvynov¹ , Alla Hrebennyk² , Elena Trunova¹  ,
Igor Skiter¹ , and Yurii Lysetskyi¹ 

¹ Chernihiv National University of Technology, 95 Shevchenko Street,
Chernihiv 14035, Ukraine

e.trunova@gmail.com, skiteris@ukr.net,
Yurii.Lysetskyi@snt.ua

² Institute of Mathematical Machines and Systems Problems,
42 Ac. Glushkov Ave., Kiev 03680, Ukraine
grebennik.alla@gmail.com

Abstract. The problems of realizing attack detection processes in real time justify the use of a set of software and hardware that acts as an attack detection system (ADS) in a corporate computer network. Recently, the traditional individual approach to protecting a corporate network, based on the processing of its network traffic, is increasingly being improved by collective protection methods. In this case, the prevention and counteraction to unauthorized intrusions is based on active reconnaissance, processing large amounts of information stored and circulating on the Internet, creating of industry and national cyber security centers. This article considers the approach to the establishment of ADS, which use Internet resources of social networks, information from DB of international and national defense centers and also prehistory of computer corporate network attacks. It allows making forecasts of the danger state of the external environment relatively to the corporate computer network on the basis of which it is proposed to use the principles of adaptive security management of a computer network. The effective using of adaptive management, reducing the time for detecting threats, increasing productivity and optimizing the load of ADS components requires the selection of operational management procedures and tuning procedures for ADS analyzers that must operate at various computing capacities. This gives an opportunity to justify the functionality of the protection centers of corporate computer networks that can take on the tasks of configuring ADS analyzers of subordinate corporate networks connected by an integration information bus.

Keywords: ADS · Attack source · Computer network

1 Introduction

Nowadays, despite of the efforts of specialists, amount of computer attacks and many of their varieties, captured by official agencies, increase from year to year. Commercial operations, business interests, confidential information and so on first of all constitute the object of injurious acts. That is why a solution to the problem of security for

corporate networks activities in conditionals of global informational environment remains extremely important for the long time.

It is not accident that special attention is paid to the development of protection methods against computer attacks and related software.

Today, a typical answer to the question of medium-sized corporate nets security is a situation, where each corporate network, its resources and software, is protected from computer attacks by its own attack detection system (ADS) (Fig. 1) [10].

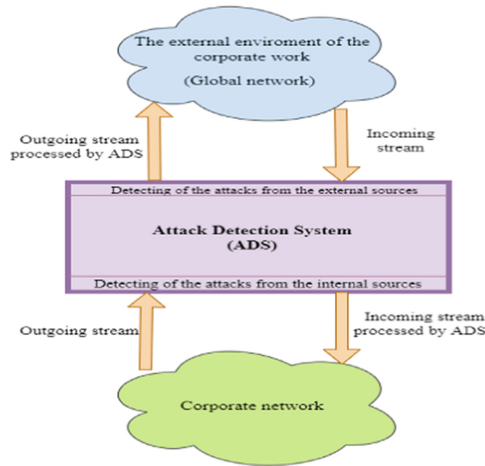


Fig. 1. Detection system of attacks on corporate networks S.

2 Tasks Solved by ADS

Typical ADS solves the following tasks [5]: analyzes the behavior of the corporate network in order to identify deviations from the standard behavior; determines whether the non-standard behavior is caused by an attack, or it can be classified as random event; determines the type of attack, if the attack is of a known type; gives recommendations on protection against attacks.

These tasks can be solved on the basis of information from network packet threads coming into the corporate computer network from the external environment (global network), as well as information from the event logs taking place in the nodes of the corporate net [2].

Accordingly, the conceptual model of ADS subject area is presented in Fig. 2 [9].

Traditionally, the architecture of the attack detection system includes (Fig. 3): Sensors for collecting primary information; Storage of primary information; Analyzers of corporate network behavior; Attack analyzers; Basic definitions.

ADS is defined as a means of protection of corporate networks from unauthorized access.

In the simplest case, a corporate computer network is a peer-to-peer local network that solves the tasks of a corporation or institution.

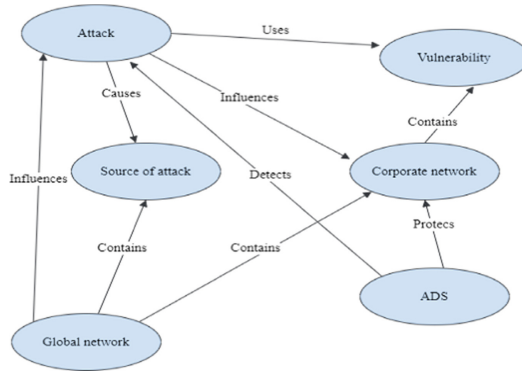


Fig. 2. Conceptual model of ADS subject area.

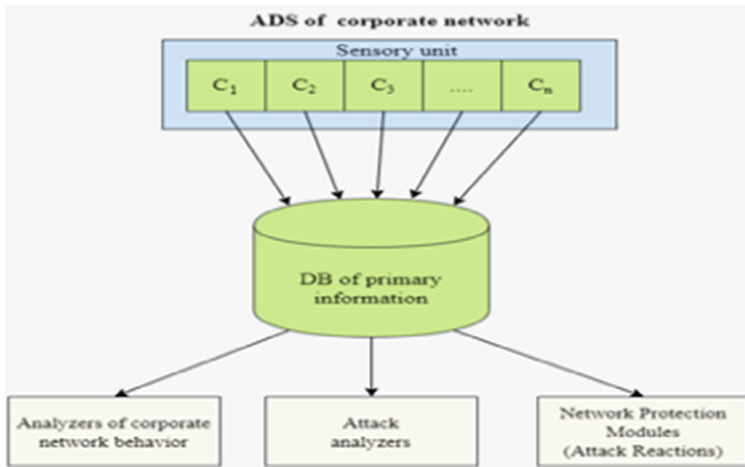


Fig. 3. Typical ADS architecture.

In general, it represents a distributed association of local networks consisting of hundreds, or even thousands of computers, intended to solve the problems of a corporation $S = \{S_1, S_2, \dots, S_n\}$.

An example of a corporate network is the association of computer networks of the fuel and energy complex of the country, consisting of computer networks of enterprises that have automated process control systems, networks for control centers of production associations, a network of the information system of the ministry.

We introduce a number of definitions related to the analysis of the corporate computer network security.

$U_s = \{u_{1s}, u_{2s}, \dots, u_{ks}\}$ – the set of types of corporate network vulnerabilities S .

Vulnerability can be used in case of unauthorized access to net computing resources during the attack on the network S .

Network vulnerabilities arise at the all stages its life cycle: projecting, implementation, configuration. In process of functioning of the network, they have nonstationary character and first of all depend on network administrator actions. Therefore, it is reasonable to consider the general network vulnerability in relation to time.

$U_s(t) = \langle u_{1s}(t), u_{2s}(t), \dots, u_{ks}(t) \rangle$ – state of vulnerability of corporate network S at time t , considered as a vector of network vulnerability indicators at the current time t , where $u_{is}(t) \in \{0, 1\}$ – the i -th vulnerability indicator at time t for corporate network S , determinate the presence or absence of a specified type of vulnerability in the network at the current time.

Let $A = \{A_1, A_2, \dots, A_n\}$ be the set of attack types, where A_i – is an attack of the i -th type, $A_i = \langle n_i, \beta_i, \gamma_i \rangle$, n_i – name of attack type, β_i – network vulnerability, used by the attack, γ_i – training sequence that is used in order to configure the network to recognize the attack A_i .

Attack A_i can be implemented, if $\beta_i \in U_s(t)$.

The level of danger state for the corporate network S from the external environment $A\Delta^{ee}(t)$ is considered as an ordered set of frequencies (probabilities) for the types of attacks on the network S by the time t for a given period of time Δ , that is:

$$A\Delta^{ee}(t) = \langle \langle A_1, p_1 \rangle, \langle A_2, p_2 \rangle, \dots, \langle A_n, p_n \rangle, t \rangle \quad (1)$$

where p_i – is the frequency (probability) of an attack of the i -th type during the time interval $[t - \Delta, t]$,

t – is the current time,

$A\Delta^{ee}(t)$ – is the set of potentially possible ($p_i \geq p_{thr}$ threshold value) types of attacks on the network S in the interval $[t - \Delta, t]$,

P_{thr} – is the threshold probability of an attack, accepted to be included in the list of possible attacks.

The indicator $A\Delta^{ee}(t)$ can be called the potential threat to the network S from the external environment at the moment of time t .

Let $M_S = \{m_1, m_2, \dots, m_k\}$ be the set of modules-analyzers of attacks in the ADS for the corporate network S , $A_{mi}(t)$ – is the set of types of attacks on the network S , which can be detected by the module-analyzer m_i by the time t during the interval Δt , $A_M(t)$ – set of types of attacks that can be detected by ADS of corporate network S by the time t during the interval Δt .

3 Basic Definitions

Obviously, for successful work of the ADS, it is required that for each moment of time t the condition $A\Delta^{ee}(t) \leq A_M(t)$ is fulfilled [4].

The attack analyzer acts as a filter, detecting and not allowing an attack from the list of potential threats of the external environment. For that it should be configured on finding of this type of attacks (Fig. 4).

Now consider the features of the operation algorithms of the ADS analyzers.

For ADS, a rather extensive series of algorithms for detection of non-standard network behavior and net attack detection has been developed [8].

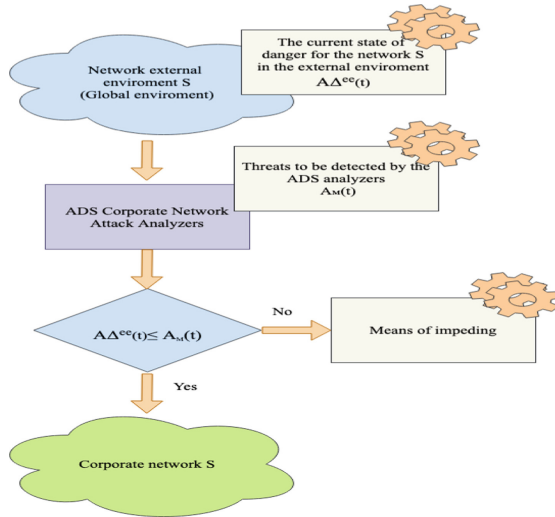


Fig. 4. Filtration of attacks with ADS analyzers.

These algorithms are divided into the following classes: statistical algorithms; signature (based on automata theories) algorithms; algorithms based on neural and immune systems; production type algorithms (based on clear or fuzzy rules); algorithms based on genetic approaches, etc.

They can be divided into two large groups: algorithms that effectively determine attacks or non-standard behavior within a specified range of attack types, algorithms that have learning properties (configured to effectively determination the specified types of attacks).

The main problems of their use in ADS are: 1) the presence of a large number of errors of the 1st and 2nd type (when the algorithm does not detect a valid attack on the network, or when the algorithm classifies the normal behavior of the net as an attack); 2) violation of the time constraints allocated to the determination of non-standard behavior or attacks (inefficient of the use of ADS).

Analyzing the first problem, it can be concluded that the main reason for this is the inconsistency between the current state of danger of the external environment of the network S ($R^{ee}(t)$) and the current recognition capabilities of the ADS analyzers ($A_M(t)$) [6]. Visual the configuration of the external environment danger and the analyzer's detecting capabilities should be regarded in different options (Fig. 5): a) the analyzer may make mistakes, b) the analyzer does not allow errors, c) the analyzer does not perform its functions.

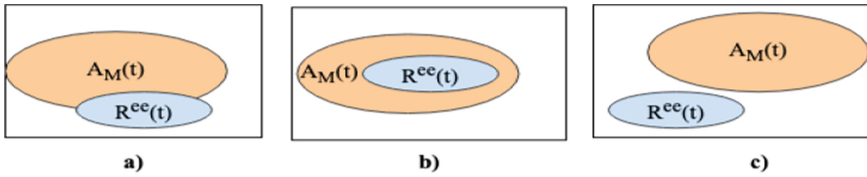


Fig. 5. Options for inconsistencies between the current state of danger of the external environment of the network S and the current recognition capabilities of the ADS analyzers.

4 Use of Adaptive ADS Analyzers

One of the ways to solve the first problem is to use custom (adaptive) modules of recognition algorithms that approximate their area of recognition capabilities to the characteristics of the current state of the external environment danger, that is, the use of the principles of adaptive security management of the corporate network.

To implement this approach, it is preferable to use algorithms for recognition of attacks and non-standard behavior in the analyzers based on the trained neural and immune systems, parametric production-type algorithms with clear and fuzzy rules, recognition algorithms based on genetic approaches.

The higher the degree of adaptive properties, algorithms for detection of non-standard behavior and attacks using neural networks and immune systems, the longer process of learning analyzers is. As a result of this disadvantage, customizing of algorithms takes quite a long time, therefore, this process may not be completed in the operational control cycle, which takes time Δt (Fig. 6).

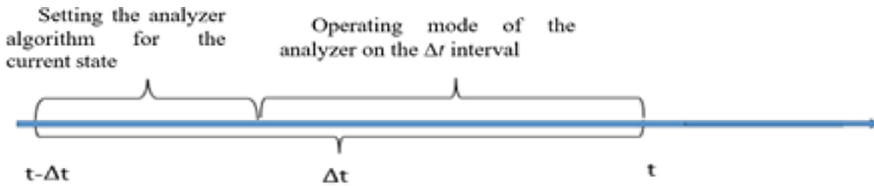


Fig. 6. Existing security operative management cycle.

5 Adaptive Corporate Network Security Management

A solution to the second problem can be found along the way of:

- involvement of the settings of a wider information from the external environment of the corporate network S and its subnets in the analysis, in order to predict the state of danger;
- prediction of the state of danger for a given time interval, during which it is possible to collect the necessary information and calculate the settings of the modules of the ADS analyzers that are most appropriate for this state of danger;

- exclusion of procedures for setting up recognition algorithms from the operational management;
- division of the analysis process into the analysis for analyzer settings and analysis for the rapid detection of attacks and non-standard behavior, allocating different computational capacity for this;
- using the methods of collective protection of agents, which are presented by the subnetworks of the corporate network S .

Consider the techniques above.

Prediction of the state of danger requires the use of a wide range of information.

If previously ADS analyzers used only information from event logs and data from network packets, in our case, it is proposed to expand the space of information retrieval contained in the global network, including websites and portals, social network information, databases of national and international cybersecurity centers and etc.

One of the most important steps in this security management scheme is finding and ranking potential sources of attacks. Until recently, the solution to this problem did not seem possible because despite the presence of sender addresses in the packet headers, changing them (setting false addresses) is not a problem. However, the latest results in mathematical linguistics and decision theory make this problem solvable.

Usually, indicators (signs) and methods are used as evidence, such as:

- the place of registration of IP-addresses and domains that participate in the attack or provide the infrastructure for the implementation of the attack;
- tracing of the attack to its source or at least localization of the area in which the source is located;
- time parameters;
- analysis of the program code in which you can find comments, links to sites, domains, IP-addresses that take part in the attack;
- style of programmers and school of programming;
- stylometry, allowing to determine the style of the language in the comments or related texts;
- deceptive systems or honeypot/honeynet;
- operative development;
- analysis of activity on forums and on social media;
- competitive exploring methods, etc.

Solution to this problem requires extensive involvement of Internet resources, as well as information about potential attackers and their inclination to use those or other types of attacks collected in national and international computer network security centers.

Thus, for each potential source of $i \in I$ (where I is the set sources of the attacks) attacks on the S network, many preferences are formed on the use of attack types, as well as information about their activity.

At the same time, determining the predicted state of danger of the external environment of a corporate network can be done by searching and integrating the most likely sources of attacks. And if there is information about the activity of the sources,

building a predictive dynamic model M_{pred} predictive of the environmental danger level for the corporate network S (Fig. 7).

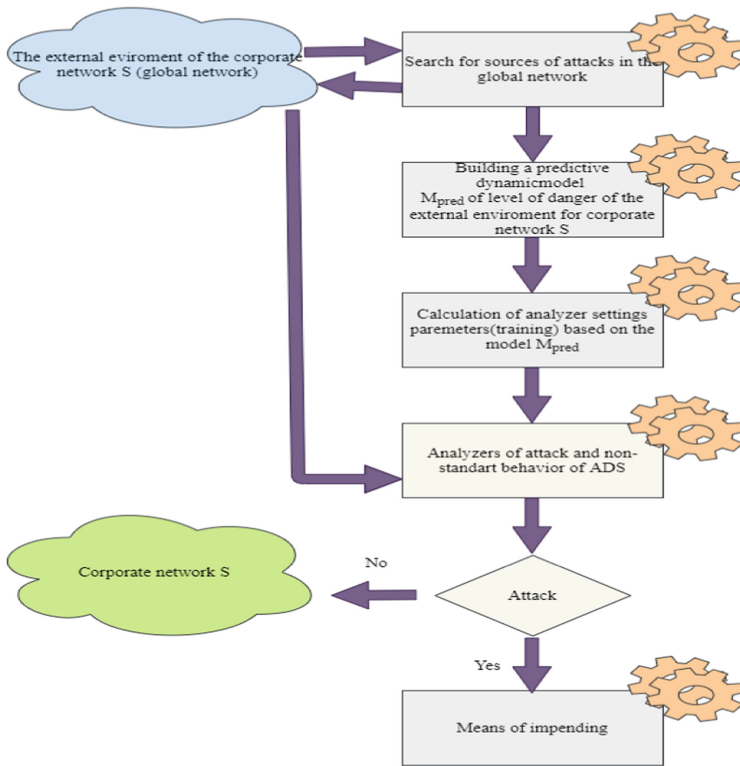


Fig. 7. Training of modules of the ADS analyzers on the basis of the predictive dynamic model M_{pred} of the environmental danger level of the corporate network S .

Availability of such model will give an opportunity to conduct warning training of the ADS analyzers, bringing their r achieve adaptive control of the recognition capabilities of the ADS analyzers, their recognition capabilities to the level corresponding to the state of environmental danger for the predicted time. It allows reaching an adaptive management with recognition capacities of analyzers of ADS.

An important point in the implementation of forecasts is the use of time rows for the events of the attacks. Such information is available in the logs of the corporate network S , as well as in national and international security centers. The latter requires of projecting its merger to corporate network. It is possible to use of traditional forecasting methods, for example, group method of data handling (GMDH) the algorithms with having specified scope of incoming information.

The result of the development of prediction algorithms are analytical expressions for the prediction of the environmental danger situation for the corporate network S , which is a predictive dynamic model M_{pred} .

The next step of adaptive management is the configuration of the ADS analyzers performed by the training methods.

An example of such setting is the training of neural networks to recognize a certain type of attack using the backpropagation method. The presence in the description of the type of attack A_i of such a parameter as γ_i makes it possible to conduct such a training. However, due to the length of the training process, it is not rational to train the network standing in the operational security management loop within the control cycle (Fig. 6).

If there is a predictive dynamic state model M_{preds} , you can warningly train the double of the neural network of the analyzer, and then in the main neural network, you can replace the settings (weights of neural elements) with the resulting duplicate settings. But the warning setting must be done at a processing power different from the power of the analyzer. The warning interval depends on the accuracy of the forecast, and the discrepancies between the current state of the environmental danger and the predicted state, can serve as a feedback signal.

Thus, the model of adaptive safety management in its general form can be represented as shown in Fig. 8.

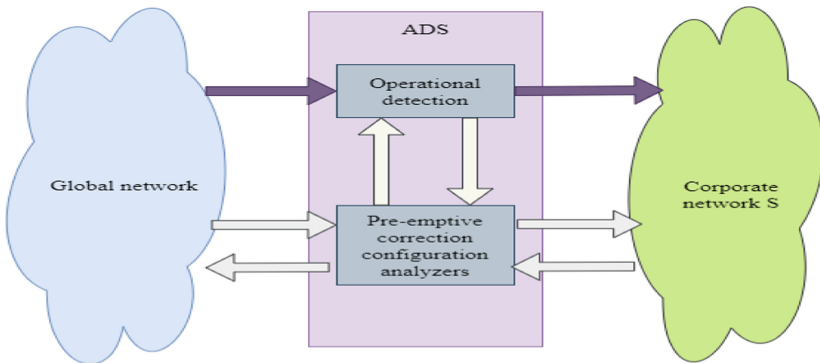


Fig. 8. Adaptive corporate network security management scheme.

Actually, management is carried out by both adjusting types of A_{mi} attacks recognized by ADS analyzers, and by elimination of the vulnerabilities β_i used by the attack.

6 Cybersecurity Industry Centers

The solution to the tasks of excluding the procedures for setting up recognition algorithms and dividing the analysis process into analysis for analyzer settings and analysis for the rapid detection of attacks and non-standard behavior from the operational control is solved by allocating different computational capacities for these purposes [1].

For corporate networks of industry-specific information systems, analyzing the settings of the SDA analyzers protecting the components of the network association can

be assigned to the Industry Cyber Security Center (ICSC), while the ADS_i of the S_i elements of the network association will assume operational control functions.

In addition to the functions described above for analyzing of the configuration of analyzers for subordinate networks (searching for sources of attacks, predicting the state of danger, calculating the settings of the analyzers ADS_i), the ICSC should carry out the following functions:

- communication with international and national computer security centers;
- monitoring of the Internet and social networks by sources and types of attacks;
- database maintenance on security issues;
- distribution of information about the ADS_i settings of the S_i elements obtained on the basis of the prediction of the state of danger among the elements of the association;
- distribution of information on non-standard behavior, attacks of an unknown type and recommendations for dealing with them among the elements of the association;
- organization of collective defense of the association of networks for the corporation.

Collective protection requires intensive exchange of information about attacks, speeds of their distribution in the corporate network, and recommendations for dealing with them.

Taking into account the intensive exchange of information about settings between ICSC and ADS_i , and about the danger of the environment, network attacks as well as the fact that S_i can be created at different times and on different computing platforms, the most preferred method of uniting industry associations $S = \{S_1, S_2, \dots, S_k\}$ is the integration information bus (Fig. 9).

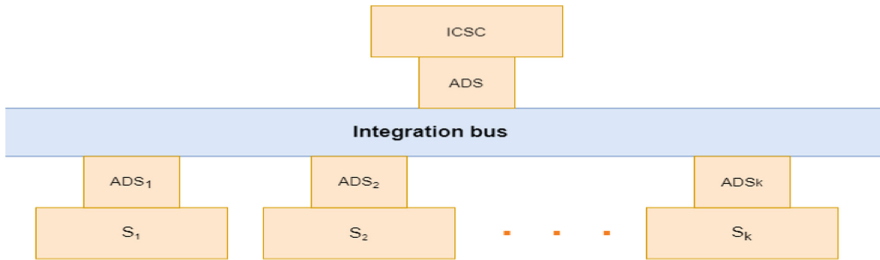


Fig. 9. Integration bus of network association.

The latter is based on the ideas of using Web services, HTTP, XML and its extensions SOAP, WHD [7].

7 Conclusion

Therefore, the use of a varied range of information of a global cyberspace in the identification of attacks allows reaching the alignment of the adaptive control with the current capabilities of the ADS analyzers with the current state of danger of the external

environment of the corporate network. This is achieved by using of adaptive network security management and justifies the need to create industry-specific security centers that provide an increasing of operational efficiency of the ADS of their subordinate networks.

Acknowledgment. The work was carried out and funded under the NATO project CyRADARS (Cyber Rapid Analysis for Defense Awareness of Real-time Situation) – grant agreement number: G5286 [3].

References

1. Abraham, A., Thomas, J.: Distributed intrusion detection systems: a computational intelligence approach. In: Abbass, H., Essam, D. (eds.) *Applications of Information Systems to Homeland Security and Defense*, chapter 5, pp. 105–135. Idea Group Inc, USA (2005). http://w.softcomputing.net/hussein_chapter.pdf
2. Buecker, A., et al.: *Understanding SOA Security Design and Implementation*. 7 ed. Redbooks: IBM Corp., November (2007). <http://www.redbooks.ibm.com/redbooks/pdfs/sg247310.pdf>
3. NATO SPS Project CyRADARS (Cyber Rapid Analysis for Defense Awareness of Real-time Situation). <https://www.cyradars.net>
4. Doynikova, E., Kotenko, I.: Analysis of the current situation and decision-making support for computer network security based on a system of security indicators. *J. Instrum. Eng.* **57** (11), 72–77 (2014). <https://cyberleninka.ru/article/n/analiz-tekuschey-situatsii-i-podderzhka-prinyatiya-resheniy-po-bezopasnosti-kompyuternoy-seti-na-osnove-sistemy-pokazateley>
5. Gazzarata, G., Gazzarata, R., Giacomini, M.: A Standardized SOA based solution to guarantee the secure access to EHR. *Procedia Comput. Sci.* **64**, 1124–1129 (2015). <https://core.ac.uk/download/pdf/82597744.pdf>
6. Indrakanti, S.: Service oriented architecture security risks and their mitigation. DTIC Document (2012). <https://apps.dtic.mil/dtic/tr/fulltext/u2/a576267.pdf>
7. Clark, K.J.: Comparing web APIs with service-oriented architecture and enterprise application integration. Integration architecture. Published on 18 March (2015). <https://developer.ibm.com/technologies/web-development/articles/comparing-web-apis-with-service-oriented-architecture-and-enterprise-application-integration/>
8. Lytvynov, V., Stoianov, N., Skiter, I., Trunova, O., Hrebennyk, A.: Using decision support in finding sources of attacks on computer networks under uncertainty. *Math. Mach. Syst.* **4**, 38–51 (2019). http://www.immsp.kiev.ua/publications/articles/2019/2019_4/04_Litvinov_19.pdf
9. Web Services and Contemporary SOA. Part II: Advanced Messaging, Metadata, and Security. <https://www.pearsonhighered.com/assets/samplechapter/0/1/3/1/0131858580.pdf>
10. White Paper. Industrial Cyber security for Small and Medium-Sized Businesses. A Practical Guide. International Society of Automation Copyright © ISA (2017). https://www.isa.org/uploadedFiles/Content/PDFs/Industrial_Cybersecurity_for_SMB_WP.pdf



Petri-object Simulation Two Level Visual Programming Language

Inna V. Stetsenko^(✉)  and Anton Dyfuchyn 

Igor Sikorsky Kyiv Polytechnic Institute, 37 Prospect Peremogy,
Kiev 03056, Ukraine
stiv.inna@gmail.com, difuchin@gmail.com

Abstract. Visual components are often used by simulation tools to represent the model, constructed of basic elements, and the simulation results. As a rule, the logic of basic elements cannot be modified by the user that is a significant restriction, especially, if a researcher creates a model of an unusual process. Petri-object simulation is the technology for constructing models with a large number of elements. This technology provides faster performance comparing with stochastic Petri net. Being based on object-oriented technology it ensures fast reproducing of model elements. However, creating a Petri-object model without specific tool is a time-consuming process. To reduce the time spent on model construction and to increase the created model accuracy the visual programming language is developed. The description of language is presented in terms of context-free grammar. The advantages of Petri-object simulation visual programming language are considered on the example of the model of rectangular communication grid.

Keywords: Visual programming · Simulation algorithm · Petri-object simulation · Stochastic Petri net

1 Introduction

Visual programming is associated with the tool that helps to write program source code by manipulating graphical objects instead of typing its text. Concerning the business process simulation, Arena software should be pointed to as the best example [1]. It contains an excellent designer of hierarchical model and provides creating 2D or 3D animation. Simio framework improved the Arena software approach by including inheritance of graphical objects, which can be constructed by the user [2]. Instead of the hierarchical structure of a model, which is specific for Arena software, creating a model in Simio is based on the object-oriented structure. However, the model in both cases can be composed only of basic elements, the set of which will never be sufficient while technical systems evolve.

It is not the only reason why researchers return to universal programming languages when they start to develop the model of a specific object of investigation. The requirements for model usage flexibility are very high nowadays. The model should be easily embedded (integrated) in a software package (or information system), be easily

modified and customized. The model created for a specific project should contain only code that provides its functioning, reducing redundant code.

Another approach is the usage of mathematical formalisms for model construction, the most powerful of which is stochastic Petri nets. The advantages of this formalism are a simple unified way of concurrent events presentation, the minimal set of basic elements needed for model description, the logical-algebraic expression of the change of the state.

Well-known software for Petri nets simulating and analyzing is CPNTools [3] supporting colored Petri net which is an extension of stochastic Petri net. To achieve model construction flexibility, this type of Petri net includes specific parameters to determine different types of tokens and to take into account the type of token in the rule of events occurring. This approach allows to reduce the total number of elements for the model construction. However, it leads to the significant complication of the model representation and as a consequence the model debugging. The large part of model functionality is hidden beyond the parameters of elements. Therefore, the benefit of visual construction is reduced.

On one side, universal programming languages give high flexibility for model construction. On the other side, the debugging of a complicated net using universal language is time-consuming task. A balance between using visualization and programming could be found if using visual language conception.

Visual language means “language for programming with visual expressions”, or “visual programming language” [4]. The tools of visual programming are especially needed when simulation model is built of a large number of elements. The multiple connections between elements are more perceptible in a graphical form.

The goal of this research is to develop the visual language for Petri-object model programming characterized by the following properties:

- using minimal set of visual elements for model creation,
- easy and fast (with minimal number of click and mouse movements) creating and manipulating of the model elements, and setting parameters,
- graphical presentation which can be easily perceived,
- support both superficial and deep levels of model construction to provide flexibility,
- generated code of the model according to the graph image.

It should be noted that the graph image built with the tool is used not only for better understanding or improving design of a model. According to the image, the code should be automatically generated. Thus, it is like model image interpretation in the form of program code.

Section 1 introduces to the problems of a simulation software. The second section describes related works. The next section represents the foundations of Petri-object approach which is used for simulation. Section 4 contains the description of developed visual language. The fifth section contains the example of creating model of rectangular communication grid by means of Petri-object simulation visual language. The last section summaries the article and gives the perspective on future research.

2 Related Works

An introduction to the theory of stochastic Petri net is given in book [5]. The authors of work [6] have done a comparison of twenty Petri net tools. Only 2 of 20 tools support object-oriented Petri net and 8 of 20 tools support stochastic Petri net. Unlike existing tools based on Petri net formalism, Petri-object simulation supports object-oriented technology and stochastic Petri net.

For the purposes of microcontrollers design, several pieces of research were devoted to the development of tool being used to generate program code automatically. The problem of automatic generated Assembly code from a Petri net is discussed in work [7]. The IOPT tool providing automatically generated VHDL code based on Petri net is described in work [8]. Being developed for designing embedded system controllers the tool uses Petri net simulator to debug model and to generate code deploying on the target hardware platforms [9]. Another example of using Petri net for automatic code generation was presented by the author of thesis [10]. It was proposed to develop the model of game mechanics by means of colored Petri net and to use the C# automatic generated code according to the model as a part of game code. Despite the fact that the tools above have been developed for specific areas of model designing the same advantages have been revealed. It is easier to percept complicated structural properties using graphical representation. In addition, the time spending on development is significantly shorter due to the smaller number of errors and simpler debugging.

Visual programming is an approach to create a program without manually typing its code. The perfect discussion of the term ‘visual language’ and the research of visual language syntactic and semantic features could be found in work [11]. The problem of grammar determination of the visual language is considered in work [12]. It was proposed using a graph grammar to specify visual grammar rules. The approach had been used for the definition of visual languages like UML diagrams. The environment GenGED supporting graph grammar concept was presented in work [13]. Visual language specification can be obtained using this programming tool. (the main purpose of which is the visual language design)

According to the classification given in work [14] there are following categories of visual languages: purely visual language, hybrid text and visual systems, programming-by-example systems, constraint-oriented systems, form-based systems.

Nowadays software engineering exploits terms no-code and low-code development platform for the tools providing the development of an application using graphical user interface instead of traditional programming. Discussing the software development future, IT professionals state such advantages of using low code development platform as a quickness, low maintenance and aids communication [15].

The main goal of using visual programming tools is to simplify the process of coding, to reduce the number of errors for the complicated structures and as a result the time spent on the program development.

In this research, we have developed the visual programming language for creating and simulating a Petri-object model. The language can be classified as two-level labelled graph using dynamic semantics. The significant difference from previous research is a strict formalization of the new language grammar rules, simplicity of the simulation model creation of the classes of typical elements, flexibility of the behavior creation for each class of typical elements using stochastic multichannel Petri net.

