Serhiy Shkarlet · Anatoliy Morozov ·
Alexander Palagin · Dmitri Vinnikov ·
Nikolai Stoianov · Mark Zhelezniak ·
Volodymyr Kazymyr   *Editors*

Mathematical Modeling and Simulation of Systems

Selected Papers of 16th International
Scientific-practical Conference,
MODS, 2021 June 28–July 01, Chernihiv,
Ukraine

Springer

Lecture Notes in Networks and Systems

Volume 344

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

More information about this series at https://link.springer.com/bookseries/15179

Serhiy Shkarlet · Anatoliy Morozov ·
Alexander Palagin · Dmitri Vinnikov ·
Nikolai Stoianov · Mark Zhelezniak ·
Volodymyr Kazymyr
Editors

Mathematical Modeling and Simulation of Systems

Selected Papers of 16th International
Scientific-practical Conference, MODS, 2021
June 28–July 01, Chernihiv, Ukraine

Editors
Serhiy Shkarlet
The Ministry of Education and Science
of Ukraine
Kyiv, Ukraine

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

Nikolai Stoianov
Bulgarian Defense Institute
Sofia, Bulgaria

Volodymyr Kazymyr
Chernihiv Polytechnic National University
Chernihiv, Ukraine

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

Dmitri Vinnikov
Tallinn University of Technology
Tallinn, Estonia

Mark Zhelezniak
Institute of Environmental Radioactivity
Fukushima University
Fukushima, Japan

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
Wrexham Glyndwr University, 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
Francisk Skorina Gomel State University, Belarus
Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
National Technical University of Ukraine "Kyiv Polytechnic Institute", Ukraine
National University «Yuri Kondratyuk Poltava Polytechnic», Ukraine
Cherkasy State Technological University, Ukraine
Institute for Safety Problems of Nuclear Power Plants of the NASU, Ukraine
Chernihiv Polytechnic National University, Ukraine

Chairs

Alexander Palagin, Academician of NASU, Ukraine, V. M. Glushkov Institute of Cybernetics of the NASU, Ukraine
Ireneusz Zbiciński, Ph.D., D.Sc., Professor, Lodz University of Technology, Poland

Anatoliy Morozov, Academician of the NASU, Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine

V. Onishchenko, D.Sc., Professor, National University «Yuri Kondratyuk Poltava Polytechnic», Ukraine

Enrique Romero-Cadaval, Dr.Sc., Professor, University of Extremadura, Badajoz, Spain

Serhiy Shkarlet, D.Sc., Professor, Ministry of Education and Science of Ukraine, Ukraine

Volodymyr Bashynskyi, Dr. Sc., Chief Sc., State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Ukraine

A. Vasiliev, D.Sc., A. N. Podgorny Institute for Mechanical Engineering Problems of the NASU, Ukraine

Dmitri Vinnikov, Dr.Sc., Tallinn University of Technology, Estonia

Ilya Galkin, Dr.Sc.Ing., Professor, Riga Technical University, Latvia

O. Demidenko, D.Sc., Professor, Francisk Skorina Gomel State University, Belarus

John Davies, Professor, Wrexham Glyndwr University, UK

Program Committee

V. Adamchuk, D.Sc., Professor, The National Academy of Agrarian Sciences of Ukraine, Ukraine

O. Azarov, D.Sc., Professor, Vinnytsia National Technical University, Ukraine

Serhiy Bushuev, D.Sc., Professor, Kyiv National University of Construction and Architecture, Ukraine

Volodymyr Dmytriiev, Ph.D., Chief Sc., State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Ukraine

Mariia Dorosh, D.Sc., Professor, Chernihiv Polytechnic National University, Ukraine

Serhiy Holub, D.Sc., Professor, Cherkasy State Technological University, Ukraine

V. Gryshko, D.Sc., Professor, National University «Yuri Kondratyuk Poltava Polytechnic», Ukraine

Janis Zakis, Dr.Sc.Ing., Riga Technical University, Latvia

Volodymyr Kazymyr, D.Sc., Professor, Chernihiv Polytechnic National University, Ukraine

V. Kharchenko, D.Sc., Professor, National Aerospace University named after N. E. Zhukovsky "Kharkiv Aviation Institute", Ukraine

V. Klimenko, D.Sc., Professor, Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine

Ivan Kovalets, D.Sc., Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine

V. Kraskevych, D.Sc., Professor, Kyiv National University of Trade and Economics, Ukraine

Dmytro Lande, D.Sc., Professor, Institute for Information Recording of NAS of Ukraine, Ukraine