3 The Basics of Petri-object Simulation

Petri-object simulation technology grounded on stochastic multichannel Petri net and object-oriented technology. It provides duplicating model elements with the same Petri net and composing the model functioning of the multiple sets of elements. Mathematical description of Petri-object model was formulated firstly in [16] as the model consisted of Petri-objects. The main feature is that the model functioning is described by stochastic multichannel Petri net obtained as a union of Petri nets describing the functioning of its elements. Moreover, the transformation of the model state is divided into transformations of the model elements states:

$$t_n = \min E(t_{n-1}), t_n \geq t_{n-1}, n = 1, 2, \dots \quad (1)$$

$$S(t_n) = \{(D^-)^m (D^+ (\tilde{S}_h(t_{n-1})))\}, h = 1, 2, \dots, L\}, \quad (2)$$

where t_n is the current moment of time which is determined by the nearest moment of time, $E(t_{n-1})$ is the state of Petri net transitions in the previous moment of time, $S(t_n)$ is the state of Petri net that describes the total model in the n -moment of time, $\tilde{S}_h(t_{n-1})$ is the state of h -element, which is a Petri-object, in the previous moment of time, modified if it has been connected with other in a second way, D^+ and $(D^-)^m$ are transformations of Petri net state describing tokens output and m -times repeated tokens input correspondingly that are described by logic-algebraic equations.

As it was defined a Petri-object is an object of the class which has a Petri net as a description of its 'rule of life'. On one side, such object is an ordinary object in terms of object-oriented programming. On the other side, it is an elementary component of discrete event functioning of the system. Elements with similar behavior can be created as instances of a class, the Petri net of which is determined. The Petri net once built can be used for creating multiple instances, therefore the replication of the Petri net is supported. Moreover, the parameters of Petri net can follow the values of instance fields.

Petri-objects can be connected by only two ways: using shared places or using event initialization. Because the second one can be established between more than two objects it can be named as multiconnection. The list of connected Petri-objects can be used for creating the Petri-object model.

The proof of Petri-object model complexity was presented in work [17]. Due to the dividing the model on structural parts we achieve significantly decreasing of computation complexity as well as simplicity of model representation. To create Petri-object model, which consists of N types of elements, following routines should be done:

- to construct Petri net for each type and to define which characteristics of Petri net are used as parameters for instance creation,
- to define places of Petri net which can be used as shared places,
- to save Petri net in database,
- to instantiate Petri-object for each element using correspondent Petri net with given set of parameters,
- to link Petri-objects in accordance to model elements connections,
- to construct model with the list of Petri-objects,
- to simulate model with given duration.

4 Petri-object Model Two Level Presentation

4.1 The Formal Grammar of Petri-object Simulation Visual Language




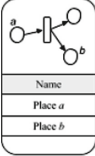

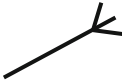
Formal grammar is used for programming language description. The common notation of context-free grammar is Backus-Naur form and its extensions [18]. Petri-object simulation visual programming language as other languages can be described by following alphabet and grammar. Alphabet consists of images of place, transition, arc, Petri-object, connection and multiconnection represented in Table 1. The description of grammar in Augmented Backus-Naur Form notation includes following rules:

```

<PetriObjModel> ::      {{<PetriObj>}      {<Connection/Multiconnection
{<PetriObj>}}
<PetriObj> :: "imageObject" <PetriNet> [ <args>]
<Connection> :: "imageConnection" <parameters>
<Multiconnection> :: "imageMulticonnection" <parameters>
<PetriNet> :: {<Place> {<Arc> <Transition> <Arc> <Place>}} | {<Place>}
<Place> :: "imagePlace" <parameters>
<Transition> :: "imageTransition" <parameters>
<Arc> :: "imageArc" <parameters>
<parameters> :: <parameter> { <parameter>}

```

Table 1. The alphabet of Petri-object model visual programming language.

Name	Image	List of parameters
Place		name, number of tokens
Transition		name, delay, type of distribution, mean value
Arc		name, input, output, quantity, is inf
PetriObj		name, PetriNet id, args
Connection		pairs of connecting places between two Petri-objects
Multi-connection		a pair of connecting transition and place between Petri-object and the set of Petri-objects

Notice that images of first level elements have the same shape as traditional Petri net elements. The images of second level elements are specific for Petri-object model and have been developed in this research. The user writes a sentence in terms of visual language when he creates a chain of connected elements. The set of sentences are a description of the model. Chains of connected elements are identical to chains of events concurrently simulated.

4.2 The First Level: Petri Net Construction

The elements Place, Transition, and Arc of the alphabet are the elements of the first level been intended to create Petri net. Visual language provides tools for creation of Petri net by picking Place or Transition element from the toolbar and placing it into canvas. Then places and transitions can be connected with arc elements. Each created element has default parameters which can be later configured with a special form. When creation of Petri net is finished user can save it for further work. The environment will perform validation of created Petri net during saving. For testing purpose, a user can run Petri net functioning. Run can be performed in two modes: simulation to closest event or simulation during given time.

Petri net editor allows creation of Petri net elements as graphical elements and manipulation with them such as drag n drop, scaling, etc. All Petri nets used for model construction should be saved in NetLib class. The example of created stochastic multichannel Petri net is represented in Fig. 1. For the purpose of Petri-object model construction, the elements which will be used for connections should be specified.

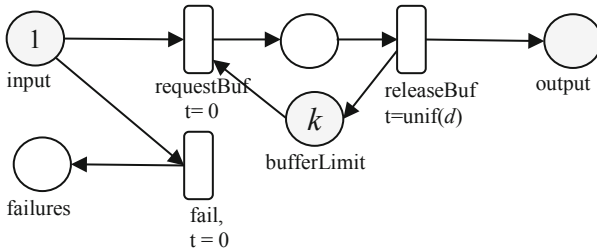


Fig. 1. The stochastic Petri net example (shared places are marked with gray color).

4.3 The Second Level: Petri-object Model Creation

The second level of visual programming environment provides a tool for creating Petri-object model. Since Petri-object model consists of connected Petri-objects, the main purpose of the second level is to create Petri-objects and make connections between them. Each Petri-object is based on Petri net created at the first level. User can pick a Petri net from the left sidebar of the window and drag it onto a canvas to create a Petri-object of specific type. After the Petri-object has been created user needs to specify its parameters by filling a table. Every row of the table corresponds to a new Petri-object with these parameters. So, user could create many objects in one table. In addition, the option of download file with array of parameters is supported. The name of parameter for each column can be specified by the user. This way of objects creation is very convenient if more than 10 objects are needed. Hundred or more objects can be imported from file.

Graphical representation of Petri-object is depicted in Fig. 2. It consists of Petri-object icon, name and list of the shared positions. User can connect Petri-objects by drawing a line between corresponding shared positions as represented in Fig. 2. After the Petri-objects creation have finished user can run created Petri-object model with given simulation time. The environment will perform validation process which includes checking if connections are made correctly and its parameters have a correct type.

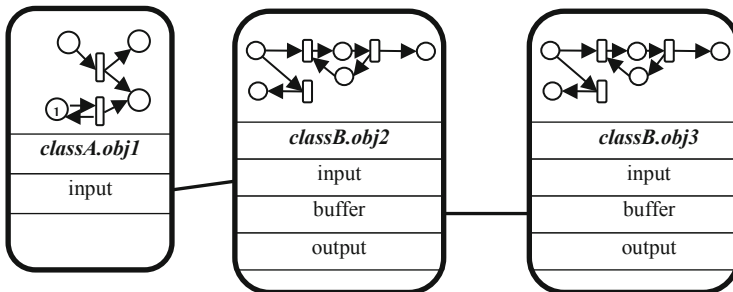


Fig. 2. Connections between Petri-object *obj1*, *obj2* and *obj3* of the classes *A* and *B* correspondently.

4.4 Model Construction Example

Let us construct a model of a rectangular communication grid consisted of four ports communication devices which are connected with each other creating a rectangular structure. The model was represented by means of basic Petri net and investigated by invariants in work [19]. Repeated elements were represented in the model by parametric description that can be applied only for regular structure which is difficult for perception and connections editing. The usage of Petri-object model provides creating a structure connecting an elements of communication grid the parameters of which can be different.

Firstly, the classes of Petri-objects should be defined. It would be rational to define classes *Port* and *Device*. An element of class *Port* reproduces receiving and sending packages. A *Device* element represents only the state of buffers of the communication device. If required buffer has enough space a package can be served by the communication device. The net of Petri-object *Port* is depicted in Fig. 3.

Secondly, connections between objects should be established. Connections between objects of *Port* and *Device* are depicted in Fig. 4. Each communication device composes of the one object *Device* and four objects *Port*. Ports of different communication devices have connections between its input and output elements as it is represented in Fig. 5. Then the structure of rectangular communication grid can be obtained by creating connections between correspondent ports (Fig. 6).

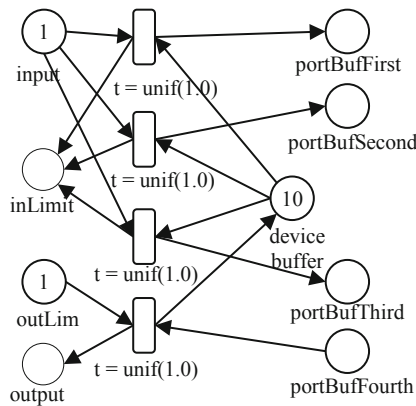


Fig. 3. Stochastic Petri net presenting the net of Petri-object *Port*.

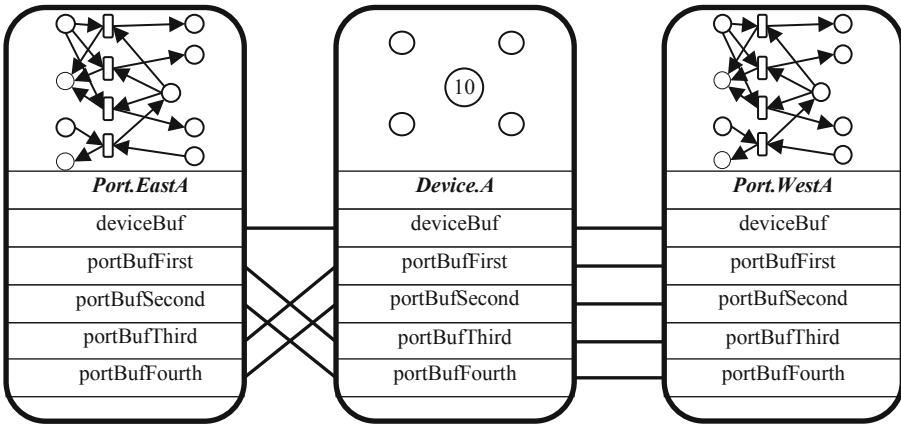


Fig. 4. Connections between a Petri-object of class *Device* and a Petri-object of class *Port*.

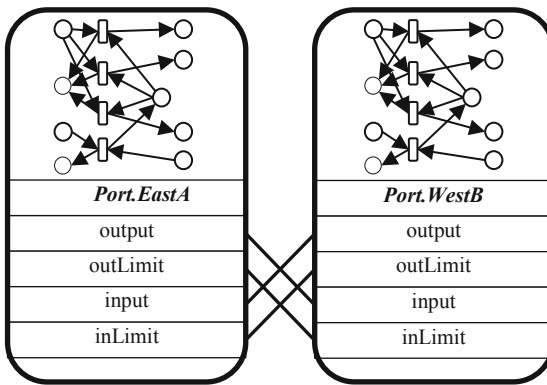


Fig. 5. Connections between Petri-objects of the class *Port*.

Thus, developer could compose the model of repeated elements duplicating them and connecting with each other in a corresponding way. The simulation algorithm will reproduce the model behavior according to the stochastic Petri net. However, the transformations are divided on transformations of Petri-objects that significantly decrease the time performance.

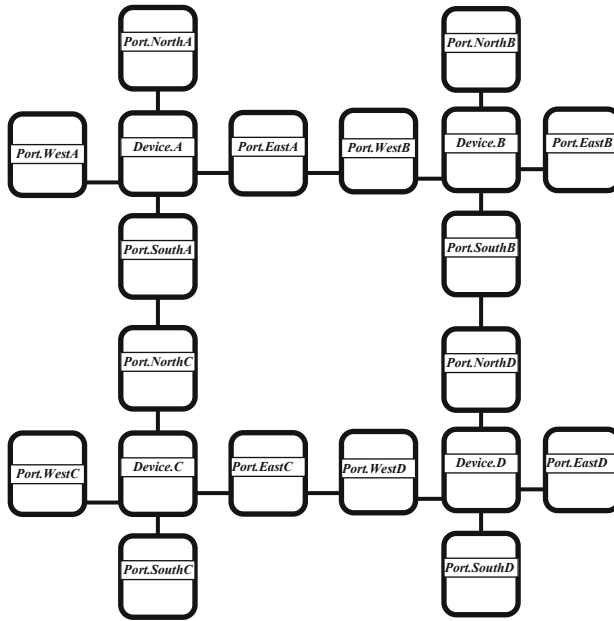


Fig. 6. The graphical image of rectangular communication grid Petri-object model.

5 Conclusions

In this research we developed the new visual programming language for discrete event simulation which is based on Petri-object simulation technology. Petri-object technology gives an opportunity for multi-level model development: Petri net level and Petri-object model level. This approach provides a clear and convenient way for developing large complex models using stochastic Petri nets. However, it needs visual model construction layer to simplify routines of connecting elements.

Visual programming language saves the balance between model flexibility and the advantages of visual construction of the model. User can develop both low-level components using stochastic Petri nets and high-level components, which have own set of parameters, can be multiplied and connected with others.

The software flexibility can be extended by serializing model for further usage in popular programming languages. The opportunity to create Petri-object composing other Petri-objects will be considered in future research.





References

1. Kelton, W.D., Sadowski, R.P., Zupick, N.B.: Simulation with Arena, 6th edn. McGraw-Hill Professional, New York City (2014)
2. Prochaska, K., Thiesing, R.M.: Introduction to simio. In: 2017 Winter Simulation Conference (WSC), Las Vegas, NV, 2017, pp. 4410–4419 (2017)

3. CPNTools: A tool for editing, simulating, and analyzing Colored Petri nets. <http://cpntools.org/>. Accessed 26 Apr 2020
4. Chang, S.-K., Ichikawa, T., Ligomenides, P.A.: Visual Languages. Plenum Press, New York City (1986)
5. Haas, P.: Stochastic Petri Nets: Modelling, Stability Simulation. Springer, New York (2002)
6. Thong, W.J., Ameen, M.A.: A survey of Petri Net tools. In: Sulaiman, H.A., Othman, M. A., Othman, M.F.I., Rahim, Y.A., Pee, N.C. (eds.) Advanced computer and communication engineering technology. LNEE, vol. 315, pp. 537–551. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-07674-4_51
7. Dezani, H., Marranghello, N., Pereira, A.S., da Silva, A.C.R., Węgrzyn, M.: Automatic code generation for microcontrollers from place-transition Petri Net models. IFAC Proc. Vol. **44** (1), 7873–7878 (2011)
8. Pereira, F., Moutinho, F., Gomes, L.: IOPT Tools User Manual Version 1.1 (2014). http://gres.uninova.pt/iopt_usermanual.pdf. Accessed 02 Apr 2020
9. Pereira, F., Gomes, L.: Cloud based IOPT Petri Net simulator to test and debug embedded system controllers. In: Camarinha-Matos L., Baldissera T., Di Orio G., Marques F. (eds.) Technological Innovation for Cloud-Based Engineering Systems. DoCEIS 2015. IFIP Advances in Information and Communication Technology, vol. 450. Springer, Cham (2015)
10. Carlsson, M.: Automatic Code Generation from a Colored Petri Net Specification for Game Development with Unity3D (2018). <http://www.diva-portal.org/smash/get/diva2:1220952/FULLTEXT01.pdf>. Accessed 02 Apr 2020
11. Erwiga, M., Smeltzer, K., Wanga, X.: What is a visual language? J. Vis. Lang. Comput. **38**, 9–17 (2016)
12. Bardohl, R.: A visual environment for visual languages. Sci. Comput. Program. **44**, 181–203 (2002)
13. Bardohl, R., Ermel, C., Ingo Weinhold, I.: Specification and Analysis Techniques for Visual Languages with GenGED. <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=076EB470823CE2C149EB847A42B42B87?doi=10.1.1.7.3498&rep=rep1&type=pdf>. Accessed 2 Apr 2020
14. Stanciu, L.: Visual Programming and Modeling. <https://www.aut.upt.ro/~adrianaa/teaching/VPM/Visual%20Programming%20and%20Modeling.pdf>. Accessed 02 Apr 2020
15. SoftJourn: Low-code Environments: The Future of Apps Development or a Fad. <https://softjourn.com/blog/article/low-code-environments-the-future-of-app-development-or-a-fad>. Accessed 12 Apr 2020
16. Stetsenko, I.V.: Theoretical foundations of petri-object modeling of systems. Math. Mach. Syst. **4**, 136–148 (2011). (In Russian)
17. Stetsenko, I.V., Dorosh, V., Dyfuchyn, A.: Petri-object simulation: software package and complexity. In: The 8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS 2015), pp. 381–385, IEEE, Piscataway (2015)
18. Might, M.: Grammar: The language of languages (BNF, EBNF, ABNF). <http://matt.might.net/articles/grammars-bnf-ebnf/>. Accessed 02 Apr 2020
19. Zaitsev, D., Zaitsev, I., Shmeleva, T.: Infinite Petri Nets: part 1, modeling square grid structures. Complex Syst. **26**(2), 157–195 (2017)



Detection of SQL Injection Attack Using Neural Networks

Oleksandr Hubskeyi¹ , Tetiana Babenko¹ ,
Larysa Myrutenko¹ , and Oleksandr Oksiuk² 

¹ Taras Shevchenko National University of Kyiv, Kiev, Ukraine
a.gubskyy@devhub.solutions, babenkot@ua.fm,
myrutenko.lara@gmail.com

² Department of Cyber Security and Information Protection, Taras Shevchenko
National University of Kyiv, Kiev, Ukraine
oksiuk@ukr.net

Abstract. The use of data mining technologies, including neural networks, has become widespread in cybersecurity. Outdated approaches for analyzing and identifying rule- or template-based attacks do not always produce the desired result. Successful implementation of a cyber-attack can lead not only to disruption of the information system, but also to potential losses. In this research a neural network model for identifying SQL injection attacks based on HTTP request analysis has been developed. The model allows classifying URL values by attributing them into one of two classes: attack or normal activity. An additional advantage is the provision of a quantitative identification value which describes the predicted accuracy of SQL injection determination.

Keywords: Cyberattack detection · SQL injection attack · Detection of SQLi attack · Neural network · Machine learning

1 Introduction

SQL injection (SQLi) attack type is among 10 types of the most critical Web-application security risks by OWASP (Open Web Application Security Project) methodology [1]. SQL injection vulnerabilities are typical not only for Web-applications. Implementation technologies of this attack type are well-known and widely used by attackers. The purpose of this attack type is the execution of different commands, fingerprinting of database management systems (DBMS), or obtaining user's data and arbitrary code execution in client's side [2–5]. SQL injection vulnerabilities exploitations are divided based on implemented DBMS type and attack realization conditions: vulnerable request (Insert, Update, Delete can be processed; injection can be performed on any part of the SQL query; blind insertion of SQL statements; SQL language features are used for various DBMSs [6]. Types of SQL injection vulnerabilities are presented Fig. 1. It should be noted that SQL injection vulnerability could be exploited both in the moment of attack and after a certain period of time. SQL injection attack vulnerability can occur in any web application that uses a SQL database.

Such bases can be: MySQL, MSSQL, Oracle, SQL Server, PostgreSQL, etc.

SQL injection attacks (SQLi) are among the 10 most common Open Web Application Security Project (OWASP) methodology.

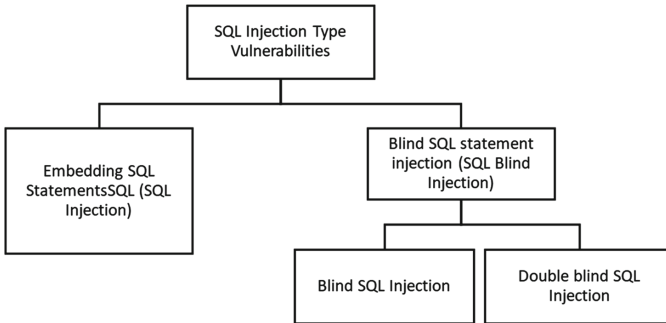


Fig. 1. Types of SQL injection vulnerabilities

Ways for identify software vulnerabilities to SQL injection are: function testing (black/white-box); fuzzing; static, dynamic, manual analysis of source code. It should also be noted, that in parallel to the vulnerabilities search in software applications, a wide range of WAF (web application firewall) is commonly used. Typically they are based on two security models: signature-based and rule-based. Each of these models has its advantages and disadvantages, but their common disadvantage is the inability to identify “zero-day” threats, and it should be noted, that the use of WAF can only partially cover the vector of attacks [7–10].

As follows, the majority of typical approaches for maintaining security from SQL injection attacks do not allow obtaining a sufficient security level due to low identification accuracy and working speed. Therefore, currently a large quantity of different technologies of computer systems and networks protection, which are based on data mining technologies, in particular on the use of neural networks, which allows effective counteracting already known attacks and “zero-day” attacks, are developed.

Artificial Intelligence – a term in the broad sense, based on the imitation of human capabilities through computers: feeling, understanding, responding.

Machine Learning – a field of artificial intelligence in the section of computer science, which often uses statistical techniques to give computers the ability to “learn” (e.g. gradually improve productivity in a particular task) [11].

Data Science – for performing (using) machine learning algorithms, there is need to define datasets, select the appropriate variables and metrics, and perform various information engineering tasks: search for hidden dependencies, data collection, training, integration, visualization, determination of the algorithms performances etc.

SQL (Structured Query Language) injection is an attack that is implemented by modifying database queries due to exploitation of vulnerabilities that are contained in

Web-applications. Successful implementation of the attack gives the attacker the opportunity to obtain sensitive information, change or even destroy it.

2 Related Works

2.1 Method

As known, SQL is query language designed to manage data stored in relational databases. Most websites work with data stored in SQL databases. In some cases, SQL can be used for launching operating system commands.

As known, a webpage or web application, which are vulnerable to an SQL injection attack, use the user-entered parameters directly in the SQL query. An attacker can modify incoming content. This content is often classified as malware and is a key part of the attack. After the attacker sends specially-created SQL query, commands, which were unplanned by developers, start to execute [1–10].

Typically, a SQL query looks as follows:

```
select id, forename, surname from authors
```

This request gives an opportunity to receive «id», «forename», «surname» columns from «authors» table. The result of database response is table rows. To detail the response from the server, the request is specified:

```
select id, forename, surname from authors where forename
= 'john' and surname = 'smith'
```

The important point is that the string parameters are separated by single quotes. Assuming that these values are formed based on user’s input (e.g. a web-form), an attacker can modify the structure of the original SQL query by entering the following parameters:

```
Forename: jo'hn
Surname: smith
```

The resulting query will look as follows:

```
select id, forename, surname from authors where forename
= 'jo'hn' and surname = 'smith'
```

When the database executes this query, an error similar to the following occurs:

```
Server: Msg 170, Level 15, State 1, Line 1
Line 1: Incorrect syntax near 'hn'.
```

The reason for this error is that the installation of single quotes breaks the request for parts. By setting the value of the field “forename” to “jo”, the database tries to

execute the next block “hn”, which causes an error. If the attacker enters the value of the ‘forename’ field as “ ‘jo’; drop table authors–”, then the authors table will be deleted, which will result in data loss. Why so? The character “-” sequence is a one-line comment that exists in the Transact-SQL extension. This is required in order to finish query without errors (the rest of the main query will be viewed by the database as a comment and will not be executed). The symbol “;” delimits the end of one request and the beginning of another. There can be a huge variation amounts in the base manipulation, which carries some damage and losses.

Developers, system administrators, DevOps must take a number of precautions to minimize the impact of this attack on business services and to the purpose of maintaining the end users’ security. Such actions include filtering and blocking queries that contain potentially malicious content and can lead to a SQL attack. This research examines the creation of an HTTP (Hyper Text Transfer Protocol) module that can identify SQL injection attacks using artificial neural networks.

Artificial Neural Networks are mathematical models and their program or hardware realizations. This term appears while studying processes, which run in the brain, and while trying to simulate these processes. The basic principles are the interpretation of sensory data through a kind of machine perception, marking or grouping of the incoming data. The recognized patterns are numeric and are contained in vectors, into which any other data is translated [12].

Every node has one or several inputs and one output. The neuron has two operation modes: training mode and using or testing mode. In training mode, the neuron learns to respond to a specific input pattern. In operating mode, the neuron responds to the input pattern and associates the output. In the case neuron receives to input not typical set of parameters, it determines whether to activate itself or not.

Each input signal has a corresponding weight, which is calculated based on the input data. If this number exceeds the threshold, the neuron will fire.

Activation of any neuron is regulated by its activation function.

2.2 Model Development

The main purpose of the work was to create a logical module that could attribute the incoming HTTP request to the SQLi attack.

Proposed model contain three main components:

- URL (Uniform Resource Locator) address generator.
- URL address classifier.
- Model based on neural network.

URL address generator is a testing module, which is needed for output dataset forming. It contains two components: common URL addresses generating by collecting information from popular sites (e.g. sitemap.xml file analysis) and generating malicious URLs by adding URL query parameters, which are typical for SQL-injection, to URLs, obtained by previous method. Another way, that was used to fill a sample with malicious parameters, is to use open source datasets.

Typical requests for SQL injection contain key words in SQL language, which are commonly used for operating with SQL database tables. Requests can be applied to the

entire database: creating, deleting tables, etc., as well as to individual tables: modifying records, refining searches, etc.

Typical parameters can be:

- those, beginning with на « '»;
- those, finishing with «-», «/* */», «#»;
- those, containing UNION, SELECT i FROM, information_schema, exec, logical operators (OR, AND,=) and expressions (1=1);
- Those, containing group phrases ADMIN DROP, CREATE TABLE, DELETE FROM, INSERT INTO etc.

In order to synthesize and analyze the model of attack identification, preliminary training, control and test datasets must be prepared. Training dataset contains the parameters of learning object; the selection of parameters was heuristically-based on the essential features analysis of the attack, which may contain a URL. As known, neural-based models work solely with numerical data represented in some numerical range, so in the first phase of the research, the URL classifier was developed. It is a software module that converts the URL into a binary format and sets the logical identifier “true (1)” if the address refers to an attack, and “false (0)” otherwise. Thus, an input vector was generated for each of the URLs. Analyzed parameters are presented in Table 1.

Table 1. Assigned vectors to SQL parameter patterns

Parameter number of the input vector	Query parameter
x ₁	'(single quotes)
x ₂	CREATE
x ₃	DELETE FROM
x ₄	DROP TABLE
x ₅	INSERT INTO
x ₆	UNION
x ₇	AND
x ₈	OR
x ₉	-
x ₁₀	(space)
x ₁₁	FROM
x ₁₂	EXEC

Thus, output vector can be presented in a format:

$$X^T = [x_1 x_2 \dots x_n]$$

For example, with n = 12 parameters, URL, which contains the expression «CREATETABLE ... AND INSERT INTO...» can be presented as: 010010100100.

Based on this approach, neural-network model will have 12 neurons at the input.

Every vector is characterized by the parameter: Benign (0) – does not refer to an attack, while Injection (1) – refers to an attack. It is enough to have only one neuron at the output, dividing the input vectors into two classes 0 and 1.

Synthesis of the neural network model was performed on the basis of Rumelhart's multilayer perceptron. Rumelhart's multilayer perceptron is a special case of the Rosenblatt perceptron, in which the neuron's weight coefficients are adjusted by the error-reversing algorithm. The peculiarity is the presence of more than one layer (usually two or three layers) [12]. Thus, a neural network in the form of a Rosenblatt perceptron divides the input vectors into two classes 0 and 1. The training sequence is formed of two arrays: an input array X and an array of purposes Y , which assigns each of the input vectors to one of two classes.

Three layers were used in the research: input, output and one hidden. Operations of back propagation neural networks can be divided into two stages: forward and reverse propagation. In the forward propagation step, an input template is applied to the input layer and its effect is propagated, layer by layer, through the network until output is obtained. The actual output value of the network is compared to the expected output, and an error signal is calculated for each of the output nodes. Since all the hidden nodes contributed to some extent to the errors detected in the output layer, the output signals errors are transmitted back from the output layer to the each node in the hidden (inner) layer, which affects the output layer. This process is then repeated in layers until each node in the network receives an error message that describes its relative contribution to the overall error.

Once the error signal for each node is determined, the error data will be used to update the weight values of each connection until the network enters a state that allows all training schemes to be coded. The back propagation algorithm looks for the minimum value of the error function in the weight space using a technique called the delta rule or gradient descent. Scales, which minimize the error function, are considered to be a solution to the learning problem [12, 13].

While defined template is given to the input layer during the training process, the weighted sum of j th node input in hidden layer is calculated by formula:

$$Net_j = \sum w_{ij}x_j + \theta_j \quad (1)$$

Equation 1 is used to calculate the total input to a neuron. θ_j , which is a weighted offset node that always has an output value of 1. The offset node is considered a "pseudo input" for each neuron in the hidden and output layer, and is used to solve problems in situations when the input template value is equal to zero. If any input template contains zero values, the neural network cannot be trained without an offset node.

To decide whether or not a neuron should be activated, the value of the action Net_j potential is passed to the corresponding activation function. The resulting value of the activation function determines the output of the neuron and becomes the input value for the neurons in the next layers that are associated with it.

Since one of the requirements for the back propagation algorithm is that the activation function must be differentiated, a typical function is the sigmoid equation.

$$O_j = x_k = \frac{1}{1 + e^{-Net_j}} \quad (2)$$

It should be noted that other types of functions, for example, hyperbolic, can be used. Equations 1 and 2 are used to determine the output value of node k in the output layer.

2.3 The Experimental Part

Node JS programming language and ancillary third-party module Synaptic were used for developing the appropriate software.

Synaptic is a library of JavaScript neural networks for node.js, which allows creation any training basically any type of first order neural network architecture and even second order [12].

Thus, the synthesized model has three layers of neurons, the neurons of the input and hidden layers have a sigmoidal activation function, while the output layer is linear. The model has the following architecture: input layer of 12 neurons; hidden layer of 6 neurons; output of 1 neuron. One input vector consisting of 12 elements X, which admissible limits of values are [0, 1] is input to the model. Initialization of the model was performed by random values from the interval [-1, 1].

The number of hidden layer neurons was determined according to following rules:

- “The optimum size of the hidden layer is the value usually between the number of nodes in the input layer and the number of nodes in the output layer”;
- “The number of neurons in this layer is the average between the number of neurons in the output and input layers.”

To set up a neural network, it is needed to define a number of parameters based on the best result with optimal configuration (training speed, error value, etc.).

Learning rate in a machine learning and statistics is a setting parameter in optimization algorithm, which determines the step size at each iteration, moving to the minimum of the error function [14, 15]. The value varies between 0 and 1.

In this research learning rate 0.2 was used.

The maximum number of iterations (epochs) is required to terminate training if a minimum error has not been reached. The number of iterations in the research did not exceed 200.

The minimum of error is the value of the error at which the neural network stops learning. In this research, the value was 0.005.

Input sample was divided into 70% (learning datasets), 15% (control sample) and 15% (testing sample).

A total of 30 233 URLs were generated to form samples.

The training set consisted of 20 182 records, including 12 913 regular and 7 269 malicious URLs.

The control sample included 5 025 records: 3217 normal and 1 808 malicious.

The test sample contained 5 026 records: 3217 normal and 1 809 malicious.

After the training the synthesized model can classify the input data to an accuracy of 95%.

The model training diagram is presented below and shows the change in learning error at each iteration. The training was conducted both on small (Fig. 2) and large samples (Fig. 3).

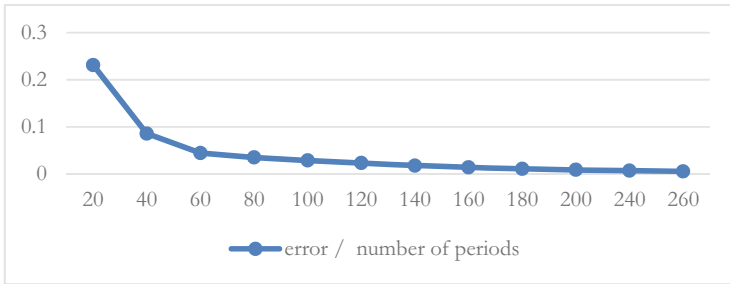


Fig. 2. Learning diagram on small sample

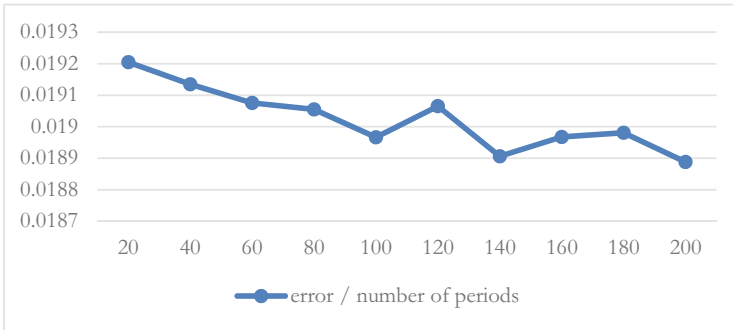


Fig. 3. Learning diagram on main sample

The graph above shows that for the main training sample, the relative learning error is 1.89%.

While operating the neural network, there may be a situation, when the studying SQL injection parameters are randomly presented in the URL and do no harm to the webserver and, in accordance, the database (False Positive), but the system defines them as an attack. In future researches, it is possible to use clear SQL injection templates and take into account the HTTP status of the webserver response code to resolve this issue.

In summary, the relative error of the synthesized model on control and test samples does not exceed 5%. The classification results are presented in Table 2.

Table 2. The result of model training

Traffic type	Training	Control (successfull/all)	Testing (successfull/all)	Relative error, %
Common	12913	3196/3217	3190/3217	0,7%
SQLi	7269	1719/1808	1721/1809	4,89%

3 Conclusions

This paper presents the results of research of the creation of a SQL injection model, based on the neural network of direct signal propagation. We have evaluated the synthesized model on a large sample of data, obtained through the synthesis of maximum quantity of diverse URLs. In our opinion, the proposed model has all the advantages, that makes it possible to obtain the use of a neural network apparatus and achieves high accuracy in solving such a complex problem as the SQL injection identification task, demonstrating stable results. This approach to identifying SQL injection is useful in integrated information security solutions, which are processed using different DBMS platforms.

References






1. SQL Injection. https://owasp.org/www-community/attacks/SQL_Injection. Accessed 21 Mar 2020
2. SQL injection in Sophos XG Firewall/SFOS. <https://www.securitylab.ru/vulnerability/507863.php>. Accessed 1 Mar 2020
3. SQL Injection FAQ. <https://www.sqlsecurity.com/faqs-1/sql-injection-faq>. Accessed 11 Mar 2020
4. SQL injection in Oracle. <http://forum.antichat.ru/showthread.php?t=40576>. Accessed 21 Feb 2020
5. SQL injection in PostgreSQL. <http://forum.antichat.ru/thread35599.html>. Accessed 2 Mar 2020
6. What is SQL Injection (SQLi) and How to Prevent It. <https://www.acunetix.com/websitesecurity/sql-injection>. Accessed 15 Mar 2020
7. Multiple Buffer Format String Vulnerabilities in SQL Server. <http://www.microsoft.com/technet/security/bulletin/MS01-060.asp>. Accessed 11 Mar 2020
8. Taking the monkey work out of pentesting. <http://pentestmonkey.net/>. Accessed 21 Feb 2020
9. The Web Application Security Consortium/SQL Injection. <http://projects.webappsec.org/SQL-Injection>. Accessed 21 Feb 2020
10. 'Microsoft SQL Server extended stored procedure vulnerability (technical explanation and exploit code)' – SecuriTeam. <https://securiteam.com/windowsntfocus/6n0010u0kw/>. Accessed 21 Mar 2020
11. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning, 2nd edn. Springer, New York (2009)
12. Haykin, S.: Neural networks and Learning Machines, 2nd edn. Prentice Hall, Harlow (2009)

13. Artificial Neural Networks Using Multilayer Perceptron. <https://www.cse.unsw.edu.au/~cs9417/ml/MLP2/index.html>. Accessed 5 Feb 2020
14. Menshawy, A.: Deep Learning By Example: A Hands-On Guide to Implementing Advanced Machine Learning Algorithms and Neural Networks, 1st edn. Pact Publishing Ltd., Birmingham (2018)
15. Murphy, K.: Machine Learning. The MIT Press, New York (2012)

Mathematical Modeling and Simulation of Special Purpose Equipment Samples



Improving Accuracy of Ground Targets Determination by Unmanned Aerial Systems

Maksym Gerashchenko , Oleksandr Isachenko ,
Sergey Nesterenko  , and Serhii Rudnichenko 

State Scientific and Research Institute of Arms Testing and Certification
of the Ukraine Armed Forces, 1 Strilets'ka Street, Chernihiv 14033, Ukraine
cranoxxy@gmail.com

Abstract. The practice of using unmanned aerial systems (UAS) shows that the accuracy of determining by them the coordinates of a ground targets (GTC) is often insufficient. However, the quantitative assessment of the degrees of influence of the individual components involved in this process was not performed yet. To assess the degree of influence on the accuracy of the GTC determination the camera geometric factors and the flight control application screen parameters have been taken into account. The relative location of the camera and the on-surface target at the moment of such determining was analyzed. To solve the problem posed the specific values of the partial derivatives of GTC with respect to most influent arguments were assessed and after that they were compared with the accuracy of the geo-positioning. By this method was assessed the sensitivity measure for the inaccuracy of determining the angular position of the camera, and the influence of inaccuracy of specifying by the operator the target's location on the flight control application screen. This may give the UAS's developer the guidances concerning which UAS component may be most promising in terms of improving the accuracy of GTC determination. The made calculations show that at this stage of the UAS development to improve the accuracy of the operational GTC determination worth to focus on improving the camera angular coordinates accuracy at the moment of the target detection.

Keywords: Unmanned aerial system · Target detection · Accuracy · Video reconnaissance · Video surveillance

1 Introduction

Unmanned aviation systems (UASs) in the last decades have become actively used as a specific accessible means of solving many civil and military problems [1, 2]. In addition to traditional surveillance, aerial photography, and reconnaissance, they are used [3] to solve many new problems. These include:

- radio and video broadcasts;
- agricultural lands cultivation;
- people and vehicles search and rescue;
- goods delivery.

In recent years there has been rapid progress in the design and manufacture of more sophisticated airborne sensors for unmanned aerial vehicles. This opens up the new opportunities for environmental monitoring, reconnaissance, subsequent elimination of the flood and earthquake, and specific military tasks as well.

It is also significant that the UAS capabilities mentioned above are accompanied by a simultaneous price decline for small UASes. This, in turn, makes them more accessible [4] due to the following benefits:

- they are capable of hanging or loitering for long periods at the desired height in the desired location;
- they do not need much time to deploy, they can be launched quickly;
- The UAS equipment is capable of transmitting images with better resolution and better quality than satellites;
- the practical flight altitudes of such UASes are lower than the typical altitudes of conventional air traffic;
- they can get close to the points of interest users need.

The tasks of aviation reconnaissance have always been solved since the aviation advent. Before the modern information and communication technologies (ICT) have appeared, aviation reconnaissance information was traditionally obtained through post-processing of aerial photographs [5], but this creates time delays and reduces operative efficiency.

The development of information and communication technology (ICT) in the late XX - early XXI centuries has led to a significant increase in the pace of events in many fields, especially in the military field. Moreover, the success of a many mission becomes critically dependent on the speed of decision-making. This, in turn, is determined by the speed of obtaining reliable input information. As a result, this has led to an increasing need to receive from UASes, along with the traditional types of service listed above, also operational information about ground targets coordinates (GTC), which are detected directly through aerial video surveillance and operational target detection in the unmanned aerial vehicle (UAV) videocamera's field of view.

In particular, the need for such operational determination of the coordinates of the targets arises when solving the following practical problems [6]:

- the defining targets for military needs, for example, to control artillery fire;
- the search for survivors and aviation flight recorders after aviation accidents;
- the inspection and search operations in the wildlands and at the sea;
- the spot spraying the dangerous weeds in agriculture by pesticides.

2 Target Setting

When performing a typical UAV flight as part of a UAS there interact coordinatively (see Fig. 1):

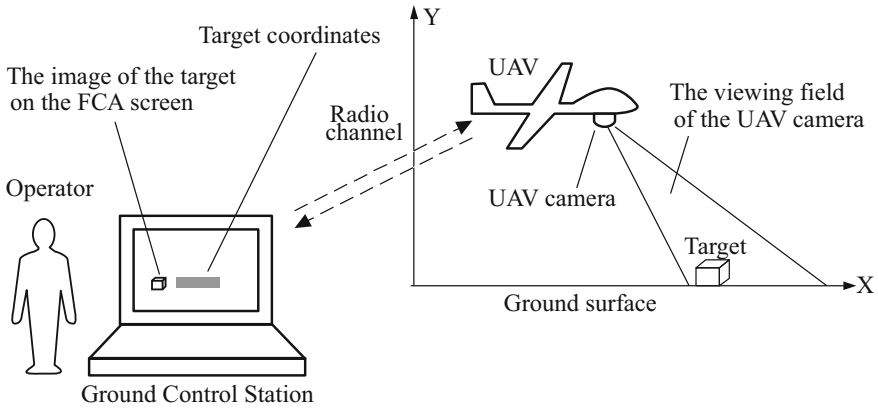


Fig. 1. General scheme of operational determination of the target’s coordinates in the process of reconnaissance UAV flight performing.

- an unmanned aerial vehicle equipped with a video camera;
 - a ground control station (GCS) which is radio-connected to the UAV. Flight Control Application (FCA) has been started and is active on GCS;
 - an operator (ground-based external pilot) who, through the GCS, controls the execution of the UAV flight and issues requests for coordinates of targets which he observes on the FCA screen.
1. The UAV is started. During the flight, the UAV camera transmits in real-time to the GCS an image of the ground surface that the operator sees in the flight control application window.
 2. During video surveillance, the operator sees a target – an object of interest on the ground.
 3. The operator with the mouse cursor clicks on a target image on the FCA screen.
 4. In response to the operator’s action, the GCS through radio channel sends a request to the UAV to transmit information about its current position in space, that is, its linear and angular coordinates.
 5. The on-board equipment of UAV (autopilot and inertial measurement unit, IMU) will read the attitude parameters of the camera from the GPS and angle sensors. An appropriate data packet is generated by onboard software and is sent to the GCS by the radio channel.
 6. At GCS, the FCA application calculates the coordinates of the ground target, taking into account:
 - a. the spatial information received from the UAV;
 - b. the optical parameters of the camera lens;
 - c. the screen coordinates of the place where (see step 3) the operator clicked the mouse.

7. The calculated coordinates are displayed on the screen of the operator, which determines the way of their further use.

The described actions are formally explained by the sequence diagram in Fig. 2.

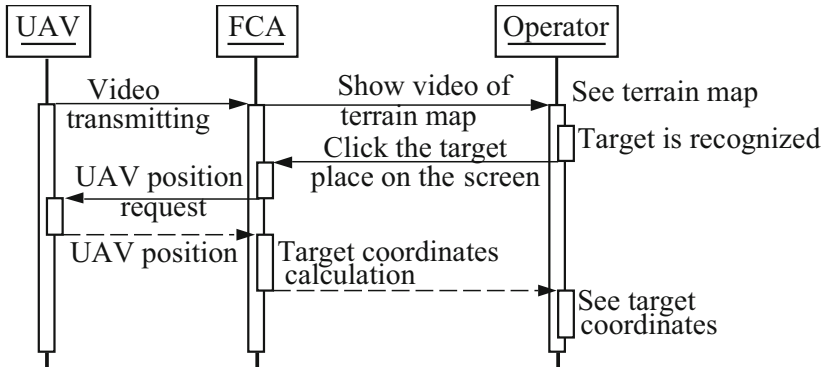


Fig. 2. The UML sequence diagram of operational determination of the target’s coordinates

The practice of using UAS shows that the accuracy of determining the coordinates of a target according to the methodology described is often insufficient. The reason for this is the organizational and technical complexity of the described technology, each component of which contributes to the inaccuracy of the final result. However, the quantitative assessment of the degrees of influence of the individual components and process factors shown in Fig. 1 and Fig. 2 on the accuracy of determining the coordinates of the targets was not performed yet. Further refinement and development of new UAS models require the knowledge of the relative impact on the accuracy of GTCs of the various UAS components and which ones need to be given priority in the further development.

Decomposing the task of defining a GTC using UAS by its logic requires the definition at least two groups of parameters:

1. Spatial and angular coordinates of the UAV’s video camera;
2. The relative coordinates of the target relative to the UAV camera.

Below is an analysis of the UAS subsystems which have influence to obtaining sufficient accuracy to solve the formulated subtasks.

3 Actual Scientific Research and Issues Analysis

3.1 The Precision of the Gyro Stabilized Platforms Positioning

At a qualitative level, it is quite obvious that the stability of the video image on the GCS depends primarily on the stability of the angular orientation of the camera during

the shooting process. In turn, stable video on the GCS screen makes it easy for the operator to identify the target visually and mark it with the mouse cursor.

An important means of enhancing the stability of images transmitted by cameras was their placement on gyro-stabilized platforms [7] (gimbals).

The gyro-stabilized camera's gimbal (Fig. 3) has its own inertial measurement unit (IMU), which includes electromechanical micro gyroscopes, among others. The output signals of these gyroscopes are proportional to the angular velocities of the camera rotation around the UAV's axes of the pitch, roll, and yaw.

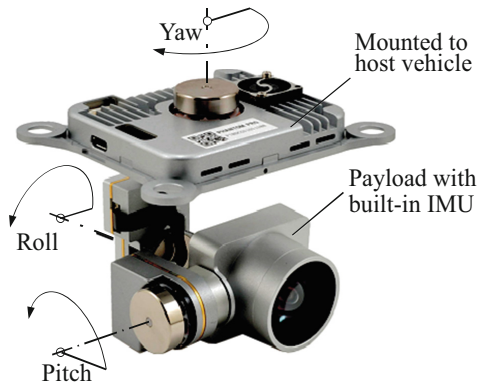


Fig. 3. Video camera on the gimbal [8]

The stability of the video image transmitted by the camera is achieved at relatively stable angular coordinates of the camera relative to the external coordinate system (ground). Since the IMU gyroscopes produce angular velocity signals, and the aim is the stability of angular coordinates, the gyroscopes outputs must be integrated over time. This function in the electromechanical gimbal system is performed by integrated motors with axes of rotation, which correspond to rotations in the pitch, roll, and yaw. The angles of rotation of the fuselage UAV are compensated by the gimbal's motors, which ensures the stability of the orientation of the camera.

Although such gimbals provide [8] precision for controlling the angular position of the camera within 0.02° , the "horizon loss" remains a technical problem for them. This phenomenon is manifested in the fact that during the gimbal work integration errors gradually accumulate and the actual position of the "horizontal" plane of the camera receives a certain deviation from the horizontal.

This disadvantage is partially compensated by the horizontal alignment calibration procedure performed immediately before the UAV flight, but it is not entirely possible to eliminate this deficiency in common gimbal models.

However, the practical accuracy of holding the horizon in the process of one flight with an electromechanical gimbal system within the deviation of $1^\circ \dots 3^\circ$ is quite real [8].

3.2 Accuracy of the Geo-Positioning Satellite Navigation System

The first and most important component of the GTC is UAV coordinates at the time of their calculation.

The most commonly used way to determine your spatial coordinates for UAVs was and still is to use geo-positioning satellite navigation systems [9] like GPS NAVSTAR, SRNS GLONASS, and others.

According to [9], the accuracy of the coordinates determination in the open field at a single-frequency GPS signal reception by different geo-positioning satellite systems can have the values shown in the Table 1.

Table 1. The values of the accuracy of coordinates determining by the different geo-positioning satellite systems

Navigation system	GPS NAVSTAR	SRNS GLONASS	BDS COMPASS
Accuracy	5 m (without differential global positioning system, DGPS)	4,5 m–7,4 m (without DGPS)	10 m (Open sygnal)

The accuracy of coordinate determination by these systems depends [10] on:

- the surrounding area (in the city, among high-rise buildings, the accuracy is worse than in the open field);
- the number of satellites whose signals are available to the UAV GPS receiver at the current time. Typical preferred number of satellites is at least 5, optimal is at least 8;
- the use or non-use DGPS;
- whether the GPS UAV receiver can or cannot receive two-frequency (L1 + L2 or L1 + L5) satellite navigation signals;
- sufficient satellite signal reception time. With insufficient satellite reception time, the coordinates will have unreliable values.

If necessary, to improve the accuracy of GPS coordinates, the users may apply the following means.

1. The most radical, but rather costly, measure to date was the additional use of differential position correction (DGPS). Normal GPS accuracy is 2–20 m, DGPS application provides accuracy 1–5 m, but an additional limitation is the distance of the object from an anchor point with well-defined coordinates. As this distance increases, accuracy decreases.
2. The second approach available to the user is to switch from receiving GPS signals on one carrier frequency (single-frequency reception, L1) to receiving signals GPS using two carrier frequencies (dual-frequency reception, L1 + L2 or L1 + L5, signal denotation presented in according to [11]). Such a transition provides [12] an increase in the accuracy of coordinate determination to the values of 0.2... 0.5 m, but it is related to the modernization of the radio equipment of the GPS UAV receiver.

Since the upgrading of the UAV equipment (item 2) is less costly than the organization of anchor points (item 1), especially in the unfamiliar site, the transition to dual-frequency GPS reception has recently become widespread.

4 The Accuracy Analysis of the Ground Targets Determination Using the UAV Video Camera

To assess the degree of influence on the accuracy of the GTC determination the camera geometric factors and the application screen of the FCA, let us consider the diagram of the relative location of the camera and the on-surface target at the moment of determining its coordinates (Fig. 4). The image generated by the lens on the camera's M sensor is displayed on the FCA window by the UAS.

Figure 4 shows: M – the light-sensitive pixel matrix (video sensor) of the camera; F – the focus of the camera lens; d – the focal distance of the camera lens; γ – the angle half of the viewing area of the camera; G – ground target; H – the current flight altitude of the UAV; GS – the surface of the earth; Δ – elevation correction according to the elevation map; α – the view angle of the camera relative to the horizontal; β – the target angle relative to the camera's view ray; C_S – the center of the FCA window; C_C – the center of the camera sensor; T_S – the location of the target point on the FCA screen; T_C – the projection of the target onto the camera matrix; h_S – the height of the FCA window in pixels; h_M – the height of the camera matrix in pixels.

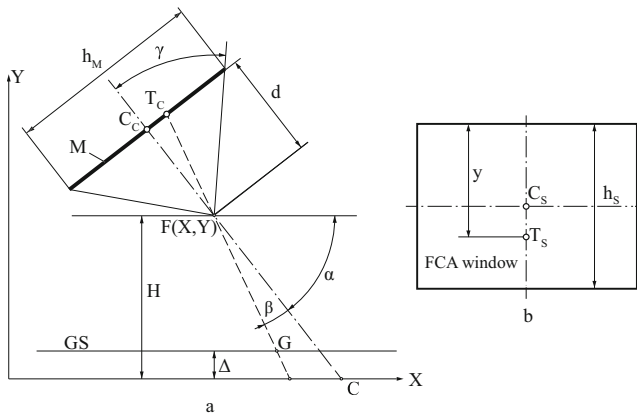


Fig. 4. Scheme for the accuracy assessment for determining the coordinates of a target by using UAS. **a** – the position of the camera relative to the ground target G ; **b** – the position of the T_S target's image in the flight control application window

Let's make some assumptions to simplify the analysis.

1. With small changes the camera angles of pitch and yaw, the offset of the camera's line of sight on the Earth's surface by course and yaw will be values although

slightly different, but in the same order. Thus, to determine the influence degree on the accuracy of the GTC estimation the camera angular orientation accuracy, it is sufficient to consider only a vertical section of the scene (pitch plane), which significantly simplifies the mathematical descriptions.

2. Assume that the camera lens has no optical distortion. However, for further analysis, it's possible to take into account that according to [13], at the UAV flight altitude of 300–400 m, the distortion error of determining ground coordinates from an aerial photograph of the Sony RX-1 camera was 12–15 cm.
3. Assume that the image on the camera's sensor M (Fig. 4) is similar to the scene in the FCA window. This condition is true when using only affine transformations of the first image to the second, which is used in almost all computer graphics systems.

Suppose (Fig. 4) that the target G has a projection on the sensor M at the point T_C , and then the latter was displayed on the FCA window at the point T_S . Then, assuming assumption 3, it's possible to write:

$$\frac{T_C C_c}{\frac{h_M}{2}} = \frac{d \cdot tg\beta}{d \cdot tg\gamma} = \frac{y - \frac{h_S}{2}}{\frac{h_S}{2}}, \tag{1}$$

that is

$$tg\beta = tg\gamma \frac{2y - h_S}{h_S}. \tag{2}$$

Then the coordinate X of the target G in the coordinate system XY will be equal

$$X_G = X_F + (H - \Delta) \cdot ctg(\alpha + \beta(y)). \tag{3}$$

To solve the problem posed in this study, it is necessary to estimate the specific values of the partial derivatives $\frac{\partial X_G}{\partial \alpha}$ and $\frac{\partial X_G}{\partial y}$, and after this compare them with the accuracy of the GPS positioning. The partial derivative $\frac{\partial X_G}{\partial \alpha}$ may be a sensitivity measure for the inaccuracy of determining the angular position of the camera, and the partial derivative $\frac{\partial X_G}{\partial y}$ will allow to estimate the influence of inaccuracy of specifying by the operator the target's location on the FCA screen. This will provide guidance for the UAV developer concerning which component may be most promising in terms of improving the accuracy of GTC determination.

From formula (3) it can be seen that the derivatives $\frac{\partial X_G}{\partial \alpha}$ and $\frac{\partial X_G}{\partial \beta}$ can be found in the same expression

$$\frac{\partial X_G}{\partial \alpha} = \frac{\partial X_G}{\partial \alpha} = \frac{H - \Delta}{\sin^2(\alpha + \beta)}, \tag{4}$$

allowing you to rate one of them.

Given the practical orientation of the task of this work, the method of approximate numerical differentiation (3) was chosen for the situation of a typical UAV flight at an

altitude of 600 m. Two variants of the camera were evaluated – with a regular lens ($\gamma = 30^\circ$) and with a long-focal lens ($\gamma = 5^\circ$), and two variants of finding the UAV relative to the target – on approach ($\alpha = 50^\circ$) and just above target ($\alpha = 90^\circ$). The results of the calculations are given in Table 2.

Table 2. The offset of the calculated CPCs with small changes in the parameters γ and α

Camera lens	Regular $\gamma = 30^\circ$		Long-focal $\gamma = 5^\circ$	
UAV position	$\frac{\partial X_G}{\partial \alpha}$, m/degree	$\frac{\partial X_G}{\partial \gamma}$, m/pixel	$\frac{\partial X_G}{\partial \alpha}$, m/degree	$\frac{\partial X_G}{\partial \gamma}$, m/pixel
On approach $\alpha = 50^\circ$	15,8	0,97	17,7	0,16
Above target $\alpha = 90^\circ$	10,39	0,64	10,39	0,1

Table 2 shows that the error of target's image positioning 1–3 pixels (point TS, Fig. 4) on the FCA screen by computer mouse with usual sensitivity 400–600 dpi gives the ground target offset 0.1 ... 0.97 m/pixel. This value is at least at the order of magnitude smaller than the error in determining the GTC due to the error of the angular position of the camera (10.39... 17.7 m/degree). The “by pixel” error has the same order of magnitude as the GPS positioning with using DGPS, especially with dual-frequency receivers (positioning error 0.1... 0.2 m).

The use of telephoto lenses or the “mark the target right above it” tactics do not give a special gain. However, there remains a practical approach whereby the detection of a target is more accurate, the closer the camera is to it, which is not always acceptable in military applications.

5 Conclusions

The made calculations show that at this stage of UAS development to improve the accuracy of the operational GTC determination rationally worth to focus, first of all, on improving the camera angular coordinates accuracy at the moment of the target detection. This is most effective for military UASes, and for civilian applications it's possible additionally recommend tactics of loitering with reducing flight UAV's altitude to refine the GTC.

The recommended improvement in precision of camera attitude angles can be achieved by using more accurate inertial orientation systems of gyro stabilized UAV camera's gimbals and their mandatory pre-flight calibration relating to horizontality.

References

1. Herlik, E.: Unmanned Aerial Vehicles (UAVs) for commercial applications global market & technologies outlook 2011–2016. Technical report, Market Intel Group LLC (2010)
2. Dworkin, A.: Drones and Targeted Killing: Defining a European Position. European Council on Foreign Relations (2013). http://www.ecfr.eu/page/-/ECFR84_DRONES_BRIEF.pdf. Accessed 07 Mar 2020

3. Chen, H., Wang, X., Li, Y.: A survey of autonomous control for UAV. In: International Conference on Artificial Intelligence and Computational Intelligence, AICI 2009, vol. 2, pp. 267–271 (2009)
4. Sun, J., Li, B., Jiang, Y., Wen, C.-Y.: A camera-based target detection and positioning UAV system for search and rescue (SAR) purposes. *Sensors* **16**, 1778 (2016)
5. Barber, M.: *A History of Aerial Photography and Archaeology*. Historic England (2011)
6. Hinas, A., Ragel, R., Roberts, J., Gonzalez, L.: A framework for multiple ground target finding and inspection using a multirotor UAS. *Sensors* **20**, 272 (2020). <https://doi.org/10.3390/s20010272>
7. Tischler, P., Glynn, K.: Stabilized Gimbal System. https://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/f2013/kjg58_pmt43/website/website/index.html. #software_design. Cornell University ECE 4760 Final Project (2013). Accessed 05 Mar 2020
8. DJI Phantom 3 Part #5 4K Camera w/Gimbal for P3 Professional. <https://www.worthpoint.com/worthopedia/dji-phantom-part-4k-camera-gimbal-p3-1813371183>. Accessed 12 Mar 2020
9. Satellite Navigation. https://en.wikipedia.org/wiki/Satellite_navigation. Accessed 13 Mar 2020
10. Antonovich, K.M.: The use of satellite radio navigation systems in geodesy. In: Monograph of SEE VPO “Siberian State Geodetic Academy”, vol. 2. FSEE “Cartgeocenter”, Moscow (2006)
11. New Civil Signals. <https://www.gps.gov/systems/gps/modernization/civilsignals>. Accessed 15 Mar 2020
12. Global Navigation Satellite System (GNSS). http://www.nngasu.ru/geodesy/classification/chastnye-klassifikatsii/10_GNSS.php. Accessed 13 Mar 2020
13. Makarov, V., Bondarenko, D., Makarov, I., et al.: Unmanned aerial vehicles for solving the tasks of surveying and monitoring of open pit mining. <https://uav-siberia.com/news/bspilotnye-letatelnye-apparaty-dlya-resheniya-zadach-marksheyderii-i-monitoringa-otkrytykh-gornyykh/>. Accessed 15 Mar 2020



Unmanned Aerial Vehicle Mass Model Peculiarities

Dmitriy Kritskiy^(✉) , Sergiy Yashin , and Sergiy Koba 

National Aerospace University Kharkiv Aviation Institute,
Chkalova Street, 17, Kharkiv 61070, Ukraine
d.krickiy@khai.edu

Abstract. The article discusses the features of the mass balance of an unmanned aerial vehicle (UAV) and the mass of UAV safety equipment; the equation of the UAV mass balance and the dependence of the UAV take-off mass on the known loads and the relative masses of the UAV parts are obtained taking into account safety and airworthiness; The main systems designed to replace the pilot on board an unmanned aerial vehicle are considered; The dependence of the total mass of the apparatus on the masses of means ensuring safe operation is demonstrated; the calculations of several options for the mass balance depending on the purpose of the unmanned aerial vehicle; the dependence of the thrust-weight ratio, flight range, maximum speed and mass of an unmanned aerial vehicle of civilian use is considered in a first approximation; the developed mass balance model was tested in the project of creating the Sparrow unmanned aerial vehicle at the National Aerospace University, Kharkov Aviation Institute.

Keywords: UAV · UAV take-off mass · UAV mass balance equation · Aviation equipment · Existence equation

1 The Problem

Mass model in CAD systems provides mass calculation and mass report. The level of report details is determined by the project stage. The model is based on relationships between the geometry of the aircraft, the loads acting on it, the features of the power point, the magnitude of the target load, the composition and placement of equipment on the one hand, and the mass of the aircraft and its components on the other [1].

Mass and size are the basis of all engineering structures design and construction. They become even more important in aviation. The high efficiency of the aircraft is achieved mainly by reducing the mass of the structure, power point, equipment systems, as well as the high density of its layout, i.e. maximum use of volumes. Therefore, mass related issues have always been among the main design challenges [2].

The importance of this problem has increased not only due to the further airframe mass decreasing issues, but in addition to ever wider application of optimization methods, the main criteria of which are mass and economics. Particular attention to mass issues is also explained by the fact that the overweighting of aircraft is an inevitable reason for the failure to fulfill the basic flight characteristics. That is why

everyone recognized that the main and most difficult task of creating a modern aircraft is the need to keep its initial equipped mass from increasing, on the basis of which all its design characteristics are formed [2].

The absence of a pilot on board defines the main problem of civil unmanned aerial vehicles (UAV) as the problem of the airspace's legal and safe usage.

According to the current Air Code of Ukraine and the Aviation Rules of Ukraine, part 21, despite the great potential of UAVs, they cannot be allowed to fly in the Ukrainian airspace for commercial purposes.

To solve this UAV issue, it is necessary to add safety features and functionally compensate the absence of a pilot on board. In the available literature on UAV design [3–5], the issues of UAVs mass design and analysis were not reflected.

The complexity of the mass problem is caused by the need to achieve extremely high aircraft payload and productivity.

2 Main Part

The take-off mass of the aircraft and its parts are the most important parameters that integrally describe all the properties of the aircraft, that is why these parameters have found a significant place in CALS technologies [6, 7].

According to V.F. Bolkhovitinov, to give the aircraft any property, it is necessary to spend a certain amount of material. This amount of material is inconsistent and depends on the device's technology level, which is used to achieve this property.

So the flight range at a constant fuel consumption per kilometer C_k is directly proportional to the mass of fuel m_F , i.e.

$$L = \frac{1}{C_k} m_F. \quad (1)$$

The aircraft property - the ability to fly – is proportional to fuel mass onboard and aircraft's degree of perfection $\frac{1}{C_k}$, i.e. technology level. As you can see, mass is the equivalent of product quality. This is a kind of “money”, with which we pay for any property attached to the product. The mass of the entire aircraft can be represented by the sum of the individual masses, which provide a certain composition and properties required for the successful aircraft tasks completion. Thus, the mass of the entire aircraft and its structure quantitatively determine the set of properties and the relationship between them.

Typically, the mass of an aircraft is considered as the sum of several terms, which are either the mass of aircraft parts or its payload. Therefore, the mass of an aircraft can be represented as the mass of its structure (m_P), propulsion system (m_M), fuel, payload (m_N), i.e.

$$m = m_P + m_M + m_F + m_N. \quad (2)$$

This equation expresses the aircraft mass balance. If on the right side of this equation we express the mass of terms using the corresponding aircraft properties, we obtain an equation relating the aircraft mass to its various properties.

Such an equation, combining the properties of an aircraft and, moreover, in certain quantities, will no longer be a mass balance equation, but an equation for the aircraft existence. It will display not only the relationships of various properties and qualities of the aircraft, but also the possibility of implementing them in certain quantities in one aircraft.

Conclusions from the existence equation:

1. The ability to provide a set of properties has a historical character, since it depends on the mass, which in its turn depends on the technology's perfection level.
2. The aircraft could be created for the first time only at a certain stage in the development of mass excellence, which made it possible to maintain the mass balance while keeping the necessary minimum flight performance.
3. For a given take-off weight, an aircraft can have only a certain, limited number of properties.
4. For a given take-off mass, each aircraft property cannot exceed a certain limiting value.
5. All aircraft properties are quantitatively interconnected in the sense that an increase in any of them leads to decrease in others for a constant take-off mass.
6. The aircraft specialization leads to the redistribution of mass costs, i.e. to changes in the aircraft weight structure.
7. An increase in the properties volume attributed to the aircraft, all other things being equal, leads to an increase in its take-off mass.

The reasons for the mass structure constancy: strict interdependence of the aircraft structures properties and masses; stable requirements range for each aircraft type and class; stable values of certain relative masses caused by special conditions; mathematical features of the relative quantities very nature.

Let's find an analytical expression for the aircraft existence equation.

Airframe mass depends on the total mass of the aircraft (the greater the mass of the aircraft m , i.e. the mass that the airframe will have to carry, the greater will be the mass of the airframe), on calculated overload (the greater the calculated overload n , the larger the construct elements sections will be, therefore, the greater the mass of the airframe), on wing elongation (the greater the elongation λ , the greater bending moments, the smaller the chord, the smaller the section height, the greater the section of the wing force elements, therefore, the greater the mass of the airframe), on relative wing profile thickness c , on contraction η etc. As a result

$$m_P = F(m, n, \lambda, \bar{c}, \eta, \dots). \quad (3)$$

Similarly, the mass of the propulsion system is a function of the aircraft mass, specific mass and specific "forehead" of the propulsion system (γ), maximum horizontal speed (V_{\max}), etc., i.e.

$$m_M = \Phi(m, \gamma, V_{\max}, \dots). \quad (4)$$

The mass of the fuel system depends on the aircraft mass, the specific consumption of the engine (c_p), altitude (H_{cr}) and speed (V_{kp}) at cruising mode, quality (K), aircraft range (L), etc. Hence,

$$m_F = \Psi(m, c_p, V_{kp}, H_{cr}, K, L, \dots). \quad (5)$$

The mass of the full load is composed of the crew mass (m_{pil}), weapons (m_w) and equipment (m_{eq}) that we want to load into the aircraft, i.e.

$$m_n = \mathcal{G}(m_{pil}, m_{eq}, m_w, \dots). \quad (6)$$

Using the obtained expressions, we can formulate the next equation

$$m = F(m, n, \lambda, \bar{c}, \eta, \dots) + \Phi(m, \gamma, V_{\max}, \dots) + \Psi(m, c_p, V_{kp}, H_{cr}, K, L, \dots) + \vartheta(m_{pil}, m_{eq}, m_w, \dots). \quad (7)$$

The resulting equation is the equation for the aircraft existence, since it shows which properties and in what quantities are available in an aircraft.

But in this form, the existence equation is inconvenient. Although it expresses the absolute aircraft mass, it does not allow to compare different aircrafts due to differences in their masses, powers, resistances, etc. It is necessary, without abandoning the meaning and essence of the aircraft existence equation, to find another form of it, bringing all aircrafts to the same scale and therefore allowing them to be compared.

If we divide both sides of the existence equation by the aircraft mass, we obtain the equation in the form of relative masses

$$1 = \bar{m}_p + \bar{m}_M + \bar{m}_F + \bar{m}_N, \quad (8)$$

$$\bar{m}_p = \frac{m_p}{m} = f(n, \lambda, \bar{c}, \eta, \dots),$$

$$\bar{m}_M = \frac{m_M}{m} = \phi(\gamma, V_{\max}, \dots), \quad (9)$$

$$\bar{m}_F = \frac{m_F}{m} = \psi(c_p, K, L, H_{cr}, V_{\max}, \dots), \quad (10)$$

$$\bar{m}_N = \frac{m_N}{m} = \vartheta\left(\frac{m_{pil}}{m}, \frac{m_{eq}}{m}, \frac{m_w}{m}, \dots\right). \quad (11)$$

In this equation, the properties of an aircraft are relative to a kilogram of its mass, so it can be used to compare aircrafts with each other. This is achieved by comparing the degree of mass unit usage, i.e. by the degree of their perfection.

In this form, the existence equation is convenient for analysis and for evaluating both existing and future aircrafts.

The meaning of this equation is that any aircraft property will be associated with the mass. This determines the condition for the quality existence in the required quantity.

The mass balance equation is written as:

$$m_0 = m_{ee} + m_F + m_{up}, \quad (12)$$

where m_0 – take-off mass, m_{ee} – empty equipped mass, m_{up} – useful payload mass. If we decompose m_{ee} into components, we will get:

$$m_0 = m_k + m_{pp} + m_{ce} + m_{om} + m_F + m_{up}, \quad (13)$$

where m_k – airframe structure mass, m_{pp} – power point mass, m_{ce} – control equipment mass, m_{om} – overload mass.

$$m_{om} = m_{erc} + m_{eum}, \quad (14)$$

where m_{erc} – mass of equipment replacing crew, m_{eum} – equipment and undeveloped fuel residue masses. Dividing both sides of (13) by m_0 , we obtain the equation of relative masses:

$$1 = \bar{m}_k + \bar{m}_{pp} + \bar{m}_{ce} + \bar{m}_F + \frac{m_{erc} + m_{eum} + m_{up}}{m_0}, \quad (15)$$

where $\bar{m}_k = \frac{m_k}{m_0}$, $\bar{m}_{pp} = \frac{m_{pp}}{m_0}$, $\bar{m}_{ce} = \frac{m_{ce}}{m_0}$, $\bar{m}_F = \frac{m_F}{m_0}$ – the relative masses of the structure, power point, control equipment and fuel respectively. Relative masses m_k, m_{pp}, m_{ce} have corresponding dependencies on the airframe units parameters and technical specifications in the form of weight formulas and statistical data. The fuel relative mass m_F depends on the range, flight duration, engine performance, flight mode, aerodynamic properties and is determined by formulas derived from the Breguet range formula. The masses of the crew, equipment and payload are given values and expressed in absolute values.

The manned aircraft flight safety and its airworthiness property are determined by the crew presence and design decisions that meet the requirements of the Airworthiness Standards. The presence of these properties is fundamentally reflected in the next masses m_k, m_{pp}, m_{ce} [8–10]. This means that decreasing these values will reduce the reliability, safety and airworthiness of the aircraft.

Let's proceed with the components of UAV take-off mass. According to the research results, to ensure the UAV safety, the following devices must be included in its composition:

- aeronautical lights with a flashing beacon having mass m_{al} and a radar transponder of an active response having mass m_{rad} for providing visual and radar visibility;
- a container with an air conditioning system to ensure the stable operation of electronic units, having mass m_{cc} ;
- an automatic pilot to provide functional compensation for onboard pilot, having mass m_{ap} , and consisting of three subsystems - a flight control system, a system for maintaining flight modes and a security system;

- a payload system to ensure the principle of aviation security as security against payload, having mass m_{ssp} .

These systems can be used in various combinations or all together depending on the purpose of the UAV and the area where it will be used [11]. UAVs layout and design differ from the ones of the manned vehicle because of the crew removal and operational components associated with it. Also the installation of additional devices that provide functional pilot compensation and safety are required for UAV.

The UAV mass balance equation can be written as:

$$m_0 = m_k + m_{pp} + m_{al} + m_{rad} + m_{cc} + m_{sps} + m_{ap} + m_{ssp} + m_F, \quad (16)$$

where m_{sps} – the mass of all on-board systems power supply. It becomes necessary to characterize each component of the take-off mass. In the first approximation the m_{sps} design of the UAV experiences the same loads during the flight as the manned aircraft. However, UAVs are subject to loads that the manned aircraft does not have. These loads are: ejection launch load, parachute landing and uncoordinated steering deviations in case of control failures. Due to this, the relative mass of the UAV design \bar{m}_k will have a slightly greater value than that of manned aircraft of the same mass category.

The operation of a power unit on a UAV does not fundamentally differ from operation on a manned aircraft. That is why the relative masses of power unit \bar{m}_{pp} and fuel \bar{m}_F will be determined the same way for both UAV and manned aircraft.

The masses of aeronautical lights m_{al} and the radar transponder m_{rad} are independent from the take-off mass.

The power supply system feeds all consumers of an empty equipped UAV. Due to the fact that UAVs have more of these consumers, the relative mass \bar{m}_{sps} for UAVs will have bigger value than for manned aircraft.

The payload system consists of the payload and the devices ensuring its operation - the payload management system, the radio data transmission system and the power supply system. The payload mass m_{up} is defined by the UAV purpose and does not depend on the take-off mass. The devices mass that provide payload operation depends on the payload power consumption, range and flight duration, i.e. from m_{up} and m_F , therefore it can be approximately defined as: $m_{ssp} = k_1 m_{up}$, where k_1 payload weight gain, $k_1 > 1$ and, in the future, it is also necessary to define the functional dependence for k_1 .

The mass of all automatic pilot systems m_{ap} can be represented as the sum of devices masses, which does not depend on take-off mass m_{dmn} and devices whose mass depends on take-off mass m_{dmt} : $m_{ap} = m_{dmn} + m_{dmt}$.

The devices mass m_{dmn} is determined by the UAV purpose and its application areas. The devices mass m_{dmt} includes the steering drives masses m_{sd} , and the parachute system for safe termination of flight and energy damping m_{safe} , which can be expressed in take-off mass fractions m_{sd} and m_{safe} .

All electronic components must be placed in a container with artificial climate. The air conditioning system mass depends on the container volume, and therefore on the masses sum: $m_{cc} = k_2(k_1 m_{up} + m_{rad} + m_{dmn})$, where k_2 – the container with the air

conditioning system relative mass coefficient according to the weight of the units to be conditioned, this coefficient must be determined in the future.

Now Eq. (16) can be written as:

$$m_0 = m_k + m_{pp} + m_{al} + m_{rad} + m_{sps} + m_{dmn} + m_{sd} + m_{safe} + k_1 m_{up} + m_F + k_2 (k_1 m_{up} + m_{rad} + m_{dmn}). \quad (17)$$

By grouping the take-off mass components into dependent and independent from m_0 and dividing them by m_0 , we will get

$$1 = \bar{m}_k + \bar{m}_{pp} + \bar{m}_{sps} + \bar{m}_{sd} + \bar{m}_{safe} + \bar{m}_F + \frac{m_{al}}{m_0} + \frac{(1 + k_2)(k_1 m_{up} + m_{rad} + m_{dmn})}{m_0}. \quad (18)$$

By solving the Eq. (18) relatively to m_0 , we will get:

$$m_0 = \frac{m_{al} + (1 + k_2)(k_1 m_{up} + m_{rad} + m_{dmn})}{1 - (\bar{m}_k + \bar{m}_{pp} + \bar{m}_{sps} + \bar{m}_{sd} + \bar{m}_{safe} + \bar{m}_F)}. \quad (19)$$

Thus, the UAV mass balance Eq. (18) and the UAV take-off mass dependence on the masses of known cargoes and the relative masses of UAV parts equation are obtained taking into account safety and airworthiness (19).

To clarify the effect of safety and airworthiness devices on the UAV take-off mass, let us see how Eq. (19) will look in the complete absence of safety devices, but with the radio control having mass m_{con} , i.e. $m_{al} = m_{rad} = m_{dmn} = 0$; $k_1 = 1$; $k_2 = 0$; $\bar{m}_{safe} = 0$. In the result we will get:

$$m_0^* = \frac{m_{up} + m_{con}}{1 - (\bar{m}_k + \bar{m}_{pp} + \bar{m}_{sps} + \bar{m}_{sd} + \bar{m}_F)}, \quad (20)$$

where m_0^* - radio-controlled UAV take-off mass that does not include safety devices.

It is obvious that the safety devices usage significantly increases the take-off mass. As an example, let's consider the effect of these devices on the hypothetical UAV take-off mass with the following parameters close to real: $\bar{m}_k = 0.35$; $\bar{m}_{pp} = 0.12$; $\bar{m}_{sps} = 0.03$; $\bar{m}_{sd} = 0.22$; $\bar{m}_F = 0.22$; $\bar{m}_{safe} = 0.05$; $k_1 = 2$; $k_2 = 0.2$; $m_{up} = 1$ kg.

Radio-controlled UAV take-off mass that does not include safety devices $m_{con} = 0.5$ kg; will be equal to: $m_0^* = \frac{0.5+1}{1-(0.35+0.12+0.03+0.03+0.22+0.22)} = 6$ kg.

When installing parachute system $\bar{m}_{safe} = 0,05$, with radio control, we get: $m_0^* = \frac{0.5+1}{1-(0.35+0.12+0.03+0.03+0.22+0.05)} = 7,5$ kg.

Next we should isolate the payload and its control from the onboard UAV systems $k_1 = 2$, i.e. we get: $m_0^* = \frac{0.5+2\cdot1}{1-(0.35+0.12+0.03+0.03+0.22+0.05)} = 12,5$ kg.

Let's additionally apply electronic components conditioning $k_2 = 0,2$: $m_0^* = 1,2 \frac{0.5+2\cdot1}{1-(0.35+0.12+0.03+0.03+0.22+0.05)} = 15$ kg.

Thus, when you consider that an unmanned aerial vehicle is launched from a catapult, as well as that a civil unmanned aerial vehicle flies under conditions in a standard atmosphere. Having performed the calculations, we obtain the dependences of the flight range on the fuel supply at different specific loads (Fig. 1).

Thus simulating the composition of a civil unmanned aerial vehicle (assigning flight ranges, maximum speeds) using the obtained formula, we obtain mass dependences, as well as all weight components of the drone. Modeling the calculated flight ranges from 200 to 300 km and maximum speeds from 50 to 80 m/s, we get the mass of the device from 22 to 35 kg.

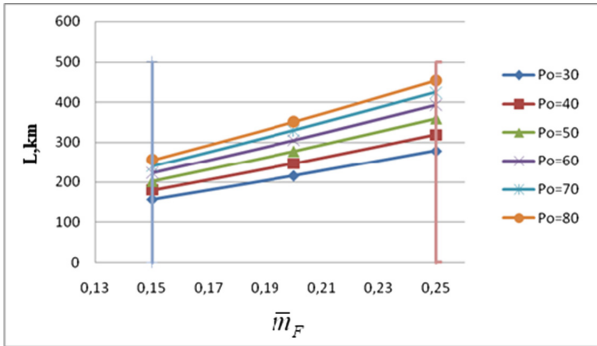


Fig. 1. The graph of the dependence of the estimated flight range L on the relative fuel reserve m_F at different specific loads P_0 .

3 Conclusion

Thus, adding safety properties to UAV leads to an increase in its take-off mass. In the same time the mass balance equation will have the following features:

- UAV mass balance structures without using safety devices and using them differ significantly;
- to ensure the UAV safety and introduce it into the airspace, it is necessary to use additional devices to increase the safety level, which significantly increase its take-off mass;
- Equation (19) shows that the safety can be achieved in several ways:
 - additional devices usage, which increase the UAV take-off mass;
 - improvement of additional devices to achieve the same effects with less take-off mass growth;
 - improvement of additional devices and design, engines, energy supply devices to achieve safety effects with less take-off mass;
 - aircraft design and its systems improvement with a slight increase in take-off mass to increase reliability;

- comprehensive improvement of all components ensuring flight safety, by clarifying the functional dependencies \bar{m}_k , \bar{m}_{sp} , \bar{m}_{sd} and \bar{m}_{safe} , k_1 and k_2 , on the specifications, geometry and technical characteristics.

The practical value of the UAV mass balance equation is that it allows to more accurately determine the take-off mass in all approximations during the design stage and shows ways to control the take-off mass.

The UAV mass balance equation and the dependence of the UAV take-off mass on the masses of known cargoes and the relative masses of UAV parts, taking into account safety and airworthiness, are an integral part of the model describing the mass in CAD systems when designing UAV. This data is necessary when describing the life cycle of an aircraft using CALS technologies.

An interesting option to consider is the design with reduced effective straining and, accordingly, with some increase in the UAV structure mass, which makes it possible in the future to either increase the resource, or install a large equipment mass or operate with large overloads. With further modernization using lighter equipment a mass reserve could be obtained. It can be used to increase payload or fuel reserve.

This approach allows to solve the problems of establishing dependencies between the UAV components masses and the possibility of using them in the corresponding sectors of the airspace. Obtaining such dependencies makes it possible to design safe UAVs, taking into account the usage conditions and combining options for on-board equipment to create multi-functional devices.






References

1. Copejans, H.H., Myburgh, H.C.: A primer on autonomous aerial vehicle design. *Sensors* **15** (12), 30033–30061 (2015)
2. Kritsky, D.N., Druzhinin, E.A., Pogudina, O.K., Kritskaya, O.S.: Decision making by the analysis of project risks based on the FMEA method. In: 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT, Lviv, pp. 187–190. IEEE (2018)
3. Kritsky, D.N., Ovsianik, V.M., Pogudina, O.K., Shevel, V.V., Druzhinin, E.A.: Model for intercepting targets by the unmanned aerial vehicle. In: Palagin, A., Anisimov, A., Morozov, A., Shkarlet, S. (eds.) *Mathematical Modeling and Simulation of Systems. MODS 2019. Advances in Intelligent Systems and Computing*, vol. 1019, pp. 197–206. Springer, Cham (2020)
4. Benić, Z., Piljek, P., Kotarski, D.: Mathematical modelling of unmanned aerial vehicles with four rotors. *Interdisc. Descr. Complex Syst.* **14**(1), 88–100 (2016)
5. Kritskiy, D., Karatanov, A., Koba, S., Druzhinin, E.: Increasing the reliability of drones due to the use of quaternions in motion. In: 9th International Conference on Dependable Systems, Services and Technologies, DESSERT, Kiev, pp. 348–352. IEEE (2018)
6. Lim, H., Park, J., Lee, D., Kim, H.J.: Build your own quadrotor: open-source projects on unmanned aerial vehicles. *Robot. Autom. Mag.* **19**, 33–45 (2012)
7. Mahmood, A., Kim, Y.: Leader-following formation control of quadcopters with heading synchronization. *Aerosp. Sci. Technol.* **47**, 68–74 (2015)
8. Pacheco, N., Resende, D., Magalhães, P.: Stability control of an autonomous quadcopter through PID control law. *Int. J. Eng. Res. Appl.* **5**(5), 7–10 (2015)

9. Phillips, W.F.: *Mechanics of Flight*, 2nd edn. Wiley, New Jersey (2010)
10. Simon, D.: *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley, New Jersey (2006)
11. Chen, H., Wang, X.M., Li, Y.: A survey of autonomous control for UAV. In: *International Conference on Artificial Intelligence and Computational Intelligence*, AICI, Shanghai, pp. 267–271 (2009)



Investigation of the Model of Testing for Weapons and Military Equipment

Ihor Korniienko¹ , Svitlana Korniienko¹ ,
Volodymyr Dmytriiev² , Anatolii Pavlenko² ,
and Dmytro Kamak² 

¹ Chernihiv National University of Technology, Shevchenka 95,
Chernihiv 14035, Ukraine

cornelukr@gmail.com, cornel@ukr.net

² State Scientific Research Institute of Armament and Military Equipment
Testing and Certification, Strilets'ka 1, Chernihiv 14033, Ukraine
{vadmitriiev, agpav}@ukr.net, dkam@meta.ua

Abstract. The effective functioning of the testing institution is related to the systematic, precise and timely tests of weapon specimens and military equipment. The planning of test activities should ensure a continuous process of testing. However, the excessive and uneven arrival rate of specimens for test might adversely affect the planned testing process.

To check the magnitude of the possible impact, the statistical parameters of the input flow of testing requests were calculated. The magnitude of the input flow is determined, the central, second and third moments are calculated, the distribution law of the request rate is determined, the coefficient of variation of the time interval between requests is determined. However, the calculated numerical values of the statistical characteristics of the input flow revealed no significant risks for significantly reducing the efficiency of the testing institution.

Analysis of the statistics revealed some deviation of the actual test time of the separate specimens from the planned value. The study of this deviation revealed the influence of some external and internal factors that accompany the test process. Such deviations can lead to negative effects: delay in the testing procedure of subsequent specimens or the testing institution inoperative in the intervals between tests.

A further solution to the problem of ensuring the effective functioning of the testing institution is to develop a mechanism that allows take into account the factors of influence and determine a priori the test time deviation of a specific specimen from the planned time with some degree of certainty.

Keywords: Testing · Queuing · Weapons and military equipment

1 Introduction

Recently the flow of testing requests for weapons and military equipment has increased significantly. This became a problem because, despite the expansion of the testing organization structure (creation of new specialized units) and the increase in staff, significant time deviations from the basic testing plan started to arise. This turned out to

be critical for the organization because it increased the risk of delays in testing other weapons and military equipment. To prevent disruption of the testing, we had to take extreme measures and modes of operation. Further increase in the size of the testing institution was unacceptable because it increased the cost of maintaining it.

In order to stabilize the planned testing process, it was proposed to study all the major processes that accompany the main stages of the testing process: test planning, testing and processing of test results. The purpose of this study was to search for problems that led to deviations from the basic testing plan and development of optimal algorithms for test planning and management. Another reason for investigating the testing processes for weapons and military equipment was the further development of a conceptual model of an automated test planning and quality management subsystem.

After analyzing a set of methods for the study of dynamical systems, we choice for the analytical apparatus of queuing theory, because in its concept it is well suited to the analysis of the system of mass tests, allowing to obtain decisions at the stage of analytical modeling and, eventually, can become the basis for simulation modeling and analysis of testing processes.

2 Background and Related Work

The issues addressed in this paper are the first steps for research and modeling of the testing institution's activities. Nowadays, the queuing theory [1] is a well-known analytical apparatus for the study of dynamic systems [2] that serve mass demand in many practical areas, such as: social infrastructure, education, production [3, 4], telecommunications, computer technology [5], military and defense technologies [6] and the like.

If the testing institution is represented by a queuing system, then the specimens of weapons and military equipment to be tested are considered to be the input flow. Analytical description of the input flow of requests is possible using the methods of mathematical statistics [7]. To do this, you need to have a data set that has the property of consistency [8].

It should be taken into account that the requests are of different importance and urgency, causing ambiguous behavior in the formed queues, which was for example considered in [9].

In [10] it is stated that recurrent Palma input flows with Poisson-like distributed input requests would be the most convenient way to describe processes in queuing systems. The hypothesis about the nature of the input flow of requests for testing weapons and military equipment was put forward in [11].

Finally, there are many examples of successful application of the theory of queuing systems in military and defense technologies, including modeling of military systems [12, 13], analysis of the effectiveness of combat methods [14], comparative assessment of weapons [15], formation of concepts for building military systems [16], etc.

3 Research Method

3.1 Description of the Object of Study

A special feature of weapon and military equipment testing is the stringent requirements for the reliability of the tests (the accuracy and reliability of the assessments obtained from the tests) and the time limits for testing.

The test process begins with the receipt of a request to test a specific prototype or specimen of weapon or military equipment. The input of requests for testing is massive and continuous in time.

Preparatory measures are taken after the test request is received:

- firstly, the program and test methods are developed, a test plan is drawn up and a test team is formed from specialists;
- then laboratory, measuring equipment and a test site are prepared;
- and eventually training of test staff is conducted.

The preparatory stage has tight time constraints and cannot last longer.

Subsequently, direct tests of a prototype or weapon specimens and military equipment are carried out, which in time is carried out in accordance with a pre-defined test plan.

The test results obtained are processed and a conclusion is drawn on the conformity or non-conformity of the test prototype or weapon specimen and military equipment.

3.2 Modeling

The baseline of the testing process at the institution is represented by the context diagram in Fig. 1 [11].

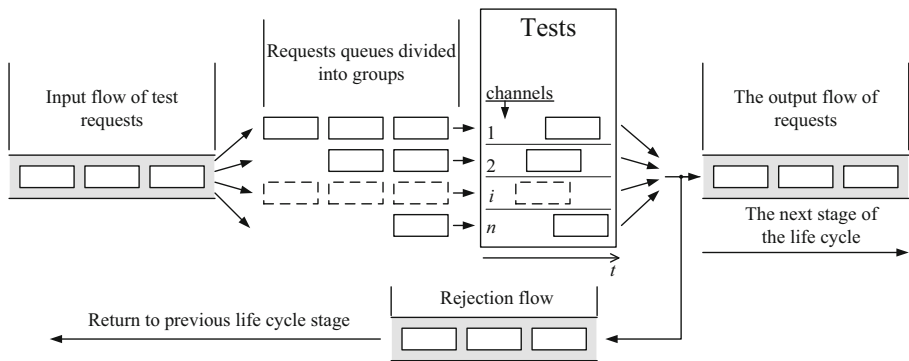


Fig. 1. Context diagram of the testing institution queuing system.

The following points are important for modeling processes in a testing institution and for analyzing problems encountered when planning and performing large amounts of tests:

- description of the parameters of the input flow of test requests;
- queue formation and order of servicing test requests;
- procedure for the distribution of test requests across the system servicing channels;
- description of the service parameters - testing of a prototype or a weapon specimen and military equipment.

The idea of modeling using a queue system is to evaluate the effectiveness of the testing institution through the utilization coefficient [6],

$$\rho = \frac{\lambda}{C}, \quad (1)$$

where λ is the intensity of the input flow of test requests and C is the traffic capacity of the testing institution. When $\rho \rightarrow 1$, we obtain the maximal usable coefficient fully using the workforce and means of the testing institution (test teams, or channels in the context diagram of Fig. 1). The mode with constant input flow and constant traffic capacity will be as close as possible to the optimum.

However, there is always a risk associated with the non-stationarity of the input flow, e.g. when there is an increase in λ , herewith $\lambda \gg C$. In this case, there will be a “congestion” on the servicing of the system, and if the request queueing time exceeds a certain limit, there is a risk of failure of the state task. As practice has shown, the risk was insignificant when the growth in λ was short in time, the channel capacity C has enough reserves to cover the peak values λ .

The worst cases were when the growth in λ was spread out in time and the reserves C were too small, then the risk of failure of the state task was great. Accordingly, the task of determining and analyzing a set of parameters of the input flow of test requests was set.

3.3 Input Flow of Test Requests

Forming a List of Input Flow Parameters. Input flow parameters were determined and analyzed with statistics using the totality of the available reporting data over three years. To estimate the parameters of the input flow, we needed the following statistical characteristics:

- the intensity of testing requests λ [1];
- the mathematical expectation value $M^*[X]$, variance $D^*[X]$ and standard deviation $\sigma^*[X]$ and the number of test requests received over a period of time t [7, 8];
- the coefficient ν of the interval’s variation between moments of occurrence of an event (the deviation of time intervals between test requests) [6],

$$\nu = \frac{\sigma_{\text{int}}}{\tau_{\text{int}}}, \quad (2)$$

where σ_{int} is the standard deviation and τ_{int} is the mathematical expectation value for the length of the interval between requests;

- the distribution law for the number of test requests received per unit of time.

Set the Time Interval for Analysis. To obtain reliable data from the analysis of statistical information, we needed to have a general data set that had the property of consistency. According to the law of large numbers [9–11], the statistical frequency of an event observation approaches the probability of an event occurring $P^*[X] \rightarrow P[X]$, if the total number of events $n \rightarrow 500$ [8]. That is, ideally, the number of time intervals for analysis (when a certain number of events occur) should be at least 500. The standard time units of the year are a month, week, day, and in three years only the number of days exceeds 500. However, when analyzing statistical material, the number of requests received more than one per day was extremely rare: as a rule, either one request per day or no request at all was received. Therefore, choosing a time interval of one day would not be indicative. If the requirement $n_t \geq 500$ is met, we are only satisfied with a two-day interval, and given the five-day workweek (established by national labor law), every fifth interval would be broken by the weekend.

Otherwise, the week unit has the characteristic feature of five working days and two weekends. In addition, the analysis of the statistics revealed a certain correlation of the receipt of applications on certain days of the week (which is related to the procedure of issuance by the top management of the order for conducting the tests). Therefore, for a time interval for analysis, to a certain extent neglecting the rule $n_t \geq 500$, a week was selected, the total number of which in three years was $n_t = 156$.

Analysis of Results. As a result of processing a set of statistics, the following parameters of the input flow of test requests were obtained:

- the intensity of testing requests $\lambda \approx 1,87$;
- mathematical expectation value for the number of test requests per week $M^*[X] = 1,79$, variance $D^*[X] = 1,93$ and standard deviation $\sigma^*[X] = 1,39$;
- the standard deviation of the length of the interval between requests $\sigma_{\text{int}} = 4,12$, mathematical expectation value $\tau_{\text{int}} = 3,79$, according to (2), the coefficient of variation of the intervals between the test requests is $v = 1,08$.

It was important for further modeling to establish the Poisson character of the distribution law for test requests, which would further substantially simplify the simulation process itself. We have constructed a statistical function of the distribution of the requirements for the Poisson law for $\lambda = 1,87$

$$P_x(\Delta t) = \frac{(\lambda \Delta t)^x}{x!} e^{-\lambda \Delta t}; x = 0, 1, \dots \quad (3)$$

The function graph is shown in Fig. 2.

As can be seen from Fig. 2, the built-in statistical function of the distribution of test requests by Poisson's law (3) converges well with the empirical histogram, which is built on a variation series of relative frequencies W_i of the number of test requests per week x_i .

Despite the importance of the input flow numerical data obtained for further modeling of processes in the testing institution, we did not detect significant deviations in the number of test requests per week x_i from their mathematical expectation value $M^*[X]$. There were also no significant prolonged periods of large increase in λ that

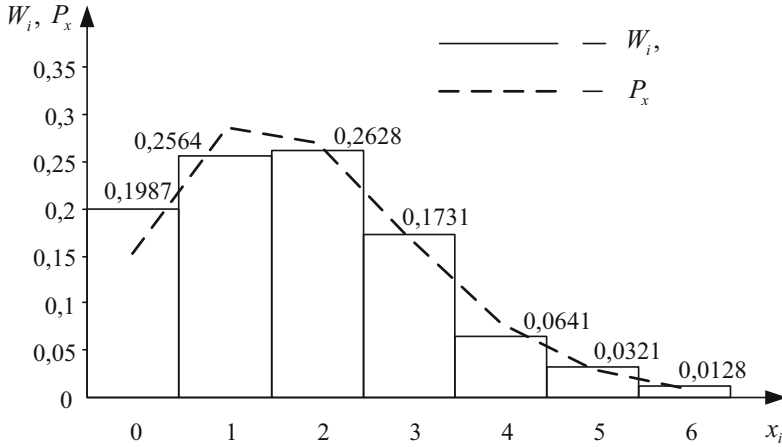


Fig. 2. Input flow parameters: W_i - histogram; P_x - statistical distribution function.

could have a big impact on the deviation from the test organization plan, and the coefficient of interval variation between requests was too small to assert a significant effect of the input flow irregularity.

4 Discussion

The conducted studies of the parameters of the input flow of test requirements did not reveal any significant risk for deviation from the test plan. But, looking at (1), we paid attention to the channel capacity of the testing institution, namely its behavior over time during the tests.

Based on the analysis of the statistical material, we have seen periodic time deviations of tests of certain samples of weapons and military equipment from the planned time indicators, which were expressed in the extended or postponed time of testing. As stated in the reports, the reasons for shifted tests were different, for example: defects of specimens discovered during the test process, weather conditions, unpreparedness of a polygon base or laboratory-measuring equipment, etc. It should be noted that the causes have both an internal character of influence (depending on the institution) and an external character of influence (independent of the institution). In addition, different time requirements for the preparatory stage were observed, which depended on the complexity of the specimen, the development of the methodological base of tests, the training of specialists and so on.

Therefore, taking into account the reasons for the time lags would allow adjustments to be made to the plans and to stabilize the institution's activities both when planning tests of a specific specimen of weapons and military equipment and planned activity of the institution as a whole at the prospective planning stage.

The problem is that, for the most part, the influx of causes, which can be called the factors of the shift in test time, is unpredictable at the test planning stage. Therefore, to

further address the problem of stabilization of the planned test process, it is necessary to investigate and establish a correlation between the reasons that may affect the test process and the magnitude of the time shift of the test process.

5 Conclusions and Future Work

5.1 Conclusions

The clear and stable operation of the testing institution in the context of an intensive flow of weapons and military equipment testing requires careful planning, consideration of all relevant baseline data and analysis of all the processes that accompany the test activity. Applying queuing theory apparatus can facilitate the process of finding optimal algorithms and test planning methods to ensure a stable test process is possible.

Based on the existing statistical material, we have analyzed the parameters of the input flow of test requests and its possible impact on the occurrence of extreme situations that threaten the planned test process. Despite some positive aspects of the study, which are related to the established parameters of the input flow requests and the distribution law of test requests, there are no high risks that would explain the extreme modes of the institution.

However, from the analysis of the statistical material, deviations in the time of testing of certain specimens of weapons and military equipment, which were caused by various internal and external factors, were revealed.

5.2 Future Work

For further research, it is necessary to formulate an approach that will allow to take into account numerically the magnitude of the influence of external and internal factors on the course of testing and to obtain an a priori assessment of the real testing time of specific weapon specimens and military equipment in specific conditions.

References

1. Kleinrock, L.: *Queueing Systems - Volume I: Theory*. Wiley, New York (1975)
2. Kleinrock, L., Gail, R.: *Queueing Systems: Problems and Solutions*. A Wiley-Interscience Publication. Wiley, New York (1996)
3. Matveev, V.F., Ushakov, V.G.: *Queueing Systems: Manual on a special course "Applied mathematics"*, p. 239. Moscow State University (1984). (in Russian)
4. Bocharov, P.P., Pechinkin, A.V.: *Queueing Theory: The textbook*, p. 529. University under P. Lumumba (1995). (in Russian)
5. Allen, A.: *Probability, Statistics and Queueing Theory with Compute Science Applications*, 2nd edn. Academic Press Inc. (1990)
6. Kazachinsky, V.Z., Levitsky, G.E.: *Mathematical methods for solving military special tasks*, p. 292. Kiev Higher Air Force Academy (1980). (in Russian)
7. Cramer, H.: *Mathematical Methods of Statistics*. Princeton University Press, Princeton (1946)

8. Ventsel, E.S.: The Probability Theory, 3rd edn. Nauka, Moscow (1969). (in Russian)
9. Saaty, T.L.: Stochastic network flows: advances in networks of queues. In: Smith, W.L., Wilkinson, W.E. (eds.) Congestion Theory, pp. 86–107. University of North Carolina, Chapel Hill (1964)
10. Gnedenko, B.V., Kovalenko, I.N.: Introduction to Queuing Theory. LKT (2012)
11. Korniienko, I., Korniienko, S., Pokhodenko, O., Kaznachei, S., Rudenko, O.: Graphical representation of the model of the testing organization functioning. Scientific works of State Scientific Research Institute of Armament and Military Equipment Testing and Certification, vol. 2, no. 2, pp. 91–98 (2019). (in Ukrainian)
12. Nikolić, N.: Monte Carlo Modeling of Military Queuing Systems - Challenge of the Initial Transience. Andrejevic Endowment/Strategic Research Institute, Belgrade (2008)
13. Jenkins, P.: Using Markov Decision Process with Heterogeneous Queuing System to Examine Military MEDEWAC Dispatching Policies. MSc thesis in Operations Research. Air Force Institute of Technology - Wright-Patterson Air Force Base, Ohio (2017)
14. Li, L., Liu, F., Long, G., Zhao, H., Mei, Y.: Performance analysis and optimal allocation of layered defense M/M/N queueing systems. Math. Probl. Eng. 1–21 (2016). <https://doi.org/10.1155/2016/5915918>. Accessed 15 Apr 2020
15. Mao, Z., Yu-juan, W., Chao, W., Sheng, H.: Analysis on aircraft sortie generation rate based on multi-class closed queueing network. In: Proceedings 2nd International Conference on Computer Science and Electronic Engineering, ICCSEE, pp. 1877–1880 (2013)
16. Leonhard, R.R., Buchanan, T.H., Hillman, J.L., Nolen, J.M., Galpin, T.J.: Concept for command and control. John Hopkins APL Technical Digest **29**(2), 157–170 (2010). <https://www.jhuapl.edu/Content/techdigest/pdf/V29-N02/20-02-Leonhard.pdf>. Accessed 15 Apr 2020



Evaluation of the Spatial Resolution of Digital Aerospace Image by the Bidirectional Point Spread Function Parameterization

Sergey A. Stankevich^(✉) 

Scientific Centre for Aerospace Research of the Earth, Kiev, Ukraine
st@casre.kiev.ua

Abstract. Actual spatial resolution of acquired digital aerospace images rapid evaluation is a topical issue for any Earth remote sensing application. A quick overview of known methods for the evaluation of spatial resolution of digital images is conducted. A model and method for the spatial resolution evaluating of digital aerospace image using automatically extracted its bidirectional point spread function (PSF) are presented. The bidirectional PSF is optimally approximated by 2D gaussoid with two parameters, which are determined by the bidirectional edge spread function (ESF) of image. The signal distribution along the ESF is restored by results of statistical filtering in a 4×4 sliding window. The final value of spatial resolution is calculated in a frequency domain at the predetermined modulation level. The proposed method was implemented as software code and was tested over a commonly used high-resolution satellite image. The quantitative estimates obtained are in good agreement with modern theoretical background.

Keywords: Digital aerospace image · Spatial resolution · Bidirectional point spread function · Gaussian approximation

1 Introduction

Remote sensing plays an important role in achieving the fundamental goals of sustainable development [1]. In particular, achieving the United Nations sustainable development goals requires high-quality remote sensing data to monitor relevant indicators [2]. Remote sensing is an universal information tool to support decision-makers with prompt and high-quality information about objects and phenomena on the Earth's surface [3]. Images of the Earth's surface contribute the largest and most important part of modern remote sensing data. And the spatial resolution is the most important and universal criterion for the quality of imaging information obtained [4]. Spatial resolution itself determines the information capabilities of remote sensing and must be taken into account during its missions planning [5]. A priori known technical specifications of the imager are not always correct: often manufacturers misrepresent the concept of spatial resolution by the ground sampled distance (GSD), i.e. the size of the photodetector projection onto the Earth's surface. The situation is further worsened by the fact that the actual resolution also depends on the radiometry of digital image

[6]. But the information capabilities of remote sensing applications associated with the objects' detection and identification are determined specifically by the actual resolution of digital image.

Spatial resolution assessing is a significant and long-developing research area. Historically, these methods inherit classic legacy for analog optical and photographic imagery. This paper does not address methods for digital imagery evaluating that are based on the reference image engagement or the human involvement. Only blind evaluation of spatial resolution is considered. Such methods include the spatial resolution evaluating by test charts using classical spatial frequency analysis [7]. The implementation of digital imaging technologies has led to the analysis of discrete evocations of image features such as contrast, sharpness and blur [8]. The need to engage the signal theory statistical methods with covariance, noise and signal-to-noise ratio for image analysis gradually jelled [9]. Finally, it became clear that the best estimates of spatial resolution of digital imagery could be obtained at the intersection of spatial frequency analysis and advanced statistical methods [10, 11].

So, the method for reliable and objective evaluation of actual spatial resolution of a digital image with consideration of one's radiometry is urgently needed. Keeping in mind the huge volumes and rapid flows of aerospace imagery, such method should be fully automatic. In addition, it should provide an opportunity for a simple solution on the suitability or unsuitability of a particular digital image for current remote sensing application.

2 Spatial Resolution Modelling

Several different models of varying complexity can be used for the spatial resolution of aerospace images estimation. Simplified naive models generally do not provide the necessary accuracy of the estimates obtained [12]. The most common of these models is discussed below.

2.1 Geometric Approach

In the simplest case, the spatial resolution r of image is considered as equal to the GSD size d :

$$r = d = a \cdot \mu, \quad (1)$$

where a is the photodetector size, μ is the image scale. For non-oblique aerospace imaging $\mu \cong f/H$ where f is the imager's equivalent focal length, H is the imaging altitude [13].

In some applications, the resolution is determined not by the GSD size, but by the period of the bar grid, which is also resolvable in the image, i.e. the sum of the light and dark bars [14]:

$$r = 2d. \quad (2)$$

2.2 Diffraction-Limited Resolution

The diffraction nature of optical image formation in imaging system limits the angular resolution γ according to the Rayleigh criterion [15]:

$$\gamma = 1.22 \lambda \cdot D, \quad (3)$$

where λ is the optical wavelength, D is the lens aperture. Next, for geometric reasons, it follows

$$r = \gamma \cdot f \cdot \mu = 1.22 \lambda \cdot D \frac{f^2}{H}. \quad (4)$$

The consequence of (4) is an expression for the optimal photodetector size in a diffraction-limited system [16]:

$$a = (0.61 \dots 0.82) \lambda \frac{f}{D}. \quad (5)$$

2.3 Statistical Approach

The concept of spatial resolution applies to digital imaging system remains intact only if the resolution of the discrete sensor array is significantly superior over the resolution of the analog part (e.g., optics) of the system. In the case of consistency of the analog part resolution with the size of the photodetector, it should rather be determining as the detection threshold of a separate discrete element of image (pixel) over the background of others by ones radiometric values [17]. In view the stochastic nature of the digital image formation, the classical statistical detection can be used to evaluate resolution [18]:

$$r = \sqrt{n_{\min}} = \frac{2\sigma_w}{|w_2 - w_1|} \Phi^{-1}(1 - \varepsilon), \quad (6)$$

where n_{\min} is the minimum number of pixels required for detection (statistical sample size), w_1 and w_2 are radiometric means of target and background, σ_w is sample standard deviation, ε is the acceptable error probability, and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$ is the standard normal cumulative distribution function [19].

2.4 Spatial Frequency Analysis

Theoretically, the resolution of imaging system is limited by its transfer function $h(x, y)$, which determines the dependence of the output signal $w(x, y)$ from the input one $w_0(x, y)$ through the convolution [20]:

$$w(x, y) = \iint_{h(x,y) > 0} w_0(x - u, y - v) \cdot h(u, v) \, du \, dv = w_0(x, y) \otimes h(x, y). \quad (7)$$

In optics, the transfer function $h(x, y)$ is known as the point spread function (PSF). To simplify operations of convolution (7) for further analysis, the frequency domain transform is often involved, where the convolution is converted into the common product. The Fourier transform of PSF $H(\xi, \eta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y) \cdot e^{-2\pi i \xi x} \cdot e^{-2\pi i \eta y} \, dx \, dy$ is called optical transfer function (OTF). It's a complex value. Its real absolute value $T(\xi, \eta) = |H(\xi, \eta)|$, called the modulation transfer function (MTF), is needed to evaluate spatial resolution [21]. The spatial resolution corresponds to the spatial frequency ξ^* for the MTF value at a certain threshold level of input modulation K , $K = T(\xi^*)$ as shown in Fig. 1; usually $K = 0.15 \dots 0.30$ [22].

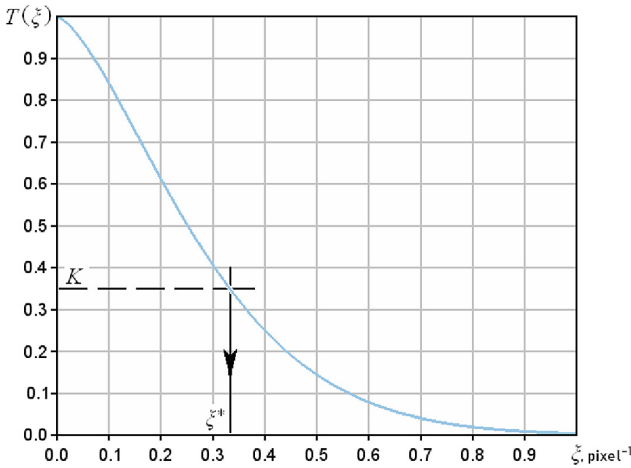


Fig. 1. Spatial resolution evaluation using MTF

The imaging system's MTF can be calculated theoretically according to model specifications of separate links of the optical signal transfer: the atmosphere, the lens, the receiving sensor array, the electronic chain, etc. [23].

2.5 Direct Measurements of Sharpness in the Image

For an accurate and objective spatial resolution evaluation, it would be worth to extract PSF directly from the analyzed image itself. Then the problem of resolution determining will be reduced to an analytical or numerical calculation of the minimum distance at which PSFs can come together without losing their discrimination. Such calculation can be performed based on the physical value of PSF signals in image or on the error probability of close PSFs confusing – see Fig. 2.

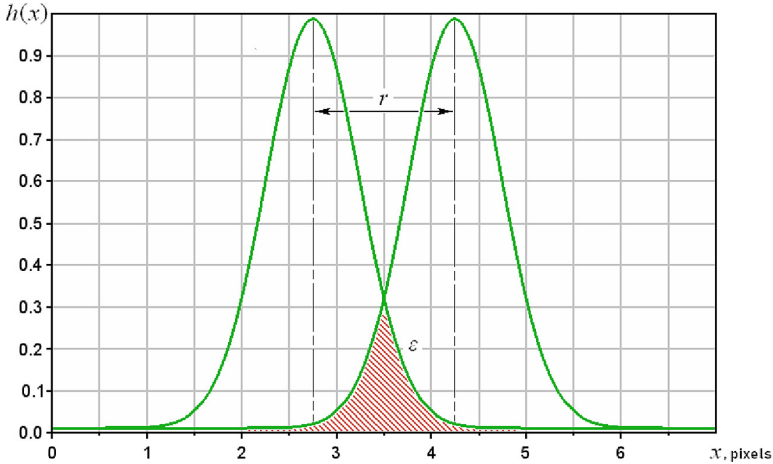


Fig. 2. Spatial resolution evaluation using PSF

The error probability ϵ is proportional to the shaded area in Fig. 2 and it depends on the resolution r chosen value.

Unfortunately, notwithstanding the direct methods for PSF measuring in digital image exist, they are not stable enough [24]. The main drawback of such methods is a strong sensitivity to noise. Much more often, the indirect determination of PSF is applied through the edge spread function (ESF), when evaluating which one the useful signal is accumulated [25]. Figure 3 illustrates the averaged ESF $\bar{w}(x)$ (right) obtaining over a digital image fragment with a brightness crossing boundary (left).

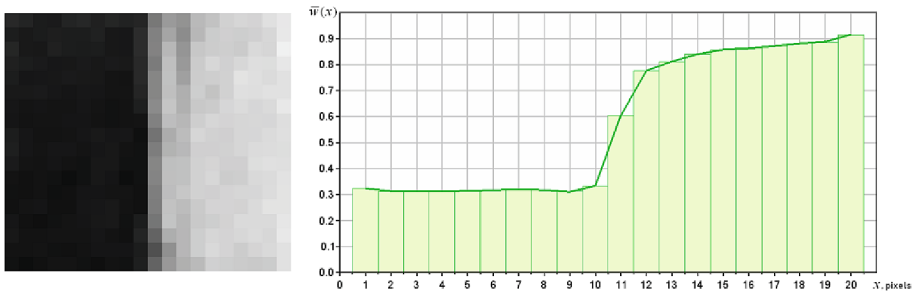


Fig. 3. ESF measuring by image

After obtaining the ESF, its spatial derivative $h(x) = \partial w(x)/\partial x$, called the line spread function (LSF), according to the projection-slice theorem, represents a one-dimensional projection of the two-dimensional PSF $h(x, y)$ [26].

3 Spatial Resolution Evaluation

An analysis of the methods for PSF determining conducted in the previous section shows that the best results provide a direct measurement in the image. Herein, the extraction of a reliable ESF from the image should be considered as the main operation.

3.1 Bidirectional ESF Extraction

The classic widely used method for ESF extraction using some edge detection technique [27] has several problems: susceptibility to noise, a small width of the brightness crossing zone (possibly within one pixel), the arbitrariness of the boundary direction relative to the raster grid.

The first problem is solved by filtering only the most suitable ESF selection. To do this, a certain measure of the ESF quality is introduced, for example, the sharpness of the brightness jump. To solve the second problem, a parametric approximation of the ESF in a small spatial vicinity of the analyzed brightness crossing is required. Quite a lot of ESF models are known [28], but the Gaussian approximation is most often used. This approximation provides a fairly high accuracy, it is easily differentiated and analytically transformed into MTF. The third problem can also be solved by simply selecting the raster-native ESF only, for which it suffices to perform scanning in window exclusively in two directions. At the significant size of the aerospace image scene, the number of required ESF will be sufficient.

3.2 Bidirectional PSF Restoration

The Gaussian approximation of bidirectional PSF is completely determined by two parameters in all, namely the Gaussian constants along the axes σ and ζ :

$$h(x, y) = \frac{1}{2\pi \cdot \sigma \cdot \zeta} \cdot e^{-\frac{x^2}{2\sigma^2}} \cdot e^{-\frac{y^2}{2\zeta^2}}. \quad (8)$$

Unknown parameters σ and ζ are calculated using one-dimensional LSF, each independently for its own:

$$h(x) = \frac{1}{\sqrt{2\pi} \sigma} \cdot e^{-\frac{x^2}{2\sigma^2}}, h(y) = \frac{1}{\sqrt{2\pi} \zeta} \cdot e^{-\frac{y^2}{2\zeta^2}}. \quad (9)$$

Chosen approximations (9) leads to the following equations for the ESFs:

$$\bar{w}(x) = \int_{-\infty}^x h(u) du = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^x e^{-\frac{u^2}{2\sigma^2}} du = \Phi(x/\sigma), \bar{w}(y) = \Phi(y/\zeta). \quad (10)$$

The EFS parameters depend on the values of normalized signals in high and low brightness segments, as well as on the degree of inclination of their distributions from the horizontal. Their optimal values can be obtained using the least squares method, the gradient method, or any other more accurate optimization method [29].

3.3 Bidirectional MTF Modelling

If the parameters of Gaussian approximation are known, then it is possible to write out the appropriate equation for bidirectional MTF:

$$T(\xi, \eta) = |F\{h(x, y)\}| = e^{-2\pi^2 \cdot \sigma^2 \cdot \xi^2} \cdot e^{-2\pi^2 \cdot \zeta^2 \cdot \eta^2}, \quad (11)$$

where $F\{\cdot\}$ is the Fourier transform operator.

Since the considered system is linear, and the PSF is separable in a good approximation [30], the spatial resolution components r_x and r_y can be estimated by MTF (11) in isolation from one another:

$$r_x = \frac{1}{\xi^*} = \sqrt{-\frac{2\pi^2 \cdot \sigma^2}{\ln K}}, \quad r_y = \frac{1}{\eta^*} = \sqrt{-\frac{2\pi^2 \cdot \zeta^2}{\ln K}}, \quad (12)$$

where $K = T^*(\xi, \eta)$ is the modulation threshold.

4 Test Image Approbation

A Fig. 4 fragment of the WorldView-3 (30 cm GSD) satellite digital panchromatic image from Maxtar (<http://www.digitalglobe.com/products/satellite-imagery>) was used for the method testing. A large number of bright crossings are present in the test image.

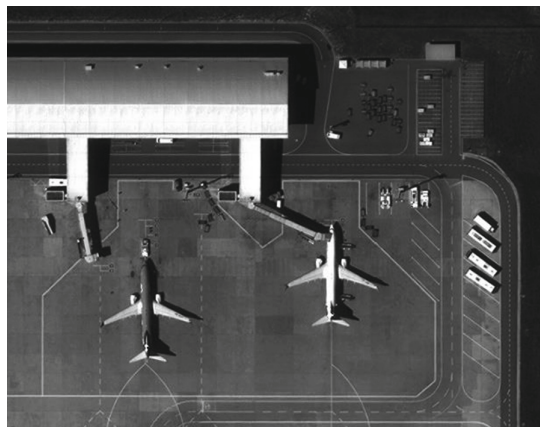


Fig. 4. WorldView-3 satellite image test fragment

The technique outlined in Sect. 3 is implemented as an image processing framework in the SciLab open source computation environment (<http://www.scilab.org>) with a graphical interface Fig. 5. The SciLab code is structured into the reusable procedure for bidirectional PSF evaluation as well as into the explicit-control codesegment for image scanning and logical operations executing.

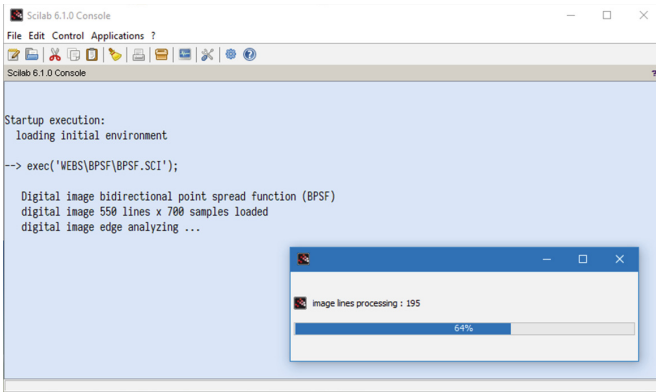


Fig. 5. SciLab application for image analysis

The input image brightness is linearly normalized; processing is carried out inside a 4×4 sliding window. The best ESF represents are selected in parallel both horizontally and vertically over the entire image. After selection is completed, the optimal approximation parameters σ and ζ are calculated. The resulting approximations of ESFs are plotted in Fig. 6.

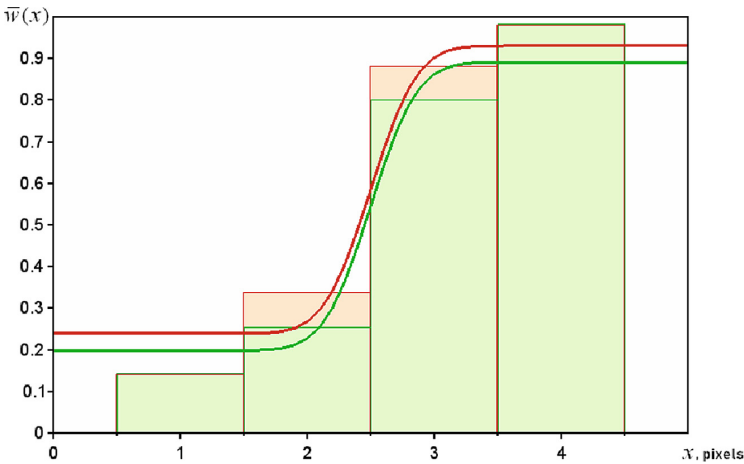


Fig. 6. Approximated ESFs of satellite image Fig. 4: — horizontal, — vertical

The obtained approximation parameters values make it possible to restore the bidirectional MTF (11) of the test image and evaluate one's spatial resolution using Eqs. (12). The calculated spatial resolution values and auxiliary intermediate entities of the test image are summarized in Table 1.

Table 1. The obtained estimates of the test image characteristics

Direction	Normalized radiometric values		Approximation parameters, pixels	Spatial resolution, pixels
	Low	High		
Horizontal	0.337	0.881	0.312	1.175
Vertical	0.254	0.800	0.322	1.214

The estimates obtained are slightly worse than the official specifications of the WorldView-3 satellite data product that is in principle confirmed by some independent tests [31].

5 Conclusions

Thus, a computer technique for the spatial resolution of digital aerospace images automatic determination has been developed and tested using statistical and spatial frequency methods combination. This technique will be quite useful for developers of aerospace imaging systems, for express evaluation of digital images acquired, as well as under the experimental evaluation of electro-optic imagers.

Future research should be expanded in two ways. Firstly, the proposed technique should become technologically mature and take the form of a complete well-proven computer framework suitable for large-scale implementation and practical service. Secondly, the problem of multidimensional imagery remains unresolved. Almost all modern aerospace imaging systems generate multi- and hyperspectral images or at least color ones. Even neglecting the issues of bands co-registration, the now available philosophy for their evaluation is not suitable. This is a significant scientific challenge, and it needs new research.

References

- Schneider, F., Kläy, A., Zimmermann, A.B., Buser, T., Ingalls, M., Messerli, P.: How can science support the 2030 Agenda for Sustainable Development? Four tasks to tackle the normative dimension of sustainability. *Sustain. Sci.* **14**(6), 1593–1604 (2019). <https://doi.org/10.1007/s11625-019-00675-y>
- Kussul, N., Lavreniuk, M., Kolotii, A., Skakun, S., Rakoid, O., Shumilo, L.: A workflow for Sustainable Development Goals indicators assessment based on high-resolution satellite data. *Int. J. Digit. Earth* **13**(2), 309–321 (2020). <https://doi.org/10.1080/17538947.2019.1610807>



3. Chang, N.-B.: *Environmental Remote Sensing and Systems Analysis*. CRC Press, Boca Raton (2012). ISBN 9781439877432
4. Wu, H., Li, Z.-L.: Scale issues in remote sensing: a review on analysis, processing and modeling. *Sensors* **9**(3), 1768–1793 (2009). <https://doi.org/10.3390/s90301768>
5. Degbelo, A., Kuhn, W.: Spatial and temporal resolution of geographic information: an observation-based theory. *Open Geospatial Data Softw. Stand.* **3**(1), 1–22 (2018). <https://doi.org/10.1186/s40965-018-0053-8>
6. Thomson, G.H.: A note on spatial resolution measurement and its implications for image radiometry. *Int. J. Remote Sens.* **30**(1), 1–8 (2009). <https://doi.org/10.1080/01431160802339480>
7. Becker, S., Haala, N.: Determination and improvement of spatial resolution for digital aerial images. *ISPRS Arch.* **XXXVI**(1/W3), 51–56 (2005)
8. Chandler, D.M.: Seven challenges in image quality assessment: past, present, and future research. *ISRN Signal Process.* **2013**, 905685 (2013). <https://doi.org/10.1155/2013/905685>
9. Dutta, J., Ahn, S., Li, Q.: Quantitative statistical methods for image quality assessment. *Theranostics* **3**(10), 741–756 (2013). <https://doi.org/10.7150/thno.6815>
10. Wang, M., Zhou, S., Yan, W.: Blurred image restoration using knife-edge function and optimal window Wiener filtering. *PLoS ONE* **13**(1), e0191833 (2018). <https://doi.org/10.1371/journal.pone.0191833>
11. He, H., Xie, X., Liu, Y., Liang, H., Zhou, J.: Exploiting the point spread function for optical imaging through a scattering medium based on deconvolution method. *J. Innov. Opt. Health Sci.* **12**(04), 1930005 (2019). <https://doi.org/10.1142/S1793545819300052>
12. Zhan, Q., Molenaar, M., Tempfli, K., Shi, W.: Quality assessment for geo-spatial objects derived from remotely sensed data. *Int. J. Remote Sens.* **26**(14), 2953–2974 (2005). <https://doi.org/10.1080/01431160500057764>
13. Jähne, B.: *Digital Image Processing*. Springer, Berlin (2005). <https://doi.org/10.1007/3-540-27563-0>
14. Sviridov, K.N., Tyulin, A.E., Pulinet, S.A.: New assessment of linear instrumental ground resolution of Earth remote sensing spacecraft for perfect design of its optoelectronic equipment. *Curr. Prob. Remote Sens. Earth Space* **17**(1), 59–67 (2020). <https://doi.org/10.21046/2070-7401-2020-17-1-59-67>. (in Russian)
15. Rees, W.G.: *Physical Principles of Remote Sensing*. Cambridge University Press, New York (2013). ISBN 9781107004733
16. Light, D.: A basis for estimating digital camera parameters. *Photogram. Eng. Remote Sens.* **70**(3), 297–300 (2004)
17. Landgrebe, D.A.: *Signal Theory Methods in Multispectral Remote Sensing*. Wiley-Interscience, Hoboken (2003). ISBN 9780471420286
18. Stankevich, S.A.: Estimating the linear resolution of digital aerospace imagery. *Space Sci. Technol.* **8**(2–3), 103–105 (2002). <https://doi.org/10.15407/knit2002.02.103>. (in Russian)
19. Pollard, J.: *A Handbook of Numerical and Statistical Techniques*. Cambridge University Press, Cambridge (1977). <https://doi.org/10.1017/CBO9780511569692>
20. Schowengerdt, R.A.: *Remote Sensing: Models and Methods for Image Processing*. Academic Press, San Diego (2007). ISBN 9780123694072
21. Vollmerhausen, R.H., Reago Jr., D.A., Driggers, R.G.: *Analysis and Evaluation of Sampled Imaging Systems*. SPIE Press, Bellingham (2010). ISBN 9780819480774
22. Smith, W.J.: *Modern Optical Engineering*. McGraw-Hill, New York (2008). ISBN 9780819470966
23. Boreman, G.D.: *Modulation Transfer Function in Optical and Electro-Optical Systems*. SPIE Press, Bellingham (2001). <https://doi.org/10.1117/3.419857>

24. Brauers, J., Seiler, C., Aach, T.: Direct PSF estimation using a random noise target. In: Proceedings of SPIE, vol. 7537, p. 75370B (2010). <https://doi.org/10.1117/12.837591>
25. Claxton, C.D., Staunton, R.C.: Measurement of the point-spread function of a noisy imaging system. *J. Opt. Soc. Am.* **25**(1), 159–170 (2008). <https://doi.org/10.1364/JOSAA.25.000159>
26. Haefner, D.P.: Best practices for imaging system MTF measurement. In: Proceedings of XVI IS&T International Symposium on Electronic Imaging, Image Quality and System Performance (EI 2019), 319. Society for Imaging Science and Technology, Burlingame (2019). <https://doi.org/10.2352/ISSN.2470-1173.2019.10.IQSP-319>
27. Zhong, J.: A knife-edge input point spread function estimation method for document images. In: Proceedings of the 2nd International Conference on Social Science and Technology Education (ICSSTE 2016), pp. 618–622. Atlantis Press, Guangzhou (2016). <https://doi.org/10.2991/icssste-16.2016.114>
28. Li, T., Feng, H., Xu, Z., Li, X., Cen, Z., Li, Q.: Comparison of different analytical edge spread function models for MTF calculation using curve-fitting. In: Proceedings of SPIE, vol. 7498, p. 74981H (2009). <https://doi.org/10.1117/12.832793>
29. Masaoka, K.: Accuracy and precision of edge-based modulation transfer function measurement for sampled imaging systems. *IEEE Access* **6**, 41079–41086 (2018). <https://doi.org/10.1109/ACCESS.2018.2856742>
30. Easton Jr., R.L.: *Fourier Methods in Imaging*. Wiley, Chichester (2010). ISBN 9781119991861
31. Longbotham, N., Pacifici, F., Malitz, S., Baugh, W., Camps-Valls, G.: Measuring the spatial and spectral performance of WorldView-3, HW3B.2. In: Proceedings of the Conference on Hyperspectral Imaging and Sounding of the Environment (HISE 2015), OSA, Lake Arrowhead (2015). <https://doi.org/10.1364/HISE.2015.HW3B.2>

Mathematical Modeling and Simulation of Systems in Project Management



Model for Human Capital Increasing via Motivation for General Training in Company Under Economic Crisis Overcoming

Liliia Nikiforova  and Anatolii Shyian  ^(✉)

Vinnitsia National Technical University, Khmelnytske schose 95,
Vinnitsia 21021, Ukraine
anatoliy.a.shyian@gmail.com

Abstract. We describe baseline model for both employees and companies to increase human capital through general training. This model is primarily aimed for companies under overcoming the economic crisis. The model describes the situation when there is a significant level of migration for highly skilled employees, too. The model uses the parameters for the company (part of the tuition fees the company pays) and for the employee (part of the maximum wage after training, the number of years from which the employee returns the loan for the company). It is shown that the motivation space always exists for the employee and the company. Therefore, the channel for motivating an employee for general training always exists for the company. Each point in the motivation space for the employee is profitable both for the company and for the employee at the same time. The company and employee can determine their own profit as a result of bargaining. Our results allow creating effective incentives for employees, which allows the company to create conditions both for increasing the productivity of employees and for continuing work in the company.

Keywords: Human capital · General training · Motivation · Wage inequality

1 Introduction

Today, the economy of most countries of the world is rapidly entering the economic crisis. It is caused by ignoring the social and economic challenges that the developed countries of the world faced. The pandemic of the coronavirus disease COVID-19 and the ever-increasing number and role of migrants proved to be a very powerful driver of the crisis too [1].

The way out of such a powerful crisis is always based on a change in the structure of the economy of each country, caused by the emergence and rapid development of innovative companies. All this requires powerful changes in the professional structure of a huge number of employees.

Therefore, in the conditions of overcoming from economic crisis, the company's requirements for the presence of human capital that have the necessary qualifications

increase. The presence of human capital in such company will determine the success of economic development in the aftermath of the crisis period.

The formation of the necessary professional skills among company employees is carried out through education and training. Becker's classic work [2] (Becker 1962) draws attention to the fact that training-on-the-job can usually be divided into general and special trainings.

General training develops the skills of an employee, which can be applied in different workplaces and in different companies. For example, this type of training includes trainings on new skills in the field of new technologies.

Specialized training develops only those skills that can be applied only to a particular company (for example, only within one particular activity). For example, this type of training refers to the skills of using specific objects of new technology.

Typically, starting with the classic Becker [2] models and Ben-Porath [3], general workplace training models use the assumption that total training is paid only by the employee (see in detail [4] and [5]). Recent models also included this assumption; see, for example, [6].

Human capital today is one of the powerful tools of economic development. The constant increase in the level of human capital is necessary for growth, because today there are rapid changes in technology, the connection between people and the migration of workers to societies with a different culture.

As a result, companies had to take part in paying for general education together with employees. This led to the emergence of a number of new models.

In [7] builds a model for firms that offer skills training. In the document, the author concentrates on the problem of identifying the reasons why companies pay for training in general skills. The answer lies in the fact that the company thus receives personal information about labor productivity.

In [6] proposed a model in which they assume that all investment in training should be funded by the firm.

In Croce's model [8] a number of proposals for the training process are used. For example, the employee can leave the company after training and the company participates in financing the training of employees only due to a payroll tax or profit tax.

In [10] and [11] found the evidence that the need for joint financing of general training on the part of the company and the employee.

The modern market is characterized by a high level of migration of people [1]. However, migrants in a new country find themselves in conditions that differ from the conditions in the domestic country. In particular, this applies to migrants from countries in which socio-economic institutions and accepted norms of behavior differ significantly from those that exist in the new country. As a result, migrants should receive general training, the task of which is to obtain new skills for him/her, which he can apply to various companies.

Staff training can play the role of a powerful motivational factor for the employee. First, staff training leads to an increase in wages. Secondly, the development of skills provides additional guarantees to the employee from dismissal. Thirdly, in the case of dismissal, the employee receives additional employment opportunities after high qualification.

The problem of joint participation of the company and its employees in raising the level of human capital has an internal contradiction. On the one hand, the company needs highly skilled workers. On the other hand, additional funds are needed to improve the efficiency of the employee. In fact, we are getting some sort of trap: the company, having no means to improve the skills of its employees, produces simpler goods. And at the same time, workers who have less wages do not have the resources to improve their skills.

Finally, the employee after general training has the opportunity to switch to another job in the company, where his wage will be higher. The situation is aggravated for migrants, especially for highly skilled employees.

Thus, today there are new areas of general education that are necessary for the functioning of a particular company. In this case, unlike the classical approach to general education, the company turns out to be an interested agent for receiving such training by the employee.

The purpose of the article is to develop a baseline model for both employees and companies to increase human capital through general training through the joint payment of general training in the workplace.

2 The Model

Employees must be continually retrained, as innovations and technologies are spreading. Previously, the company always recruited a new employee who knows new technologies or innovations. But today employees who may know new technologies or innovations are absent (as higher education in these countries has greatly decreased). The company needs a situation where the employee continues training in the company. Therefore, the company requires great motivation for employees to save their workplace after training.

An effective scenario for training employees is the case when the cost of training will be simultaneously paid by the employee and the company. This scenario simultaneously solves the problems of increasing the productivity, motivation and retention of employees in the company. Moreover, we will describe below that the profit for the company will also be received.

The model lasts two periods. In the first period the employee pays part of the payment and receives training. The employee has training and at the same time continues to carry out his work in the company. The company pays part of the payment (like lending to an employee). In the second period after training, the employee receives a wage increase and pays for loans. The company increases the wage for the employee (which increases productivity after training) and receives repayment of the loan from the employee. Of course, the company does raise wages only in part.

The described model for communication between the employee and the company is beneficial for both parties.

The employee has access to training to improve their competence. In addition, the employee receives a guarantee that he will not be dismissed on the initiative of the company during the loan repayment period. Finally, the wage will be increased

(although not in the same degree, of course, as it would be if it had the appropriate level performance).

The company also receives its “bonus”. First, he “connects” the employee with the enterprise in a certain period. Secondly, it receives higher productivity (through training) of employees at lower wages (it can provide greater savings by discounting). Thirdly, the company can support only a part of employees: the overall productivity of workers will be increased due to competition between employees. Finally, the company can switch to the production of more competitive products due to the fact that its employees will receive certificates from the relevant firms (for the case when training is special).

We introduce the following notation. c is training fee; $0 < p < 1$ is part of training fee, which the company pays; n is the number of years, from which an employee returns of a lending to company; w_L is (annual) wage of an employee before training, $\Delta w > 0$ is (annual) wage increase for employee after increasing his/her skills (after from training), $\Delta P > 0$ is an annual increase of profit of a company from employee after training; $f > 0$ is a total cost for finding an employee with high productivity and loss for adaption of an employee in a company, $w_H > w_L + \Delta w$ is (annual) wage for employee with high productivity (such after training).

An employee solves the problem of maximizing his/her utility function U_e .

$$\max \left\{ U_e = pc \frac{1 - \beta^{n+1}}{1 - \beta} - (1 - p)c - (w_H - w_L - \Delta w)n \right\}. \quad (1)$$

Here β is a discount factor $\beta = (1 + r)^{-1}$, where r is a market discount rate (in absolute terms).

The first term in (1) corresponds to the bonus of the employee, which he receives as a result of the fact that the company pays a part of the tuition fee. It is calculated as a loan that an employee must repay a loan for n years. The second term is the cost that the employee pays for training. The third term is the loss of an employee due to the fact that he is already a skilled worker and receives less payment for n years.

A company maximizes a utility function U_f .

$$\max \{ U_f = \Delta P \cdot n - pc + (w_H - w_L - \Delta w)n \}. \quad (2)$$

In (2) the first term corresponds to the additional profit for the company, which is due to the increase in employee productivity after training. The second term is the company's payment for training an employee. The third term is the company's profit due to the fact that the employee receives wages lower than his productivity.

It should be noted that the company could choose an alternative way. It can to fire this employee and to hire an employee with high productivity. A company gets a profit U_f^N by this way.

$$U_f^N = \Delta P \cdot n - f - w_H n. \quad (3)$$

The first term in (3) is the company’s profit from using an employee with high productivity. The second term is the cost of finding an employee with high productivity and loss for the adaptation of the employee in the company. The third term is the wage of an employee with high productivity. Formula (3) is written on the assumption that the term for finding a high-productivity employee is small (and therefore it is omitted).

A company chooses training for employee, when a requirement will be satisfied.

$$U_f > U_f^N. \tag{4}$$

There are the variables, which an employee and a company may use for maximization of a utility function. A part of training fee, which the company pays p and annual wage increase for employee after increasing his/her skills Δw can play a role of such variables. A company controls a variable p , and an employee controls a variable Δw .

A condition (4) we can rewrite as first condition for the control variables.

$$pc + \Delta w \cdot n < (2w_H - w_L)n + f. \tag{5}$$

Please note that the obsolescence of the technology lasts for several years. Accordingly, typical values will be equal to $n = 1, 2$.

An employee must get a profit from training, too. Therefore a second condition for the control variables obtains from the condition $U_e > 0$.

$$pc \left(1 + \frac{1 - \beta^{n+1}}{(1 - \beta)} \right) + \Delta w \cdot n > c + (w_H - w_L)n. \tag{6}$$

For convenience of analysis we introduce the relative values of the control variables of the problem. We introduce the quantity $W = (w_H - w_L)$, which is the maximum significance for the increasing employee’s wage. Thus the following dimensionless variables we introduce: $\alpha = \Delta w/W$, where $0 \leq \alpha \leq 1$, $\gamma = c/W$, $F = f/W$, $H = w_H/W$.

With using these variables inequalities (5) and (6) can be rewritten in a dimensionless form.

$$\frac{\gamma p}{n} + \alpha < 1 + \frac{H + F}{n}, \tag{7}$$

$$\frac{\gamma p}{n} \left(1 + \frac{1 - \beta^{n+1}}{1 - \beta} \right) + \alpha > 1 + \frac{\gamma}{n}. \tag{8}$$

The inequalities (7) and (8) on the plane of the control parameters (p, α) and with the restrictions for them ($0 \leq p \leq 1$ and $0 \leq \alpha \leq 1$) will define the space, in which the each point (i.e., each value of p and α) will get the profits for both an employee and a company. This space we call as a motivation’s space for employee. Every point from this space is corresponded to situation, which can be described as a contract between an employee and a company (and this contract will be profitable for both agents).

Unlike the models Acemoglu [5], Bassanini et al. [9] and Brunello [10] we took into account the different characteristics that are characteristic of countries with transition economies. We have implemented a discount rate, wage differentials (between employees with high and low productivity) and the total cost of searching for an employee with high productivity and loss for the adaptation of the employee in the company. Unlike the models discussed above, our goal is to find parameters about which the employee and the company can agree and before training. These parameters form a motivational space for the employee.

3 Model Analysis

Two theorems are valid for a motivation space for employee in general case.

Theorem 1. For any $n \geq 1$ the inequality (7) do not restricts the motivation's space for employee, when the condition $c < w_H + f$ is valid (and this inequality must be omitted).

Proof. Restrictions $0 \leq \alpha \leq 1$ and $0 \leq p \leq 1$ set a square on the plane (p, α) . The intersection of the line from (7), which sets by changing the sign of inequality to sign of equality, with the α -axis is defined as $\alpha_{p=0} = 1 + (H + F)/n > 1$. This point always lies above the motivation's space for employee.

Let's we consider the conditions, under which a line from (7) $p\gamma + n\alpha - n - H - F = 0$ intersects with a line $\alpha = 1$. This intersection occurs at a point $p_0 = (H + F)/\gamma$. In a situation, when $p_0 > 1$, can not be intersection with the square of possible values for p and α . In other world, only when $\gamma < H + F$ (or, in absolute units, $c < w_H + f$) the restriction (7) is irrelevant for our purposes. *End of theorem.*

The inequality $\gamma < H + F$ (or, in absolute terms, $c < w_H + f$) is performed only if training's payment is less than a sum of an annual wage of the employee with high qualification (i.e. after training) and the cost, which spends on finding of skilled worker and on preparing his/her for work in a company. This inequality always performs, because we consider only short-term trainings (a payment on them does not to exceed an annual wage for employee).

Theorem 2. The motivation space for the employee always exists for any conditions of the problem under consideration (which are given by inequalities (7) and (8)). A straight line (9) defines one of the boundaries in the motivation space for the employee.

$$\frac{\gamma p}{n} \left(1 + \frac{1 - \beta^{n+1}}{1 - \beta} \right) + \alpha - 1 - \frac{\gamma}{n} = 0. \quad (9)$$

Proof. We use the theorem of analytic geometry which can be stated as follows: two points $M_1(x_1, y_1)$ and $M_2(x_2, y_2)$ are always located on the same side of the line $Ax + By + C = 0$ if and only if the inequality $(Ax_1 + By_1 + C)(Ax_2 + By_2 + C) > 0$ exists. Let's consider two points with $p = 0, \alpha = 1$ and $p = \alpha = 1$ (which define the top

horizontal border for the motivation’s space for employee). Substituting these values in (9), we obtain the following inequality:

$$\begin{aligned} & \left(1 - 1 - \frac{\gamma}{n}\right) \left[\frac{\gamma}{n} \left(\frac{1 - \beta^{m+1}}{1 - \beta}\right) + 1 - 1 - \frac{\gamma}{n}\right] \\ &= \left(-\frac{\gamma^2}{n^2}\right) \left(1 + \frac{1 - \beta^{m+1}}{1 - \beta} - 1\right) = \left(-\frac{\gamma^2}{n^2}\right) \left(\frac{1 - \beta^{m+1}}{1 - \beta}\right) < 0. \end{aligned}$$

Thus, the points under consideration lie on opposite sides of the line (9). This means that the line (9) crosses the upper boundary of the motivation space for the employee. As a result, the space of motivation for the employee in the frame of the model always exists, and Eq. (9) defines one of its boundaries. *End of theorem.*

Numerical calculations of the influence of variations in model parameters on the characteristics of the motivational space that is used in negotiations are presented on Figs. 1 and 2. Figure 1 shows the effect of variation γ on the intervals of variation of α and p , which are used in negotiations between the company and employees. Figure 2

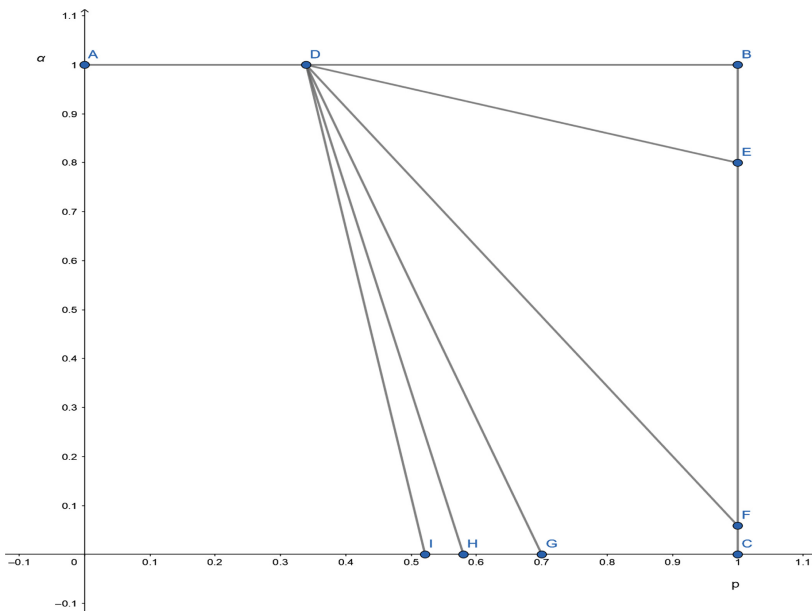


Fig. 1. The motivation space (MS) for bargaining between company and employee. Variable values are $n = 1, \beta = 0,91$ ($r = 0,1$). $DE - \gamma = 0,1$ (MS = DBE), $DF - \gamma = 0,5$ (VS = DBF), $DG - \gamma = 1$ (MS = DBCG), $DH - \gamma = 1,5$ (MS = DBCH), $DI - \gamma = 2$ (MS = DBCI).

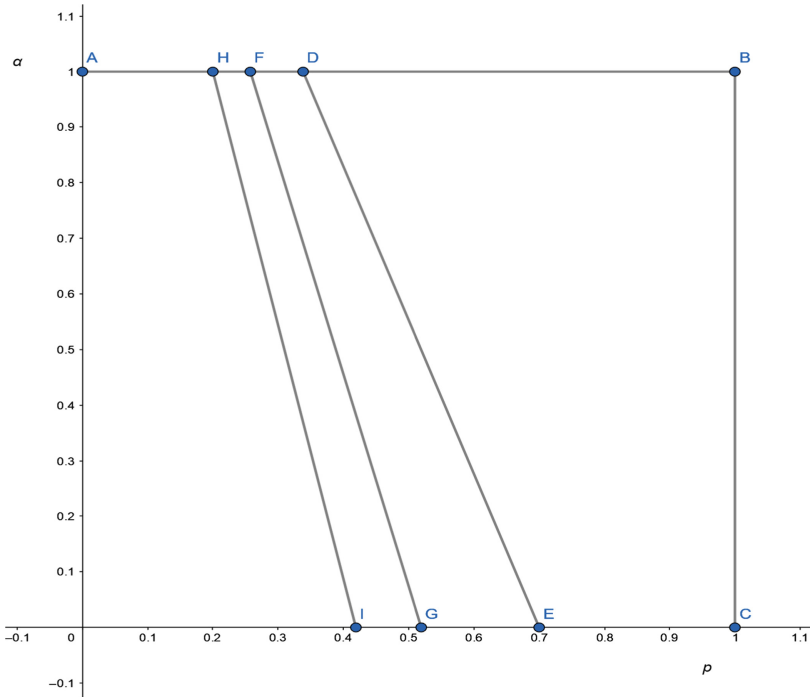


Fig. 2. The motivation space (MS) for bargaining between company and employee. Variable values are $\gamma = 1$, $\beta = 0,91$ ($r = 0,1$). $DE - n = 1$ (MS = DBCE), $FG - n = 2$ (MS = FBCG), $HI - n = 3$ (MS = HBCI).

shows the effect of the number of periods of reduction in workers’ wages on the intervals of variation of α and p .

Finally, we note that if the inequality $c < wH + f$ is not satisfied, then inequality (7) becomes relevant. In this case, the bargaining set will be limited above by a line segment $\gamma p + n\alpha - 1 - HF = 0$. Appropriate formulas are easy to obtain. However, this case corresponds to a rather limited range of situations (for example, for higher education of employee, etc.), and therefore this case will not be discussed in detail.

4 Discussion and Outlook

We see that the motivation space for both the employee and the company always exists within the framework of Theorem 1. In other words, each point from this space is beneficial both for the employee and for the company.

Of course, the Nash equilibrium is points with $\alpha = 1$ and point D (Fig. 1) or points D, F, H (Fig. 2). The Pareto optimum is point with $p = \alpha = 1$.

However, for real situations, Nash equilibrium and Pareto optimum are often not applicable. Because we do not take into account all the features of the real situation in

which the employee and the company worked. Moreover, we often have to take into account that the company (or employee) can have power.

If a company has a power, that a company can to choose the point D for case at Fig. 1, and D, F or H (when $\alpha = 0$) for case at Fig. 2. If an employee has a power, that an employee will always choose $p = \alpha = 1$. Different cases can be, too.

The company can “exchange” part of the loan for the cost of further savings in wages. In other words, the company has enough opportunities to decide what is more important for it: to spend a lot of money once or to save on the employee’s wage in the long term.

This situation is typical for the employee. He/she can also “exchange” his expenses: large expenses today or a larger increase in wages after.

The important facts are the following.

- 1) A minimum value for p always exists for the payment from a company for employee’s general training. It is calculated by the formula (9) with $\alpha = 1$ and depends only on the discount factor and on the loan term. In other words, the company makes a profit when the loan amount for the employee will exceed a certain minimum cost.
- 2) The minimum wage for an employee exists only in the case, when the inequality (10) is satisfied.

$$\gamma < \frac{n(1 - \beta)}{1 - \beta^{n+1}}. \tag{10}$$

In absolute dimensions the inequality (11) looks as following.

$$c < n(w_H - w_L) \frac{1 - \beta}{1 - \beta^{n+1}}. \tag{11}$$

An inequality (11) exists in two cases. First, if training is inexpensive. Second, if a term, during which an employee returns of a lending to company, is sufficiently long.

The obtained results make it possible to clearly take into account the difference between the training programs that provide the employee with different skills. Within the framework of the model, such differences will be reflected in different prices of general training and on different wages of an employee after training. Thus, the developed model can be used by the company to select the optimal training for employees.

The modern market is characterized by a high level of migration of people [1]. However, migrants in a new country find themselves in conditions that differ from the conditions in the domestic country. In particular, this applies to migrants from countries in which socio-economic institutions and accepted norms of behavior differ significantly from those that exist in the new country. As a result, migrants should receive general training, the task of which is to obtain new skills for him/her, which he can apply to various companies.

For example, in [11] the relationship between the success of Polish service firms and the availability of appropriate manpower was investigated. It has been established that the motivational effect of relations between people and organization is less than on the relation “man-work”. Thus, Polish service firms may have a chance to increase their

success when the company's attractiveness is enhanced by the possibility of increasing human capital through training.

5 Conclusions

Unlike existing models, the paper considers partial financing of training by the employee, as well as repayment of the company's expenses for general education because of an incomplete increase in wages to the employee after training for a certain period of time. Thus, there is an opportunity for bargaining between the employee and the company, which allows taking into account both the current economic state of the company and the individual characteristics of the employee. In particular, the management characteristics for bargaining are: part of the payment for general training by the company, part of the total increase in the employee's wage after general training, the time of the employee's return of the loan to the company. The results obtained can be used for different training costs, different levels of differences between the employee's wage before and after general education.

Acknowledgements. We thank Anatoliy Fisenko for both helpful discussions and improving of our English.

References

1. Ramirez, M.D., Liebig, Th., Thoreau, C., Veneri, P.: The Integration of Migrants in OECD Regions: A First Assessment. OECD Regional Development Working Papers 2018/01 (2018)
2. Becker, G.S.: Human Capital: A Theoretical and Empirical Analysis, with Special Referencet o Education, 3rd edn. University Chicago Press, Chicago (1993)
3. Ben-Porath, Y.: The production of human capital and the life cycle of earnings. *J. Polit. Econ.* **75**(4/1), 352–365 (1967)
4. Cahuc, P., Carcillo, S., Zylberberg, A.: Labor Economics. MIT Press, Cambridge (2014)
5. Acemoglu, D., Autor, D.: Lectures in Labor Economics. <https://economics.mit.edu/files/4689>. Accessed 25 Mar 2020
6. Acemoglu, D., Pischke, J.-S.: The structure of wages and investment in general training. *J. Polit. Econ.* **107**(3), 539–572 (1999)
7. Autor, D.: Why do temporary help firms provide free general skills training? *Quart. J. Econ.* **116**(4), 1409–1448 (2001)
8. Croce, G.: A model of training policies in an imperfectly labour market. Working Paper. Dipartimento di Economia Pubblica, Sapienza Università di Roma, vol. 90 (2005). <http://dx.doi.org/10.2139/ssrn.1361212>. Accessed 25 Mar 2020
9. Bassanini, A., Booth, A., Brunello, G., De Paola, M., Leuven, E.: Workplace training in Europe. In: Brunello, G., Garibaldi, P., Wasmer, E. (eds.) *Education and Training in Europe*, pp. 143–342. Oxford University Press, New York (2007)
10. Brunello, G., Rocco, L.: The labor market effects of academic and vocational education over the life cycle: evidence based on a British cohort. *J. Hum. Capital* **11**(1), 106–166 (2015)
11. Bednarska, M.: Complementary person-environment fit as a predictor of job pursuit intentions in the service industry original article. *Contemp. Econ.* **10**(1), 27–38 (2016)



Modelling of Emotional Infection to the Information System Management Project Success

Sergey Bushuyev^(✉) , Denis Bushuiev ,
and Victoria Bushuieva 

Kiev National University of Construction and Architecture, Kiev, Ukraine
sbushuyev@ukr.net

Abstract. Emotional infection is considered as a social and psychological model of transferring the mental mood of the manager to other stakeholders of the information system creation projects, emotional impact in the face of direct contact, and the inclusion of the individual in certain mental states that affect to the success of management. The object of emotional infection technology is the manager or group, the task was to educate, educate and create the organization of new behaviour in neither adverse, complementary but nor contradictory internal and external conditions. In times of crisis, the emotional behaviour of the Information system project manager and his infection with the project team is exacerbated by external uncertainty. The stakeholder infection model is based on an understanding of the life cycle of the project manager, which is presented as a curve for personal changes of the manager of Information system projects and programs. Emotions are considered in content, reflecting the various aspects and meanings that caused them. To apply the psychophysiological model for assessing the impact of the emotional state of the stakeholders of Information system projects success, these influences have been transformed into a competent model for managing these projects. Examples of changes in the competence of the manager and the project team in the case of a wall and asthenic emotions are given.

Keywords: Emotional infection · Model · Information system · Project success · Stakeholders

1 Introduction

The vast majority of human-created technologies are based on the imitation and copying of various natural processes and phenomena. Information system projects and management behaviour are no exception. In management processes, managers try to model the creative behaviour and build on the historical traditions of different cultures. The traditions of the research schools cover various aspects of activity: philosophy, commerce, intelligence, diplomacy and politics. Due to the rapid development of information technologies, a new association has emerged, which is to use computer systems and networks more deeply in modelling activities: artificial systems, databases, processing of big data and emotional intelligence. The trend of such penetration is

growing and expanding, so there is a need for a new organization of innovative activities with broad involvement of information technology and management of emotional processes. Creatively, innovative thinking is the most valuable, open-ended part of human thinking that manifests itself in the form of certain emotions. The health of innovative thinking is a delicate balance between the order and the chaos of stakeholder behaviour. Behaviour is often manifested through a variety of emotions. Scientific advances in this field are limited and far from complete, but at the same time, several common features inherent in this process and the behaviour of innovative project managers can be distinguished. Mission (goal structure) is defined as the result to which the activities of managers are directed. The goals are arranged in a certain sequence, which regulates the rational and emotional activities of the manager. As soon as one of them is reached, a new one arises and so on until the final goal is reached. Each step towards the main goal has a local purpose. Therefore, thinking can rationally organize (profile) the mission. Such an organization serves to manage actions in an innovation project. Many goals and stages of their achievement are profiled in the form of a graph having the structure of a tree. The behaviour of a modern project manager is organized in such a way that thinking is the cause, and action is the consequence (think first, then do it), though it is often the other way around. The peculiarity and a priori uncertainty of the innovation goal is its new quality. In turn, the innovative qualities of the purpose of the concept are relative and depend on subjective evaluation, experience, erudition, emotional intelligence, the benevolence of expertise, public recognition. In times of crisis, the emotional behaviour of the project manager and his infection with the project team is exacerbated by external uncertainty.

The statement of the research problem. The paper examines the model of emotional infection of stakeholders of information system projects and programs in a crisis environment.

The purpose of the article is to formulate the emotional infection model of stakeholders for Information system projects and programs management in a crisis to form rational structures and processes.

2 The Primary Research Materials

Human consciousness contains and uses a wealth of data and knowledge. Human nature reveals the taste and needs to acquire new knowledge and apply it to behaviour in new, circumstances and situations arise every minute. In general, intelligence can be modelled by a pair of sets, or a set of facts and rules, or methods of applying them to achieve a goal. This model was called productive and was used in the early stages of artificial intelligence development [1, 2]. Examples of facts and rules: fact - work is paid; the rule is that if you do the job, you can be rewarded. Here, the rule is conditional: if a condition is met, then some action will occur. The rules establish the necessary sequence of cause and effect to achieve the intended consequence, that is, the result. The facts and rules have different complexity and are organized into a knowledge base. To achieve this goal, you need to be able to link complex sets of facts and rules. The mission of innovative projects is often unclear, so intermediate goals are also vague, if they are bound by fuzzy rules, the task of achieving such a goal is

significantly complicated and unclearly achievable [2, 12]. This circumstance influences the emotional state of the manager and requires certain competencies to handle uncertainty in the context of emotional infection of stakeholders. This creates the conditions of creative risk. Such risks are critical in times of crisis when an innovative project is adversely affected by external factors [3, 4]. Here, it is almost impossible to abandon the simplification, the selection of the main, most essential features of the facts and rules and the rejection of minor ones [16–18].

As mentioned above, in order to share the emotions of other people, it is not necessary to see them in person. The experiments, which were conducted on the basis of Facebook and Twitter, showed that emotions are transmitted in the virtual space.

In some cases, emotional infection can be controlled consciously. Usually we are talking about the fact that the leader of opinions or a group of people, pursuing their own goals, seek to cause some kind of emotions in society. As you know, great power gives rise to great responsibility, so it depends on the leader whether he will use influence in order to inspire people to constructive activities or, say, sow panic, and then offer a “cure” (this strategy is used to get more control).

Managing Information system projects very often appear emotional resonance.

The technique of emotional resonance can be defined as a way of creating a certain mood among a wide audience while transmitting propaganda information. Emotional resonance allows you to remove the psychological defence that a person builds on a mental level, deliberately trying to protect himself from propaganda or advertising brainwashing. One of the basic rules of propaganda is: first of all, you need to turn not to the mind, but to the feelings of man. Defending themselves from propaganda messages, on a rational level, a person is always able to build a system of counterargument and reduce all efforts to “special treatment” to zero. If the propaganda influence on a person occurs on an emotional level, outside his conscious control, no rational counterarguments in this case work.

The effect of emotional infection manifests itself especially strongly in the crowd - a situational set of people who are not connected by a conscious goal (like in case Covid-19). The crowd is a property of a social community, characterized by the similarity of the emotional state of its members. In the crowd there is a mutual infection of emotions and, as a result, their intensification. The nature of mass emotional infection is almost unstudied. One of the interesting hypotheses states that the main role in this is played by the appearance of resonant oscillations in the structure of electromagnetic fields formed by the human body.

The mechanism of human behaviour in a crowd is described in many sources, they all coincide in the fact that a person, becoming part of the masses, falls under the power of passions. Typical signs of human behaviour in a crowd are the prevalence of situational feelings (moods), loss of intelligence, responsibility, hypertrophic suggestibility, easy controllability. These conditions can be enhanced by various means. The necessary moods are caused by the appropriate external environment, a certain time of day, lighting, light stimulants, various theatrical forms, music, songs, etc. In psychology, there is a special term - fascination, which denotes the conditions for increasing the efficiency of perceived material through the use of related background effects. Most often, fascination is used in theatrical performances, game and show programs, political and religious (cult) events, etc. - to infect people in the crowd with a

special emotional state. Against this background, relevant information is transmitted, and we must strive to ensure that it is not too much.

The possibility of emotional infestation opens up scope for manipulation, including in interpersonal relationships. People can aggressively include others in the space of their own experiences in order to receive attention and some benefits for themselves. It is important to separate the request for support and the desire to manage others. However, sometimes manipulations, like the tendency to succumb to them, are unconscious. Recognizing emotions and recognizing their origin can be one way to avoid emotional contagion.

The mere possibility of emotional infestation is not negative or positive. This is a feature of our psychophysiology, because people are primarily social animals [5].

It is the ability to notice the condition of relatives and quickly respond to it, as well as to signal about their condition, has become an important factor in the development of the human species. Warn others about the danger, inspire for joint activities or carry away the game - all this we learned at the dawn of mankind.

However, with the development of intelligence and society, conscious control of these processes has become possible. That is why in some cases it is better to allow yourself to show emotions, and in some - to restrain. This is especially true for managers, teachers and media workers. And, of course, self-control is important for personal and family relationships. When people are in a close emotional relationship, their experiences inevitably affect each other's experiences.

Emotional processes include a wide class of processes, internal regulation of activity. They perform this function, reflecting the meaning that objects and situations have on the subject, their values for the implementation of his life. In humans, emotions give rise to experiences of pleasure, displeasure, fear, timidity, etc., which play the role of orienting subjective signals. The simplest emotional processes are expressed in organic, motor and secretory changes and are among the innate reactions. However, in the course of development, emotions lose their direct instinctual basis, acquire a complex nature, differentiate and form diverse forms of the so-called higher emotional processes: social, intellectual and aesthetic, which a person makes up the main content of his emotional life.

The result is the following classification of emotional process:

- affects - short-term and intense emotional processes, accompanied by pronounced motor manifestations and changes in the work of internal organs;
- emotions are longer and less intense than affects, emotional processes that reflect the subjective meaning of situations, but not of specific objects in themselves;
- feelings are longer and less intense than affects, emotional processes that reflect the subjective meaning of specific objects. For example, hatred;
- moods are quite prolonged emotional processes of low intensity.

Interestingly, in this case, infection occurs without the participation of non-verbal signals. However, the inability to see the poses and facial expressions, as well as to hear the voice of other people, is compensated by new means of transmitting emotions that are used when communicating on the Internet. This is a certain style (you know exactly a certain number of people who abuse capsule and punctuation marks), typical mistakes

that people make when they are nervous and in a hurry, as well as the specifics of using emoticons and stickers. This area is still waiting for its researchers - we give an idea.

3 Emotional Infestation in Project Management

Emotional infection is a socio-psychological mechanism of transmitting the mental attitude to other people from one person or group of people, emotional impact in conditions of direct contact and the inclusion of the individual in certain mental states.

In the study of the emotional component of the processes of managing projects recognition was created by the psychophysiology Simonov P.V. [5] formula, in a short symbolic form represents a set of factors that affect the emergence and nature of the effects of emotions.

$$E = f(P * (In - Is)), \quad (1)$$

where E - emotion, its degree, value and impact; P - the power and influence of the actual need; $(In - IS)$ - assessment of the possibility of meeting the need based on innate and ontogenetic experience; In - information on cost, meeting the need; Is - information about existing values that the manager actually owns.

This formula is practically used to obtain specific quantitative values, but only to illustrate the very principle of the formation of positive or negative emotions of varying strength.

The factors listed above are decisive, necessary and sufficient, but a time factor should also be considered. Emotion can be either short-lived or long-lasting. At the same time, it is necessary to take into account the peculiarities of emotions and individual-typological features of managers of innovative projects. From the formula, it follows that the possibility of satisfying the need influences the sign of emotion. A function that reflects emotions is the same as an evaluation function.

Emotional infestation is a social and psychological mechanism of transfer of mental mood to other people from one person or group of people, emotional influence in the conditions of direct contact and inclusion of a person in certain mental states.

Emotions are different in content, reflecting different aspects of the significance of their situations. To apply the psychophysiological formula for assessing the impact of the emotional state of the stakeholders of innovative projects, we transform these influences into a competent dream model of managing innovative projects.

Consider the intensity of emotions in the implementation of information projects within the competence of the knowledge system ICB in an Agile world [12]. Emotions may vary in intensity (strength). The intensity of emotion in each case is, of course, influenced by a large number of factors within the competence of managing information projects. In general, their contribution makes it possible to estimate Simon's formula.

Emotions are different in content, reflecting different aspects of the innovation project and the situations that caused them situations.

Consider how emotional infection of stakeholders in information projects is formed through the competency system of the IPMA Reference Guide ICB4 in an Agile World Version 2.3 example [12].

Criteria for assessing competence in managing innovative projects and programs are determined based on a taxonomy consisting of 10 criteria [12].

1. Self-reflection and self-management.
2. Personal Integrity and Reliability.
3. Personal Communication.
4. Relations and Engagement.
5. Leadership.
6. Teamwork.
7. Conflict and Crisis.
8. Resourcefulness.
9. Negotiation.
10. Results Orientation.

Define Leadership and Teamwork according (1) as the power P . Part of competences Self-reflection, Personal Integrity, Personal Communication, Relations and Engagement define like value Is . Competences Conflict and Crisis, Resourcefulness, Negotiation, Results Orientation defines like cost In .

These criteria focus on individual values, psychology and ethics. These competencies are drivers of emotional infection.

Four key principles: focus on customer satisfaction, foregrounding, teamwork, and focus on excellence, have been imported into project management and are the prevailing principles of intellectual project space [6–8]. Adherence to these key principles directs the efforts of the project team and assists them in creating new, unique management ideas and methods [9, 10]. For example, the desire of the project team to comply with the terms of the contract to avoid litigation and compromise with the counterparty can lead to both positive and negative results. However, the negative can be minimized if the intellectual space of the project works properly [11, 13, 14]. Good and balanced teams are a good example to follow and a lesson for colleagues in other projects. Not only members of the project team, but also other participants involved in the project in one way or another, including service companies, agencies responsible for forming a temporary staff, etc. - all of them is directly affected by the project product, or the project implementation process. The stakeholder is, therefore, a generic term that defines all institutions, companies and individuals who are directly or indirectly affected by the project. How to resist of penetration into project team emotional infestation?

If you feel that after a short conversation with a person or communication on social networks, there was discomfort, most likely you were infected with negative emotions from the interlocutor. How to counter this?

1. Calm down. If you are angry, it is best to stop communicating and try to calm down. Then you should ask yourself the question: what is happening now? What is the purpose of your communication? Perhaps you need to stop communicating at all or

try to shift the conversation into a more constructive channel a little later, when emotions calm down.

2. Develop critical thinking. No matter what the interlocutor tells you, it is important to keep a cool head, to collect and analyse information, to compare facts. Then it will be harder to manipulate you.
3. To be here and now. This is a good habit, helping to realize that at the moment you are infected with negative from the interlocutor. Try to concentrate on something positive, pleasant, in order to regain your good mood.
4. Spread your shoulders. Our brain is designed in such a way that if we straighten our shoulders, inhale deeply and straighten our back, it reads this as information that everything is fine with us. After some time, we actually begin to feel much better. Therefore, if you feel that you have become infected with negativity, simply straighten your shoulders and smile.

During the implementation of the program, professionals from different fields of knowledge with different skills cooperate to achieve the mission of the program [15, 16]. The community is a space of partnership and competency, in which the professional competencies of the participants are concentrated, the competence of the team is formed, as well as the cooperation between professionals is encouraged to create a strong teamwork potential. At the program level, the most important thing is to set up interfaces for interaction between organizations and program team members.

Let look on the conceptual model of organizational development programs in a turbulent environment (Fig. 1).

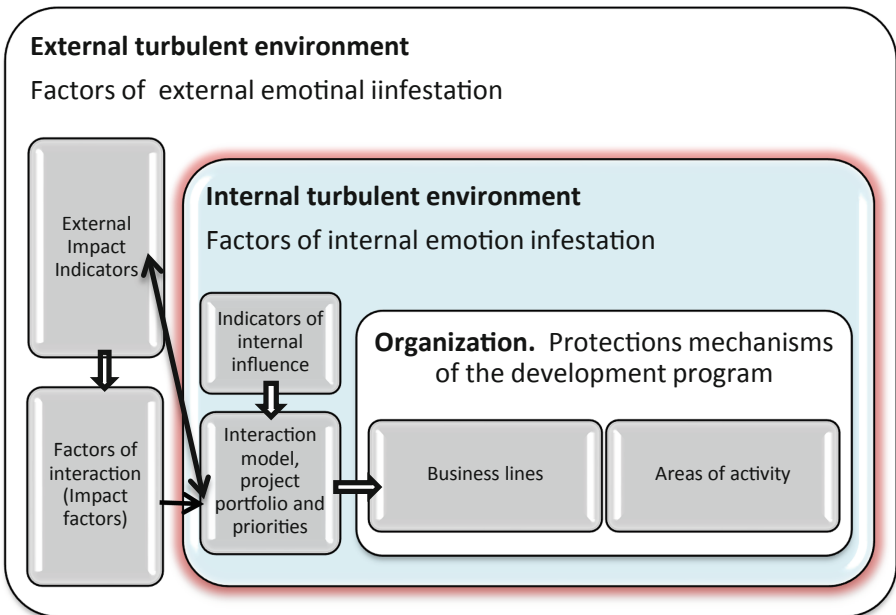


Fig. 1. Conceptual model of organizational development programs in a turbulent environment.

Corporate consistency between programs, projects, and individual tasks must be ensured in the area of organizational development management. This has to do with the coherence of corporate action on the project, part of the resources, knowledge and synergy of programs and projects.

The stakeholder infestation model is based on an understanding of the life cycle of the project manager, which is presented as a Kubler-Ross curve for personal changes of the manager of innovative projects and programs (Fig. 2).

On this curve, we see the initial phase of change of effective activity within three steps - “shock, surprise, reflection”, “insensitivity” and “denial”. This is a short-term phase where the wall and asthenic stains are infected. These infections usually do not extend beyond the project management team. The second phase is related to a significant drop in performance. These are “blaming yourself and others,” “panic and fear,” “depression and danger.” At this stage, an asthenic infection is formed that goes beyond the project management team.

The third phase involves the transition from asthenic to wall infection of the stakeholders of the innovation project. In this phase, the following factors are formed - “acceptance of ignorance”, “testing and verification”, “feeling of optimism, hope and restart”, “opening, learning”, “feeling of satisfaction” and “integration and new understanding” of the innovation project.

A negative impact “asthenic” is blindly following someone else’s mood and reactions, for example, when a “breeding ground” of gossip and negativity appears in the team, which changes the mood of the masses, causing anxiety and distrust of colleagues, which, of course, affects labour efficiency. Unconsciously “mirroring” the emotions of other people, we really begin to feel the same emotions as they are.

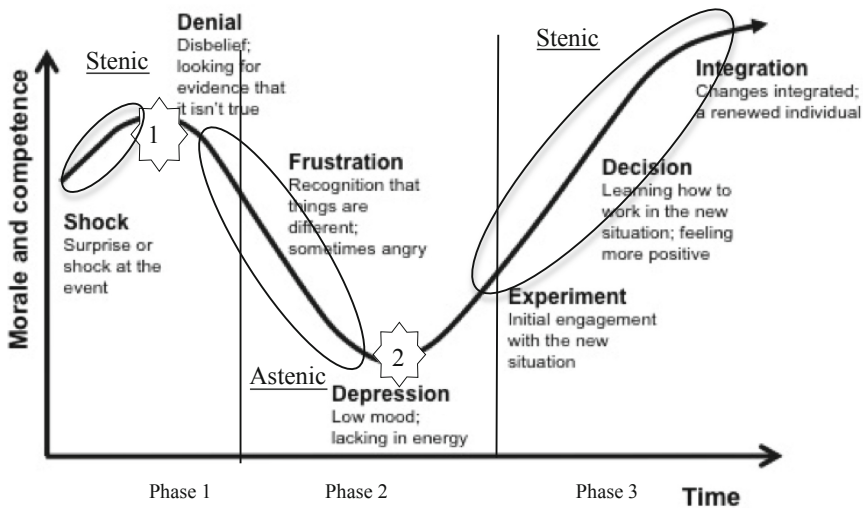


Fig. 2. The curve of personal changes of innovative projects and programs manager.

The positive effect “stenic” is associated with self-control. We try to manage our own emotions, notice changes and control our condition. For example, a confident leader can easily convey an optimistic attitude to his team. Here for the manager it is important to give people a positive charge, set up and motivate them to succeed in their work.

The simulation results of the implementation of the Knowledge Base Creation Project at the Ukrainian Ministry of Finance are shown in Table 1 and Table 2. At the same time, “score 1” was considered as an assessment of the current state of the information project before infection, “score 2”, after infection of the project stakeholders.

Table 1. Assessment of changes in the competence of the manager and team of the project with “asthenic emotions”

№	Criterion name	Score 1	Score 2
1	Self-reflection and self-management	6	4
2	Personal Integrity and Reliability	5	3
3	Personal Communication	5	3
4	Relations and Engagement	6	3
5	Leadership	7	6
6	Teamwork	7	6
7	Conflict and Crisis	6	3
8	Resourcefulness	5	4
9	Negotiation	6	4
10	Results Orientation	6	4

Table 2. Assessment of changes in the competence of the project manager and team with “stenic emotions”

№	Criterion name	Score 1	Score 2
1	Self-reflection and self-management	5	8
2	Personal Integrity and Reliability	5	9
3	Personal Communication	4	7
4	Relations and Engagement	6	9
5	Leadership	6	8
6	Teamwork	6	8
7	Conflict and Crisis	5	9
8	Resourcefulness	5	8
9	Negotiation	5	8
10	Results Orientation	5	9

Key competencies for creating intelligent program support and stakeholder engagement for innovative projects are Strategic Thinking, Integral Thinking, Leadership, Coordination, Communication Skills (provided by information systems, databases and knowledge), motivated, professional initiatives.

The results of the assessment of changes in the competence of the manager and team of the project with “asthenic emotions” are shown in Fig. 3.

The chart of changes in the level of competence of managers of innovative projects (Fig. 3) shows significant drop competence in the context of the influence of asthenic emotions. In this case, the coefficient calculated as $K = (\text{Score 1})/(\text{Score 2})$ in this case is equal to 1.55, which indicates the negative impact of infection on the application of the competence of the project manager.

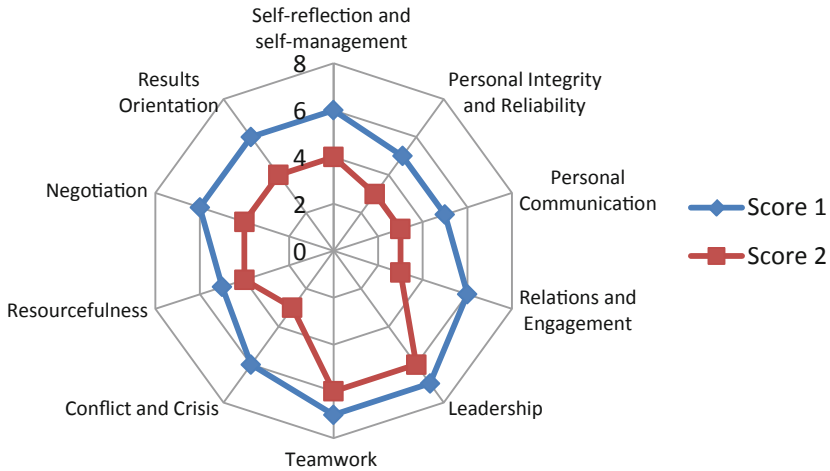


Fig. 3. Assessment of changes in the competence of the manager and team of the project with “asthenic” emotions”



Fig. 4. Assessment of changes in the competence of the project manager and team with “stenic emotions”

The graph of changes in the level of competence of managers of innovative projects (Fig. 4) shows significant increase incompetence in the context of the influence of wall emotions.

In this case, the coefficient K is 0.63, which indicates the positive impact of stakeholder infection on the competence of the innovation project manager.

The application of the proposed model and the approach to assessing the competencies of the manager and team of the innovation project is conceptually understandable and fully proven as to the adequacy of the model.

4 Conclusion

Development of a model of emotional infection of stakeholders of innovative projects and programs in crisis conditions allows specifying the trajectories of project implementation in the conditions of crisis of internal and external environment.

The results of modelling in the implementation of an information technology Knowledge base creation project at the Ministry of Finance determined that the key competencies for creating intellectual support for innovative projects and stakeholders infection are strategic thinking, integrated thinking, leadership, coordination, communication skills systems, databases and knowledge) as well as motivated, proactive and professional leaders of the program team.

References

1. Bushuyev, S., Kozyr, B., Rusan, N.: Empathy, emotional intelligence and transformational leadership to the project success. In: *Advances in Intelligent Systems and Computing*, vol. 1019, pp. 209–222. Springer (2020)
2. Verenyich, O., Dorosh, M.: Blended mental space: the methodology for creation and approaches for it's management (in English). In: *Computer Sciences and Information Technologies (CSIT)*, vol. 2, pp. 30–37 (2017). <https://doi.org/10.1109/STC-CSIT.2017.8099423>
3. Covy, S.: *7 Skills of Highly Effective People: Powerful Personal Development Tools*, p. 375. Alpina Businessbooks, Moscow (2007)
4. Azarov, N., Yaroshenko, F.: *Innovative Mechanisms for Managing Development Programs*, p. 564. Summit Book (2011)
5. Krysko, V.G.: *Social Psychology: Lecture Course*, 3rd edn, p. 352. Omega-L (2006)
6. Goleman, D.: *Emotional Intelligence in Business*, p. 356. Mann, Ivanov and Ferber, Moscow (2013)
7. Imai, M.: *Gemba Kaidzen: Way to Reduce Costs and Improve Quality*, p. 345. Alpina Business Books, Moscow (2007)
8. Decety, J., Ickes, W.: *The Social Neuroscience of Empathy (Social Neuroscience)*. MIT Press, Cambridge (2011)
9. Obradović, V., Todorović, M., Bushuyev, S.: Sustainability and agility in project management: contradictory or complementary?, pp. 522–535. Springer (2018)

10. Rusan, N., Bushuyev, S., Bushuyev, D.: Emotional intelligence—the driver of the development of breakthrough competences of the project. In: IEEE International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, pp. 1–7 (2017)
11. IPMA Organisational Competence Baseline (IPMA OCB). IPMA, p. 67 (2013)
12. IPMA Reference Guide ICB4 in an Agile World Version 2.3, p. 71 (2018)
13. Guide to managing innovative projects and programs, P2M, V.1.2, p. 173 (2009)
14. Individual Competence Baseline for Project, Programme & Portfolio Management, Version 4. International Project Management Association, p. 415 (2015)
15. Bushuyev, S., Verenych, O.: The blended mental space: mobility and flexibility as characteristics of project/program success. In: IEEE 13th International Conference on Computer Sciences and Information Technologies, CSIT (2018)
16. Archibald, R.: Management of High-Tech Programs and Projects, p. 472. DMK Press (2004)
17. Goleman, D., Boyatzis, R., McKee, A.: Primal Leadership, With a New Preface by the Authors: Unleashing the Power of Emotional Intelligence. Harvard Business Review Press, p. 336, 2 July 2013
18. Brackett, M., Rivers, S., Salovey, P.: Emotional intelligence: implications for personal, social, academic, and workplace success. *Soc. Pers. Psychol. Compass* **V5**(1), 88–103 (2011)



Modeling of Project Portfolio Management Process by CART Algorithm

Nataliia Yehorchenkova¹  and Oleksii Yehorchenkov² 

¹ Kyiv National University of Construction and Architecture,
Povitroflotsky Avenue 31, Kyiv 03037, Ukraine
realnata@ukr.net

² Taras Shevchenko National University of Kyiv, Volodymyrska Street, 64/13,
Kyiv 01601, Ukraine

Abstract. In the paper authors propose to use a machine learning approach, namely decision tree algorithm, for project selection in the portfolio of the project-oriented company. The analysis of references showed that today there are exist a lot of research is dedicated to selection and evaluation projects in a project portfolio. And this means that the problem is actual and requiring further investigation for searching the best solution. The paper's objective is the investigation of using a decision tree and evaluate the efficiency of the developed model for the task of project selection. Tasks of the paper: to develop a decision tree model for the project portfolio dataset, to evaluate the decision tree model, to describe the advantages and disadvantages of the developed model. As a result of research, it was creating a decision tree model, that allows predicting a project successful with accuracy 99%. This is a good result that allows the management of the company to make a quick decision that will allow with a high degree of probability to determine which project is best to invest in.

Keywords: Decision tree · Machine learning · Project portfolio management · CART

1 Introduction

1.1 Problem Statement

The implementation of a large number of projects in a project-oriented company in limited resources demands clear rules of investing funds. When a new project is initialized in a portfolio of the company it becomes necessary to check the feasibility of financing: the importance of the project for the company (its priority), probability of its successful finishing, resource availability, and so on. The success of the company depends on how correctly the decision on the projects was made. Evaluating and selecting a project in a portfolio is a complex process involving many factors and considerations from the time the project is proposed to the time it is finally selected. Given that making a good evaluation and selection is of crucial importance, it is essential to develop well-founded mathematical models to lead the organization to its final goal. To achieve this, such models have to reflect as closely as possible both the real situation of the organization as well as its targets and preferences. However, since

the process of selecting and implementing project portfolios occurs in real environments and not in laboratories, uncertainty and a lack of knowledge regarding some data is always an important issue due to the strong interdependence between the projects and the political, economic, social, and legal conditions in which they are carried out [1]. So, the main issue for management of the project-oriented company is debugging the process of selecting or rejecting a project for a portfolio. So, the scientists and developers are faced with the task of creating a system that will help decide on the inclusion of projects in the portfolio.

1.2 Relevant Scientific Researches

Today there is a lot of research for solving such issues. For example, in the paper [2] a mathematical model that extends the classical approach incorporating the inherent uncertainty to these problems is proposed. In the paper [2], the project portfolio has been modeled using four basic dimensions including technology, complexity, innovation, and time sensitivity. The aim is to plan and control the progress of the project portfolio while maximizing the strategic adaptation subject to the changes of the human resources. In the paper [3], authors extend earlier preference elicitation approaches by allowing the decision-maker to make direct statements about the selection and rejection of individual projects. Authors convert such project preference statements to weight information by determining the weights for which (i) the selected project is included in all potentially optimal or non-dominated portfolios, or (ii) the rejected project is not included in any potentially optimal or non-dominated portfolio. Authors [4] found that the existing methods for selecting and managing the projects within a portfolio do not focus enough on how interdependencies can disrupt the balance of the whole project portfolio system. To tackle the problem, the authors propose a systemic approach using the Viable System Model to enhance the way that these interdependencies can be effectively managed because of the overall portfolio system. Besides scientific work there is practical development, for example, Project Expert [5]. So, the analysis of references shows that today a lot of researches is dedicated to selection and evaluation projects in a project portfolio. And it is mean that this problem is actual and requiring further investigation for searching the best solution.

In the paper authors propose to use a machine learning approach – classification and regression trees (CART) algorithm for project selection in the portfolio of the project-oriented company. Machine learning methods are popular in the modeling of decision-making processes, including project selection. For example, in paper [6] authors show that structure at the instance level is tightly connected to algorithm performance, and demonstrate that different machine learning and modeling methodologies, specifically Decision Trees (DT), Case-Based Reasoning (CBR) and Multinomial Logistic Regression (MLR), can be used to perform effective algorithm portfolio selection. And authors of the paper [7] compare the results of carrying out portfolio analysis and optimization of a collection of exploration projects that have been modeled using both decision tree and stochastic simulation approaches.

Authors have an idea to provide research that will check by experiment what approach of machine learning (decision tree, random forest, logistic regression, etc.) will have a result for the task of project selection to the portfolio of the project-oriented

company. In perspective such research will allow us to develop a computer system that can evaluate and select the most profitable projects based on research results.

1.3 Paper's Objective and Tasks

The paper's objective is the investigation of a CART using for the task of project selection and evaluation of the efficiency of the developed model.

Tasks of the paper:

- to develop a CART model for project portfolio dataset;
- to evaluate the decision tree model;
- to describe the advantages and disadvantages of the developed model.

2 Modeling of a Decision Tree for Project Portfolio Selecting

2.1 Definition of a Decision Tree and CART Algorithm

Decision tree learning is one of the predictive modeling approaches used in statistics, data mining, and machine learning. It uses a decision tree (as a predictive model) to go from observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves). Tree models where the target variable can take a discrete set of values are called classification trees; in these tree structures, leaves represent class labels, and branches represent conjunctions of features that lead to those class labels. Classification and regression trees are a non-parametric decision tree learning technique that produces either classification or regression trees, depending on whether the dependent variable is categorical or numeric, respectively [8]. At the heart of the popular algorithms for decision tree construction lies the principle of greedy maximization of information gain: at each step, the algorithm chooses the variable that gives the greatest information gain upon splitting [9].

2.2 Decision Tree Model Construction for Project Portfolio Selection

Every portfolio consists of a plurality of projects defined by different features that considered in a project application. Every project application can consist of quantitative (cost, duration, data of project start/finish, project profit) and qualitative (project name, name of project initiator, project priority, country, city, and so on) values and information about project accepted or rejected. So, when a company implements a lot of projects during a time, it accumulates the database of project applications and when a new project is initialized there is a possibility to take advantage of this database for deciding the fate of this project. In reality, management of the company when deciding on project selection is guided by many factors, that can be as formal (base on the features) as not formal (intuition, personal relationships, and so on). The influence of not formal factors on the decision-making process often can reduce the probability of a successful project selection, so it is necessary to minimize these factors. This

minimization can be implemented by constructing a decision tree. Decision tree allows one to make a solution based only on formal factors.

For the construction of the decision tree model was developed in the next algorithm. An original algorithm [10] was adapted under the project environment, namely, there were changed arguments for the possibility of using the model for creating a decision tree based on the project portfolio dataset.

1. Compute the entropy for project portfolio dataset:

$$\begin{aligned}
 H(S) &= \sum_{c \in C} -p(c) \log_2 p(c), \\
 C &= \{accepted, rejected\},
 \end{aligned}
 \tag{1}$$

where, $H(S)$ - the entropy of project portfolio;
 $p(c)$ - probability of project acceptance or rejection.

$$\begin{aligned}
 p_{acc} &= -\left(\frac{n^a}{n^a + n^r}\right) \cdot \log_2 \left(\frac{n^a}{n^a + n^r}\right), \\
 p_{rej} &= -\left(\frac{n^r}{n^a + n^r}\right) \cdot \log_2 \left(\frac{n^r}{n^a + n^r}\right), \\
 H(S) &= p_{acc} + p_{rej},
 \end{aligned}
 \tag{2}$$

where, p_{acc} - entropy of accepted projects;
 p_{rej} - probability of rejected projects;
 n^a - number of accepted projects;
 n^r - number of rejected projects.

2. To calculate the entropy and information gain for every project feature.
 2.1. Calculate entropy for every feature value.

$$\begin{aligned}
 E(f_i(v_j)) &= -p(v_j^a) \log_2 p(v_j^a) - p(v_j^r) \log_2 p(v_j^r), i = 1..N, j \\
 &= 1..M.
 \end{aligned}
 \tag{3}$$

where, v_j - value of project feature;
 f_i - project feature;
 $E(f_i(v_j))$ - entropy of value v_j ;
 $p(v_j^a)$ - probability of value acceptance of the feature;
 $p(v_j^r)$ - probability of value rejection of the feature.

2.2. Calculate average entropy information for project features.

$$I(f_i) = \sum_{j=1}^M \frac{n_{ij}}{n^a + n^r}, \quad (4)$$

where, $I(f_i)$ - entropy information;
 n_{ij} - number of value j in feature i ;
 M - number of values.

2.3. Calculate Information Gain.

$$IG(f_i) = H(S) - I(f_i). \quad (5)$$

where, $IG(f_i)$ - Information Gain.

3. Calculate Information Gain for every feature of the dataset and pick the highest gain attribute. The root will be with the highest Information Gain.

Repeat the same thing for sub-trees till will get the tree.

3 Experiment Conducting and Result of Research

3.1 Experiment Conducting

For model checking an experiment was conducted. For this experiment a dataset with Kickstarter projects [11] was chosen. The dataset has 15 columns (features) and 378661 rows. Kickstarter is an American public-benefit corporation [12] based in Brooklyn, New York, that maintains a global crowdfunding platform focused on creativity [13]. The company's stated mission is to "help bring creative projects to life" [14]. People who back Kickstarter projects are offered tangible rewards or experiences in exchange for their pledges [15].

For development the decision tree was used programming language Python and environment Jupiter Notebook.

Figure 1 is shown the logical structure of the experiment.

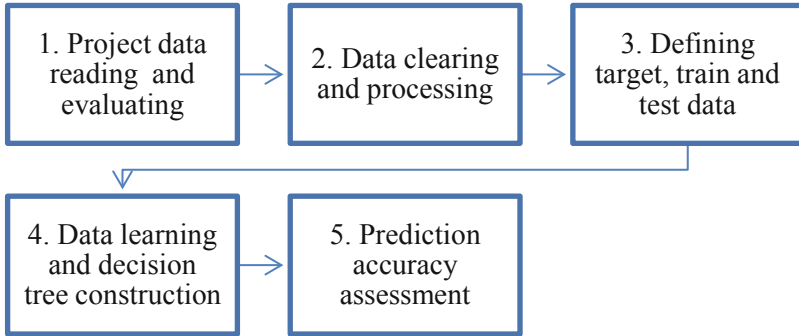


Fig. 1. The logical structure of the experiment

1. Project data reading and evaluating

Table 1 is presented a fragment of the row dataset.

Table 1. Fragment of the dataset (transformed table)

	1	...
ID	1E+09	...
name	Greeting From Earth: ZGAC Arts Capsule For ET	...
category	Narrative Film	...
main_category	Film & Video	...
currency	USD	...
deadline	2017-11-01	...
goal	30000	...
launched	2017-09-02 04:43:57	...
pledged	2421	...
state	failed	...
backers	15	...
Country	US	...
usd pledged	100	...
usd_pledged_real	2421	...
usd_goal_real	30000	...

2. Data clearing and processing

Preparing dataset to learning and development of decision tree dataset was processed. There was removed unnecessary features: ‘ID’, ‘name’, ‘currency’, ‘launched’, ‘deadline’, ‘currency’, ‘usd pledged’, ‘usd_pledged_real’, ‘usd_goal_real’. Features were removed to reduce collinearity in the dataset (Fig. 2). An important stage of

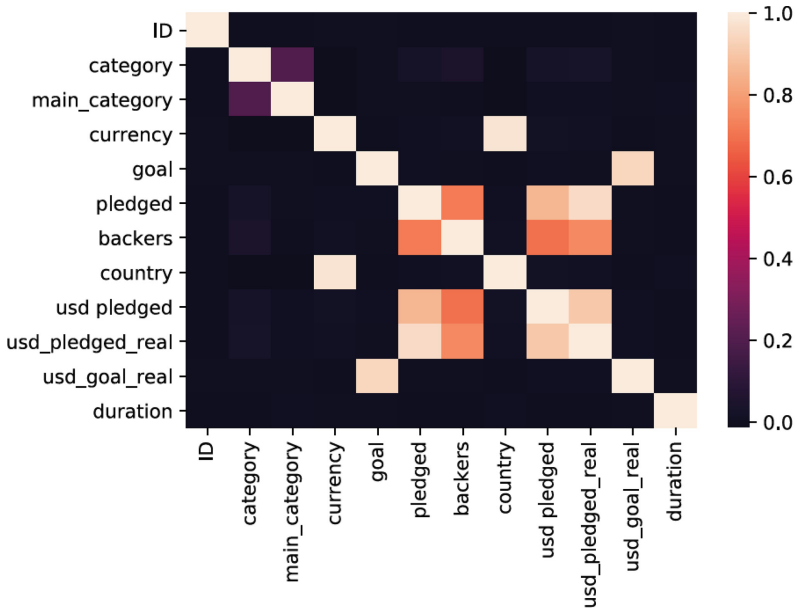


Fig. 2. Heatmap of collinearity of dataset features

Table 2. Fragment of processed dataset (transformed table)

	1
goal	30000
pledged	2421
state	0
backers	15
duration	59
category_3D Printing	0
category_Academic	0
category_Accessories	0
category_Animals	0
country_JP	0
...	...

Note: full table consists from 202 rows \times 331675 columns

dataset processing is the transformation of categorical variables to columns using one-hot encoding. Also, there was added column ‘duration’ that defined as a difference between ‘deadline’ and ‘launched’ saved in days. The fragment of the processed dataset is presented in Table 2.

3. Defining target, train, and test data

As a target was defined as a feature “State”. This feature shows the status of projects. “State” has categories: failed (197719 projects), successful (133956 projects), canceled (38779 projects), undefined (3562 projects), live (2799 projects), suspended (1846 projects). For dataset learning was left only “failed” and “successful” projects.

For the learning decision tree and evaluation of its accuracy dataset was split in the ratio of 70% train to 30% test data.

4. Data learning and decision tree construction.

An algorithm of decision tree learning is presented in Fig. 3.

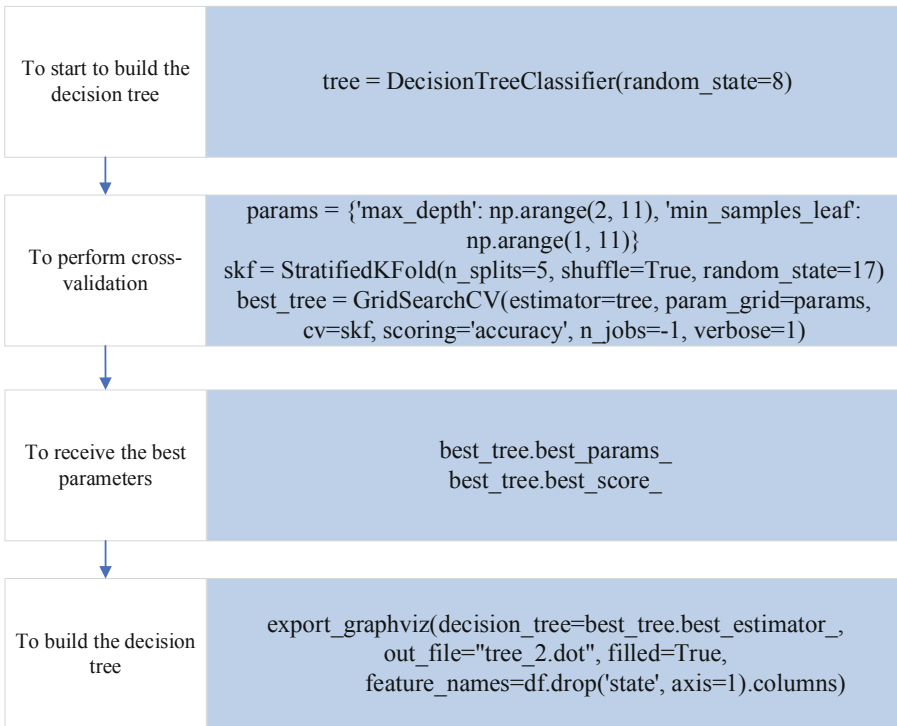


Fig. 3. An algorithm of decision tree

Best model hyperparameters are found with cross-validation:

`{'max_depth': 9, 'min_samples_leaf': 10}`

Final decision tree is presented on Fig. 4.

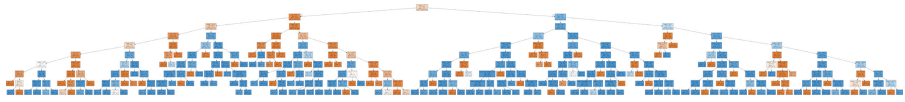


Fig. 4. Final decision tree

As far as the decision tree has a complex structure a fragment is shown in Fig. 5.



Fig. 5. Fragment of the decision tree

5. Prediction accuracy assessment

In the process of decision tree construction was defined prediction accuracy assessment by ROC AUC metric (0.998280954146071). According to this metric, the accuracy of the decision tree prediction is 99%.

Note: all experiment data and algorithms are available in the Github repository: https://github.com/Yehorchenkov/MODS2020?fbclid=IwAR2l_7F9_CcWM5xv9wjmF4FkGjGJc_UUPTwcMpqDKxjfwsvAoTCF0u8QW_c

3.2 Advantages and Disadvantages of Using the Decision Tree in the Project Portfolio Selecting Process

Using a decision tree for project selection in a portfolio of a company has as advantages and disadvantages:

Advantages:

- Decision tree allows one to make a decision only based on formal indicators.
- High accuracy of decision tree prediction.
- The decision tree has a simple algorithm of construction.
- It is possible to develop a decision tree fast and cheap.

Disadvantages:

- To select a project in the portfolio, the decision tree makes only a superficial prediction at the stage “should be considered carefully” or “reject” the project. If a tree predicts success for several projects, and you need to select only one, then such a tool as a decision tree is no longer enough. It is necessary to look for deeper mechanisms for decision making.
- Decision trees can be unstable because small variations in the data might result in a completely different tree being generated.
- Decision tree learners create biased trees if some classes dominate. It is therefore recommended to balance the dataset before fitting with the decision tree.

3.3 Summary

The experiment shows that when initiating a new project using a decision tree with an accuracy of 99%, it can be predicted whether this project will be successful. This is a good result that allows the management of the company to make a quick decision that will allow with a high degree of probability to determine which project is best to invest in. Also, a decision tree is available and easy to use tool which has a simple and cheap implementation. All you need it to prepare a dataset and construct a decision tree. For the construction decision tree there is no need for a complicated program environment or software. Besides, thanks to the decision tree, it becomes possible to understand what features of the project and their values are important for the success of the project. But on the other hand a decision tree can help only at the first stage of project selection and for more deep analysis there is a need for a more complex tool for machine learning. Also, although decision tree modeling does not need a complex program environment, its construction demands special competence from the developer.

4 Conclusion

In the paper authors propose to consider using of decision tree for the task of project selection to a portfolio of a project-oriented company. It was developed a decision tree model and evaluate this model through an experiment for a project portfolio dataset. In the result of the experiment it was defined that prediction accuracy for the project portfolio dataset is 99%. Also, in the paper authors describe the advantages and disadvantages of the developed model.

Such research allows us to consider using machine learning approaches for project portfolio management. for the suitable approach selection, it is necessary to model and






check every from them. The best solutions will be a base for an information system that will be used for improving project portfolio management efficiency.

References

1. Pérez, F., Gómez, T., Caballero, R., Liern, V.: Project portfolio selection and planning with fuzzy constraints. *Technol. Forecast. Soc. Chang.* **131**, 117–129 (2016)
2. Rad, F.H., Rowzan, S.M.: Designing a hybrid system dynamic model for analyzing the impact of strategic alignment on project portfolio selection. *Simul. Model. Pract. Theory* **89**, 175–194 (2018)
3. Tervonen, T., AhtiSalo, J.L.: Modeling project preferences in multiattribute portfolio decision analysis. *Eur. J. Oper. Res.* **263**(1), 225–239 (2017)
4. Bathallath, S., Smedberg, A., Kjellin, H.: Project interdependency management in IT/IS project portfolios: from a systems perspective. *Procedia Comput. Sci.* **100**, 928–934 (2016)
5. Project Expert - a program for developing business plans and evaluating investment projects. <https://www.expert-systems.com/financial/pe/>. Accessed 16 Apr 2020
6. Gebruers, C., Guerri, A.: Machine learning for portfolio selection using structure at the instance level. In: Wallace, M. (eds.) *Principles and Practice of Constraint Programming*. Springer, Berlin (2004)
7. Erdogan, M., Mudford, B., Chenoweth, G., Holeywell, R., Jakubson, J.: Optimization of Decision Tree and Simulation Portfolio. <https://www.onepetro.org/conference-paper/SPE-68575-MS>. Accessed 16 Apr 2020
8. Kamiński, B., Jakubczyk, M., Szufel, P.: A framework for sensitivity analysis of decision trees. *Central Eur. J. Oper. Res.* **26**(1), 135–159 (2017). <https://doi.org/10.1007/s10100-017-0479-6>
9. Topic 3. Classification, Decision Trees and k Nearest Neighbors. <https://mlcourse.ai/articles/topic3-dt-knn/>. Accessed 16 Apr 2020
10. Sanjeevi, M.: Decision Trees Algorithms. <https://medium.com/deep-math-machine-learning-ai/chapter-4-decision-trees-algorithms-b93975f7a1f1>. Accessed 17 Apr 2020
11. Kickstarter Projects. <https://www.kaggle.com/kemical/kickstarter-projects>. Accessed 16 Apr 2020
12. Kickstarter Focuses Its Mission on Altruism Over Profit. <https://www.nytimes.com/2015/09/21/technology/kickstarters-altruistic-vision-profits-as-the-means-not-the-mission.html>. Accessed 16 Apr 2020
13. Kickstarter crowdfunding site officially launches in Canada. <https://www.cbc.ca/news/business/kickstarter-crowdfunding-site-officially-launches-in-canada-1.1703774>. Accessed 16 Apr 2020
14. Stats. <https://www.kickstarter.com/help/stats>. Accessed 16 Apr 2020
15. The Trivialities and Transcendence of Kickstarter. https://www.nytimes.com/2011/08/07/magazine/the-trivialities-and-transcendence-of-kickstarter.html?_r=1&pagewanted=all. Accessed 16 Apr 2020



Modeling of Information Security Management System in the Project

Serhiy Shkarlet , Mariia Dorosh , Oleksandr Druzhynin ,
Mariia Voitsekhovska , and Iryna Bohdan 

Chernihiv National University of Technology, 95 Shevchenko Street,
Chernihiv 14035, Ukraine

rector@stu.cn.ua, mariyaya5536@gmail.com,
druzhyninalex@gmail.com, m.voitsekhovska@gmail.com,
irakirienko@gmail.com

Abstract. With the intensive development of project management processes, the increasing number of project stakeholders which demand constant access to the up-to-date project information and the large territorial distribution of project modules, the issue of ensuring the information security of a project management system becomes a significant indicator of its quality and reliability. Maintaining the confidentiality and integrity of the project information can significantly increase the level of trust in the project, reduce the risks of delayed execution and cost overruns.

The paper presents the model of formation of information security management system in the project based on the PMBoK and ISO 27001 standards. The model defines the processes of ensuring the information security in all the subject areas of project management. Such a model can be used in the formation of a common information security system in a project, taking into account different methodologies and project management systems. The paper proposes to carry out such an assessment by determining the degree of convergence of the information security systems of the project participants for their further integration.

Keywords: Information security · Project management · Information systems · Distributed team · Convergence

1 Introduction

Today the information society is developing dynamically which causes the increase of the number of different types of interactions using modern information technologies. Interdisciplinary project management practice, by synthesizing advanced information and communicative technologies and by using the active development of computer technology, allows to reduce time and the cost of project implementation significantly, to ensure the high quality of the project itself and to reduce the risk management requirements which are the main factors of a successful project.

Project activity in any field is characterized primarily by an innovative component. It makes the most valuable contribution due to which the latest achievements and

technological developments are created and implemented. Therefore, proper consideration ought to be given to the issues of information security while performing the project activity.

On the other hand, the information space of the projects is a part of the artificially created environment as the result of the information-production activity of the project participants and the information space of the project itself. And each participant of this process influences the security level of business processes which are inextricably linked to the information assets of the project.

The constant increase in the volume and types of cyber-attacks has recently demanded adequate awareness and certain competencies in the field of security of all the information processes participants in the projects. By all means, a project manager does not have to be an expert in information security, but without understanding the importance of organizing it at all stages of the project lifecycle, the risks of reducing trust between the participants are significantly increased and the sense of privacy and security of each individual is reduced.

Effective implementation and the use of project information security management systems while creating the project allows to:

- Manage the availability of information about the project and the internal (personal) information about the project participants including the company secured documents, digital database, data and information devices and cloud servers etc.
- Manage the resistance to different security threats, including data breaches and cyber-attacks.
- Create a single security platform for all the project stakeholders.

According to statistics [1], a significant number of cyber-attacks falls on small projects, therefore, not only information security management systems have to be implemented in large projects, but also the economic effectiveness of the activities needs to be carefully calculated, according to the overall project efficiency.

Almost all of the given features require further research on implementing Information Security Management in the Project Management processes, both in technical and technological directions and in the direction of the personal information security culture of the project stakeholders.

2 The Aim of Research

The aim of research is modeling the information security management system in project based on PMBoK and ISO 27001 standards, which determines processes of information security providing in all Project Management domains.

3 Related Works

For the first time an information security issue was mentioned in BS 7799 and then in ISO/IEC 27001:2013 “Information Security Management System” [2] as an essential part of PM activity.

Today there are three major design approaches to information security management depending on the subject of the study.

The first one defines the peculiarities of managing information security projects in organizations where the object is the security system of the organization.

The second approach determines the security system for development information systems where the object is a product of the project.

The third approach explores exactly the processes of integrating security management systems into the project management system.

Antonio Jose Segovia indicates correctly that it's not about establishing "a methodology to manage projects in the field of information security (for example, use a methodology such as PRINCE2 project management to implement a project of ISO 27001)", as to point out establishing "a methodology to treat the information security in project management (for example, to use a risk management methodology to analyze IS-risks relating to a project)" [1]. The ISO 27001 Standard is straight devoted to the second issue.

As Sunita Verma emphasized in [3], "considering the latest edition ISO 27001:2013, the inclusion of information security is a totally new feature which aims to integrate within different Project Management processes and procedures". According to ISO 27001 Annex A.6.1.5, information security integration and implementation should be carried out despite the project's type and scale.

She also highlights that "all projects whether internal or external need resources, activities to progress, and estimated time for completion of each project as per assigned milestones" [3]. Information Security can be integrated and implemented in different Project Management activities such as:

- Include and properly integrate information security within project objectives and deliverables.
- Implementation of risk assessment in the initial stages of the project.
- Identify and apply treatment for the identified risks during the initiation phase and make sure to implement required security measures for each identified risk.
- Make sure to make the information security policy an obligatory part of all the phases and stages of a project.

Defined activities are inherent in projects with a cascading lifecycle that have clear measurable security goals in order to have a well-secured plan with minimal lope holes for security breach or threat within the methodologies such as PRINCE2 [4], P2M or PMBoK [5].

In the [6] authors defined project management based on a security failure, a failure to maintain the CIA (Confidentiality, Integrity, Availability) triad of information security [7].

Project managers have special interests in all three components of the CIA triad. IT projects warrant special consideration for maintaining confidentiality. The business case for any IT project will include strategic business goals whether the project delivers an exciting new technology or a mundane but essential upgrade to maintain enterprise productivity. IT project documentation also frequently includes intimate details of network and systems architecture that presents an attractive target for industrial espionage and hackers. Failed changes to IT systems can also impact availability and

integrity. Special attention to backups, back-out plans and security risks early in the project will pay big dividends when project rollout leaves little time to consider how to undo the changes made during a Go-Live or react to an unexpected risk occurrence that may cause systems to go down, or cause data loss, corruption or breach. Project managers should develop plans to mitigate risks to the project documentation and methodology itself.

Survey [8] states that InfoSec is one of the critical factors for project success. Disappointingly, the major PM standards i.e. PProjects IN Controlled Environments (PRINCE2), Project Management Body of Knowledge (PMBOK) Guide and ISO 21500 (Guidance on Project Management) all disregard InfoSec. They only discuss risk management, which is a much broader area. Likewise, the various PM methodologies (Agile, Waterfall, Six Sigma etc.) all exclude discussion on InfoSec.

One of researches [1] conducted in 2014 had the following objectives:

- To measure the level of knowledge of IT Project Managers (P-Ms) in InfoSec.
- To measure the importance of InfoSec to P-Ms.
- To investigate how P-Ms rank InfoSec among the critical PM success factors.

Out of the 70 respondents, 75.7% had adequate knowledge in InfoSec and 98.6% regarded InfoSec as an important factor in managing projects. On the contrary, when respondents were asked to rank scope, time, cost, InfoSec and quality in order of importance, scope, cost, quality, time and InfoSec were ranked 1st, 2nd, 3rd, 4th and 5th respectively. It is evident from the findings that, although P-Ms regard InfoSec as an important factor in PM, they place the least value on it.

The study [9] mainly focuses on the importance of information security and its integration with project management in the form of a framework. The purpose of designing this framework is to provide project managers a clear picture of security controls to be adopted in each phase of project management. The designed framework is basically a combination of project management, system development life cycle (SDLC) and essential information security controls. As the basis, the authors indicate the Project life cycle defined by the PMBoK methodology.

As indicated in [10], Scrum, Kanban, Scaled Agile – we will hardly find a conference, a publication or a discussion in the area of digitization and project management that does not deal with these topics. But what does the daily practice of ‘agile’ really look like? How successful are agile approaches? What are the real challenges? What are the methods most commonly used? How many organizations are using scaled agile approaches like SAFe, Nexus, Spotify, LeSS?

“Status Quo (Scaled) Agile” 4th survey gave some interesting answers helping organizations to position their own agile activities and helping to define an appropriate strategy where and how to use the agile on the team level or scaled agile on the level of programs or organizations. The results are based on the answers of more than 600 participants from more than 20 countries. The survey has been designed and conducted by the team of the University of Applied Sciences of Koblenz directed by Prof. Ayelt Komus together with IPMA.

- The majority of users of agile approaches use them selectively or in a mixed form.
- The success rate of agile approaches continues to be rated much more positively than that of classic project management.
- However, a time series comparison shows that agile and classic users are converging.
- Once again, the assessment is very clear that the application of agile approaches has led to improvements in results and efficiency.
- 34% of agile respondents use a scaling framework.
- 74% say that improvements in results and efficiency have been achieved through the use of scaling frameworks.
- Only 15% of the Scaling users consistently observe the specifications of the respective Scaling standards.
- The top 3 reasons to use agile approaches are the time to market, quality and risk reduction.
- The most important challenges for the successful implementation of agile approaches are “internal processes” and “top management”.
- 74% of “consistently agile” employees see change as an integral part of the culture, at least in individual areas, compared to only 38% for classic PM.

The leading frameworks are Nexus, Large-Scale Scrum (LeSS), Scaled Agile Framework (SAFe) and Disciplined Agile (DA) [11].

4 Main Methods and Research

The main methods are structural and mathematical modeling of the information security management system of the project. Determining the technical characteristics of the security of project management information systems in different methodologies was provided by structural comparative analysis.

5 Building a Model of Project Information Security Management System (PISMS)

Including information security systems to specification can significantly increase the confidence of its stakeholders. Paying attention to such a pressing issue will create safe conditions for the collaboration of geographically distributed teams, as well as increase investment attractiveness. Working information access restriction provides an adequate security level to project information assets (as innovative developments, intellectual property, confidentiality of inside information, etc.). E.g., information within the project may be classified as confidential, innovative, personal, and business. This classification is a common basis for project stakeholders’ ISSs. It’s also clear that information security depends on chosen PM methodology.

The PISMS implementation can mainly be carried out within the framework of information communication management processes in the project, which contain PMBoK’s four main process groups:

- Information communication planning.
- Distribution of information.
- Report on project implementation.
- Administrative completion of the project.

The Project's Information Security Culture (PISC) is an integral part of the PISMS. Like organizational (corporate) culture, PISC is a complete system, and the concept of "state" represents for it an ordered set of essential properties that the system has at any given time. When forming a common model of stakeholders' values in international projects, the Project Manager has to deal with such features as national mentality, traditions, religions and cultures diversity. The same applies to the manifestation of human nature in the form of PISC, as indicated in works [12–14] devoted strictly to the influence of traditions, religions, and society environment on ICS. Therefore, the issue of developing an ISC in international (as well as multinational) projects is relevant against the background of a steady increase in their quantity, and requires special research.

When developing plans to manage the various subject areas of the project, the aspects related to the InfoSec must be integrated and sufficiently detailed. The analysis of ISO/IEC 27001 made it possible to form a model of information security within the fields of knowledge of PM (Fig. 1). Thus, at the stage of determining the convergence of the mission, strategy and values of the stakeholders of an international project, the attitudes towards restricted information common to all project participants should be highlighted and recorded. Standardized processes for InfoSec monitoring and management of the project must be standardized and developed.

When designing the estimates, the necessary equipment to ensure proper InfoSec is taken into account at each stage of the project lifecycle. On the other hand, the CyberSec technical system must be supported by the relevant InfoSec competencies of the stakeholders.

Another important step which follows risk analysis (including InfoSec) is choosing a risk management strategy. So, there are several options: eliminate the risk; reduce risk; accept the existing risk; or shift the risk to a digital partner (for example, use cloud solutions). When evaluating the information, the P-Ms decide on possible outsourcing or full autonomy.

6 Security Analysis of Support Systems for Different Project Management Methodologies

It is also worth mentioning that the PISMS can and should be cost-effective by increasing the efficiency of the project activity.

Today, the introduction of information systems namely Trello [15], Asana [16], Jira [17], and YouTrack [18] allows to:

- Increase project lifecycle control, reducing the risk of project failure.
- Improving resource efficiency.
- Due to the proper allocation of resources and participants workload, the work deadlines and the response to problems on the early stages are reduced.

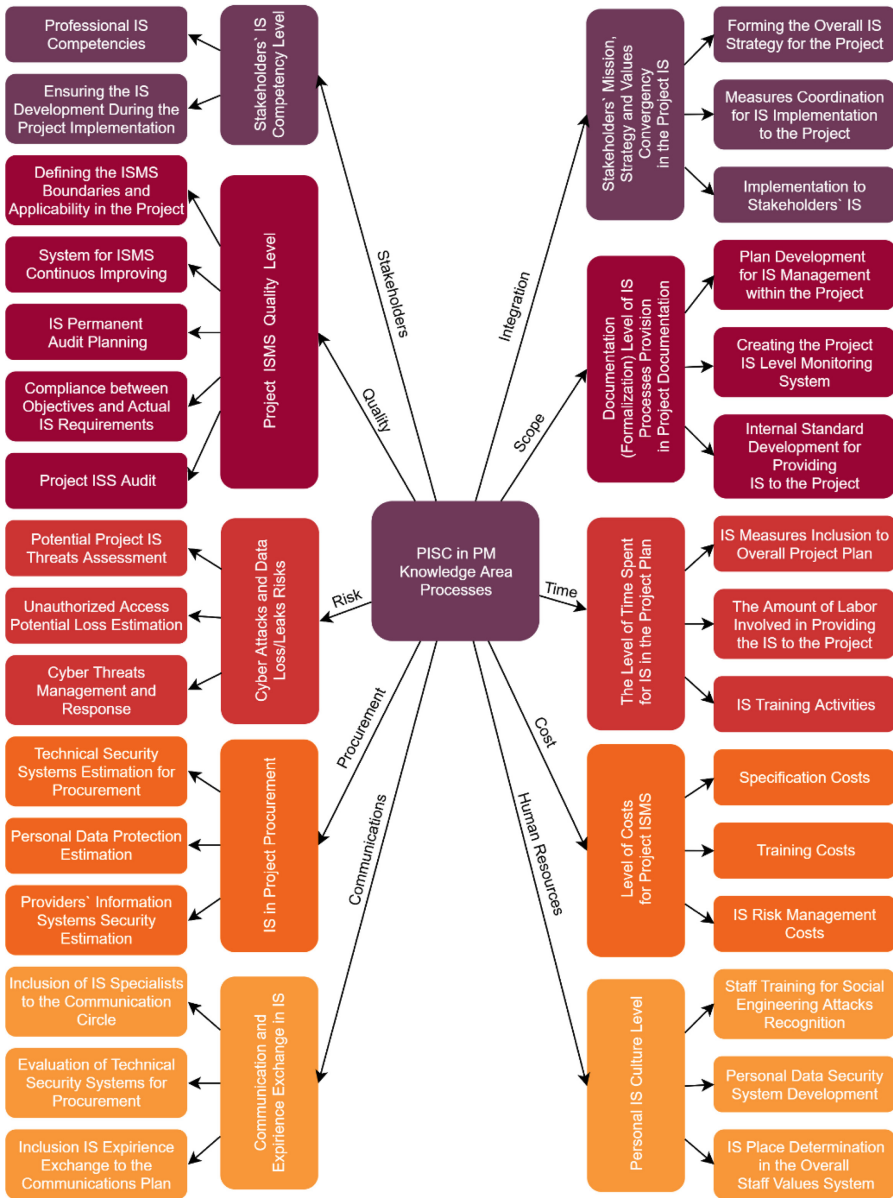


Fig. 1. Model of PISC formation in the Project within PMBoK and ISO 27001

- Reduce budget overruns.
- Reduce the time spent by P-Ms.
- Reduce time to collect data and manually generate project status reports to free up temporary resources for more priority tasks.

Therefore, when implementing project InfoSec, it is necessary to ensure that PIS is implemented in accordance with the known PM methodologies, and PMS supports the required InfoSec level.

Today, the most popular methodologies used by teams are Agile (Scrum, Kanban, Lean). This choice is made for various reasons. However, such a system should not be an obstacle to project development, but rather to significantly increase team productivity. That is why it is critically important that the PMS maintains the team-based methodology. Different information systems are used to provide these methodologies. Let's consider technological security aspects of some project management support systems.

The first step in using an online PMS is registration (gain access). At this point, it is important to remember that transmitting personal sensitive information through unsecured communication channels or in an open form reduces the overall security of the system. It is normal to allow the user to independently create/modify his/her personal data that will be used to identify this user in the system without disclosing it. All of the systems discussed in the paper use the "unique link principle" to enable the users to create or modify their data so that they comply with customized policies and rules in the PMS.

Login/password authentication is one of the simplest and most common methods of user authentication in the WEB environment. If the login and password entered are the same as those stored on the system, that user is granted access, if not - the authentication is considered unsuccessful. All systems discussed in this article support login/password authentication. The systems under review have OTP and/or U2F authentication capabilities. Both of them inherit some OTP vulnerabilities such as phishing and various forms of attack.

Another important security factor is the **communication channels**. HTTPS is already the de-facto standard for WEB systems that handle personal data and valuable information. It enables secure communication in secure communications channels through transparent encryption for the user. So, data transmitted over the network is encrypted, and even if it becomes available to a third party, it will be very difficult to decrypt.

It is also important where the **data will be stored**. Not all PMSs provide the user with the implementation of file storage, while relying on existing solutions. In some cases, the ability to deploy PMS within organization becomes very important as organizations do not trust cloud solutions and their ongoing availability.

In the Table 1 all of the systems listed below allow the use in the cloud, and only Jira and YouTrack can be installed on hardware.

Table 1. Agile supported PM-systems comparison

Params	Agile supported PM-systems			
	Trello	Asana	JIRA	YouTrack
<i>Authentication</i>				
Authentication login/password	+	+	+	+
OTP/U2F authentication	OTP/U2F (via plugins)	OTP only	OTP only	OTP only
<i>Data transfer</i>				
HTTPS support	+	+	+	+
Flexibility of project access restriction policy	Board	Ticket	Ticket	Ticket
Self-sufficiency	Storage (third-party services)	Storage (third-party services)	Storage (own & third-party services)	Storage (own & third-party services)
<i>Accessibility</i>				
Standalone	-	-	+	+
On-Demand in-cloud	+	+	+	+

7 Mathematical Modeling of the PISMS

In forming the information security management system of the project (PISMS) it is possible to use the list of possible variants of interaction of elements of information security systems of project participants formed for morphological analysis.

The search in this list is carried out taking into account the degree of convergence between these elements. The absence of ready-made solutions in the form of defined interactions (when the interaction does not take place) leads to the need to develop (select) new elements of systems and methods of their interaction.

Optimization of the choice of options is proposed for the formation of the PISMS in order to ensure the quality of the overall projected PMS. The method of integer linear programming with Boolean variables is used. We introduce a Boolean variable $x_{ij} = 1$, in case when for i^{th} interaction component (by values or/and technical indicators) of new PMS the j^{th} interaction way is chosen. In other case (when we didn't make decision on j^{th} interaction way) $x_{ij} = 0$.

Taking into account the interaction systems convergence degree the objective function will be as follows:

$$F = \sum_{i=1}^n \sum_{j=1}^m (a_i MCON_{1j} + a_i MCON_{2j} + a_i MCON_{3j} + a_i MCON_{4j}) x_{ij}, \quad (1)$$

where α_i is weight coefficient which takes into account the importance of convergence degree on i^{th} parameter of new PMS;

n – number of parameters for which the convergence degree is determined;
 m – number of possible options system elements interaction by the i^{th} parameter of convergence estimation;
 $MCON_{1j}$ – convergence degree of project stakeholders’ values system in InfoSec domain;
 $MCON_{2j}$ – convergence degree of InfoSec technics support systems of project stakeholders;
 $MCON_{3j}$ – convergence degree of PM methodologies used by project stakeholders;
 $MCON_{4j}$ – convergence degree of project integration systems security.

The restrictions in order of importance are shown as follows:

- Quality assurance requirements to PISMS being created:

$$Q = \sum_{i=1}^n \sum_{j=1}^m q_{ij}x_{ij}. \tag{2}$$

- Costs limitations on introduction of new PMS elements:

$$C = \sum_{i=1}^n \sum_{j=1}^m c_{ij}x_{ij}. \tag{3}$$

- Timing requirements to PMS development:

$$T = \sum_{i=1}^n \sum_{j=1}^m t_{ij}x_{ij} \tag{4}$$

- Requirements for PMS development risks:

$$R = \sum_{i=1}^n \sum_{j=1}^m r_{ij}x_{ij} \tag{5}$$

where q_{ij} is expert evaluation of quality $MCON_{ij}$;

c_{ij} – PISMS costs development for $MCON_{ij}$;

t_{ij} – time spent on PISMS development for $MCON_{ij}$

r_{ij} – risks associated with PISMS elements changes, estimated $MCON_{ij}$ degree, and its relationships/connections.

During optimization, the requirements of the information security project should be fully met (in other words, maximize convergence F): $max F$ subject to the following restrictions:

$$Q \geq Q'; C \leq C'; T \leq T'; R \geq R' \tag{6}$$

where Q', C', T', R' – permissible means for quality, cost, time, and risk metrics for $MCON_{ij}$.

The degree of convergence here is the area of convergence of the information security systems of project participants in a multidimensional space, which determines the completeness of the approximation of these systems.

In general, the formula of calculation can be represented as:

$$Mcon = \sum_{j=1}^N \sum_{i=1}^M \frac{1}{2} CON_i(j; j+1) * CON_i(j+1; j+2) * \sin \frac{360}{N}. \tag{7}$$

where $Mcon$ is a convergence degree of the ISSoPs of stakeholders;

- CON_i , – convergence on the i^{th} element of the ISS of j^{th} and $(j+1)^{th}$ stakeholders;
- N – number of compared elements of stakeholders’ ISSs;
- M – number of stakeholders.

The degree of convergence in this case determines the degree of approximation of information security management systems of project participants in a multidimensional space. Convergence rate estimation (CON) may be due to an Euclidean distance [19].

$$CON = 1 - \sqrt{\sum_{i=1}^N (A_i - B_i)^2}, \tag{8}$$

where A_i and B_i is completeness degree of i^{th} element of ISS for two different project stakeholders;

- N – number of project stakeholders interaction pairs by the i^{th} element.

8 Conclusions

The necessity to take into account information security in project activities is associated with the intensive development of project management processes and is regulated by international standards such as PMBoK and ISO 27001. This paper is devoted to process modeling of project’s information security management system based on mentioned standards. The resulting model allows to ensure an information security through all stages of the project creation, regardless of the chosen methodology and taking into account each project participants’ information security system.

Further research may be aimed at:

- Modeling the integration of security systems of project participants.
- Design of technical security systems in projects.
- Development of modules for existing Project Management software tools, including the information security management processes.

Acknowledgement. This research was supported by CyRADARS project (SPS G5286 “Cyber Rapid Analysis for Defense Awareness of Real-time Situation”) in the frame of the NATO Science for Peace and Security program.

References

1. Segovia, A.J.: How to manage security in project management according to ISO 27001 A.6.1.5. <https://advisera.com/27001academy/blog/2015/07/06/how-to-manage-security-in-project-management-according-to-iso-27001-a-6-1-5/>. Accessed 14 Apr 2020
2. ISO/IEC 27001:2013: Information technology – Security techniques – Information security management systems – Requirements
3. Verma, S.: ISO 27001 Information Security in Project Management. <https://www.sync-resource.com/blog/iso-27001-information-security/>. Accessed 13 Apr 2020
4. PRINCE2 Methodology. <https://www.prince2.com/uk/prince2-methodology>. Accessed 13 Apr 2020
5. Turner, O.: Guide to the Project Management Body of Knowledge PMBOK Guide 6th Edition. Project Management Institute, Pennsylvania (2017)
6. Biswas, P.: ISO 27001:2013 A.6.1.5 Information security in project management. <https://isoconsultantkuwait.com/2019/12/25/iso-270012013-a-6-1-5-information-security-in-project-management/>. Accessed 13 Apr 2020
7. SANS Institute: Security Best practices for IT Project Managers, SANS Reading room (2013). <https://www.sans.org/reading-room/whitepapers/bestprac/paper/34257>. Accessed 13 Apr 2020
8. Information Security in Project Management (2018). <https://www.myjoyonline.com/news/technology/information-security-in-project-management/>. Accessed 13 Apr 2020
9. Sangi, S., Ilkan, M., Tokgöz, H.: Incorporating information security into IT project management (a proposed framework). *IJCCIE* 4(1), 219–226 (2017). <https://doi.org/10.15242/IJCCIE.ER0317112>
10. Wagner, R.: Agile Approaches: Successful – mostly hybrid – often scaled – Results of “Status Quo (Scaled) Agile” Survey (2020). <https://www.ipma.world/agile-approaches-successful-mostly-hybrid-often-scaled-results-of-status-quo-scaled-agile-survey-available/>. Accessed 13 Apr 2020
11. Comparison of Scaling Agile Frameworks: Which one Should you Choose? <https://www.visual-paradigm.com/scrum/scaling-agile-frameworks-comparison/>. Accessed 13 Apr 2020
12. Kruger, H., Flowerday, S., Drevin, L., Steyn, T.: An assessment of the role of cultural factors in information security awareness. In: Information Security South Africa Conference 2011, pp. 1–7 (2011). <https://doi.org/10.1109/ISSA.2011.6027505>
13. ALArifi, A., Tootell, H., Hyland, P.: Information security awareness in Saudi Arabia. In: CONF-IRM 2012 Proceedings, vol. 57 (2012). <http://aisel.aisnet.org/confirm2012/57>

14. Alfawaz, S., Nelson, K., Mohannak, K.: Information security culture: a behaviour compliance conceptual framework. In: Boyd, C., Susilo, W. (eds.) Information Security 2010: AISC 2010 Proceedings of the Eighth Australasian Conference on Information Security, vol. 105, pp. 51–60 (2010). https://www.researchgate.net/publication/40496478_Information_security_culture_A_behaviour_compliance_conceptual_framework
15. Trello. <https://trello.com/>
16. Asana. <https://asana.com/>
17. Jira. <https://www.atlassian.com/ru/software/jira>
18. YouTrack. <https://www.jetbrains.com/ru-ru/youtrack/>
19. Dorosh, M., Trunova, O., Itchenko, D., Voitsekhovska, M., Dvoieglazova, M.: The study of participants' values convergence on the example of international scientific project on cyber security. Eastern-Eur. J. Enterp. Technol. **6/3**(84), 4–10 (2016). <https://doi.org/10.15587/1729-4061.2016.85215>

Author Index

A

Andriichuk, Oleh, [197](#)

B

Babenko, Tetiana, [277](#)

Barabash, Olena, [56](#)

Bohdan, Iryna, [364](#)

Boiko, Serhii, [121](#)

Boychenko, Andriy, [175](#)

Brovchenko, Igor, [36](#)

Buival, Liliia, [101](#)

Burmaka, Ivan, [237](#)

Bushuiev, Denis, [341](#)

Bushuieva, Victoria, [341](#)

Bushuyev, Sergey, [341](#)

C

Chertov, Oleg, [248](#)

Chumak, Anton, [101](#)

Chupryna, Volodymyr, [132](#)

D

Davies, John N., [209](#)

Denisov, Yuri, [121](#)

Didenko, Natalia, [92](#)

Dmytrenko, Oleh, [197](#)

Dmytriiev, Volodymyr, [309](#)

Dorosh, Mariia, [237](#), [364](#)

Druzhynin, Oleksandr, [364](#)

Dyfuchyn, Anton, [266](#)

Dymerets, Andrii, [121](#)

F

Fedorov, Eugene, [156](#)

G

Gerashchenko, Maksym, [289](#)

Gorodny, Alexey, [121](#)

Grebenikov, Oleksandr, [101](#)

Grechaninov, Viktor, [186](#)

H

Hamidov, Galib, [147](#)

Hnatiienko, Hrygorii, [217](#)

Holub, Serhii, [47](#), [227](#)

Hrebennyk, Alla, [255](#)

Hubin, Anton, [69](#)

Hubskyyi, Oleksandr, [277](#)

Humennyi, Andrii, [101](#)

I

Isachenko, Oleksandr, [289](#)

K

Kamak, Dmytro, [309](#)

Khalchenkov, Aleksander, [13](#)

Khmara, Leonid, [92](#)

Kholodov, Anton, [92](#)

Kholodov, Mykhailo, [92](#)

Khoshaba, Oleksandr, [186](#)

Khrutba, Viktoriia, [24](#)

Klets, Dmytro, [92](#)

Koba, Sergiy, [299](#)

Kombarov, Volodymyr, [82](#)

Konyshch, Dmytro, [101](#)

Korniiienko, Ihor, [309](#)

Korniiienko, Svitlana, [309](#)

Kotsiuba, Iryna, [24](#)

Kovalets, Ivan, [3](#), [13](#)

Kritskiy, Dmitriy, [299](#)
 Kruglov, Alexander, [217](#)
 Kunytska, Svitlana, [227](#)

L

Lande, Dmytro, [175](#), [197](#)
 Leontiev, Dmytro, [69](#)
 Litvinchev, Igor, [111](#)
 Lysetskyi, Yurii, [255](#)
 Lytvyn, Svitlana, [237](#)
 Lytvynov, Vitalii, [186](#), [237](#), [255](#)

M

Maderich, Vladimir, [36](#)
 Makarov, Volodymyr, [69](#)
 Malchykov, Volodymyr, [248](#)
 Morozova, Tetiana, [24](#)
 Myntiuk, Vitalii, [82](#)
 Myrutenko, Larysa, [277](#)

N

Neskorodieva, Tatiana, [156](#)
 Nesterenko, Sergey, [289](#)
 Nikiforova, Liliia, [331](#)

O

Oksiuk, Oleksandr, [277](#)

P

Pankratov, Alexander, [111](#)
 Pavlenko, Anatolii, [309](#)
 Plankovskyy, Sergiy, [82](#), [111](#)
 Porplenko, Yaroslava, [197](#)
 Prykhodko, Natalia, [166](#)
 Prykhodko, Sergiy, [166](#)

R

Rieznikov, Oleksandr, [92](#)
 Romanenko, Oleksandr, [3](#)
 Romanova, Tatiana, [111](#)

Rudnichenko, Serhii, [289](#)
 Rudyk, Andriy, [132](#)

S

Shamrai, Volodymyr, [24](#)
 Shkarlet, Serhiy, [364](#)
 Shuklinov, Serhii, [69](#)
 Shyian, Anatolii, [331](#)
 Shypul, Olga, [111](#)
 Skiter, Igor, [255](#)
 Solomakha, Iryna, [209](#)
 Stankevich, Sergey A., [317](#)
 Stetsenko, Inna V., [266](#)
 Stoianov, Nikolai, [237](#)

T

Talakh, Mariia, [47](#)
 Tmienova, Nataliia, [217](#)
 Trunova, Elena, [255](#)
 Tsegelnyk, Yevgen, [82](#), [111](#)
 Tsyganok, Vitaliy, [197](#)

V

Verbitskiy, Victor, [69](#)
 Verovko, Mariya, [209](#)
 Verovko, Oleksandr, [209](#)
 Voitsekhovska, Mariia, [364](#)

W

Weigang, Ganna, [56](#)

Y

Yashin, Sergiy, [299](#)
 Yehorchenkov, Oleksii, [353](#)
 Yehorchenkova, Nataliia, [353](#)
 Yershov, Roman, [121](#)

Z

Zadorozhnyi, Sergiy, [82](#)
 Zavertailo, Kostiantyn, [186](#)
 Zaychenko, Yurii, [147](#)