Olexandr Liakhov, D.Sc., Professor, National University «Yuri Kondratyuk Poltava Polytechnic», Ukraine
V. Maderych, D.Sc., Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
V. Mozharovskiy, D.Sc., Professor, Francisk Skorina Gomel State University, Belarus
V. Myronenko, D.Sc., Professor, Member of the NAASU, The National Academy of Agrarian Sciences of Ukraine, Ukraine
A. Nosovskyi, D.Sc., Professor, Corresponding Member of the NASU, Institute for Safety Problems of Nuclear Power Plants of the NASU, Ukraine
V. Snytiuk, D.Sc., Professor, Taras Shevchenko National University of Kyiv, Ukraine
Inna V. Stetsenko, D.Sc., Professor, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
V. Tarasenko, D.Sc., Professor, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
A. Verlan, Dr. Sc., Professor, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
O. Voloshyn, Dr. Sc., Professor, Taras Shevchenko National University of Kyiv, Ukraine
V. Zatserkovnyi, D.Sc., Professor, Taras Shevchenko National University of Kyiv, Ukraine
Mark Zhelezniak, Ph.D., Institute of Environmental Radioactivity, Fukushima University, Japan

Steering Chairs

Volodymyr Kazymyr, D.Sc., Professor, Chernihiv Polytechnic National University, Ukraine
Igor Brovchenko, D.Sc., Institute of Mathematical Machines and Systems Problems of the NASU, Ukraine
Iryna Bilous, Ph.D., Associate Professor, Chernihiv Polytechnic National University, Ukraine
Volodymyr Bazylevych, Ph.D., Associate Professor, Chernihiv Polytechnic National University, Ukraine

Local Organizing Committee

Anna Usik, Chernihiv Polytechnic National University, Ukraine
Iryna Yakymenko, Chernihiv Polytechnic National University, Ukraine

EasyChair Support

Mariia Voitsekhovska, Ph.D., Chernihiv Polytechnic National University, Ukraine

Sponsors

«S&T Ukraine», Kyiv, Ukraine

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

Preface

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 National Academy of Sciences of Ukraine (IMMSP of the NASU) in 2006. MODS is now an annual international conference held by Chernihiv Polytechnic National University with the assistance of the Ministry of Education and Science of Ukraine, the National Academy of Sciences of Ukraine and the State Research Institute for Testing and Certification of Arms and Military Equipment, universities and research organizations from UK, USA, Spain, Bulgaria, Poland, Latvia, Estonia, Belarus and Ukraine participating as co-organizers of the conference.

The XVIth International Scientific-Practical Conference MODS'2021 was held in Chernihiv, Ukraine, during June 28–July 01, 2021.

MODS'2021 received 73 paper 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'2021, which are included in *Lecture Notes in Networks and Systems* 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'2021 made great efforts to ensure the success of this conference. We hereby would like to thank all the members of MODS'2021 Advisory Committee for their guidance and advice, the members of Program Committee and Organizing Committee, 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 *Lecture Notes in Networks and Systems* for their great support in publishing these selected papers.

Kyiv, Ukraine	Serhiy Shkarlet
Kyiv, Ukraine	Anatoliy Morozov
Kyiv, Ukraine	Alexander Palagin
Tallinn, Estonia	Dmitri Vinnikov
Sofia, Bulgaria	Nikolai Stoianov
Fukushima, Japan	Mark Zhelezniak
Chernihiv, Ukraine	Volodymyr Kazymyr

About This 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 event 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 16th International Scientific-Practical Conference, MODS, June 28–July 01, 2021, Chernihiv, Ukraine.

Contents

**Mathematical Modeling and Simulation of Systems in Information
Technology and Information Security**

Mathematical Modeling and Simulation of Special Purpose Equipment Samples

Mathematical Modeling and Simulation of Systems in Project Management

Contributors

Roman Andrushchenko Chernihiv Polytechnic National University, Chernihiv, Ukraine

Kirill Bashinsky 1285Th Office of Military Representatives in the Defense Ministry of Ukraine, Kyiv, Ukraine

Denys Berestov Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Iryna Bilous Chernihiv Polytechnic National University, Chernihiv, Ukraine

Iryna Bohdan Chernihiv Polytechnic National University, Chernihiv, Ukraine

Maya Bozhilova Defence Institute "Prof. Tsvetan Lazarov", Bul. "Prof. Tsvetan Lazarov" 2, Sofia, Bulgaria

Maksym Brazhenenko Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Igor Brovchenko Institute of Mathematical Machine and System Problems, Kyiv, Ukraine

Ivan Burmaka Chernihiv Polytechnic National University, Chernihiv, Ukraine

Oleksii Bychkov Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Serhii Choporov Zaporizhzhya National University, Zaporizhzhya, Ukraine

Volodymyr Chupryna State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Chernihiv, Ukraine

John N. Davies CARDS, Glyndŵr University, Wrexham, UK

Oleh Dmytrenko Institute for Information Recording, National Academy of Sciences of Ukraine, Kyiv, Ukraine;
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Volodymyr Dmytriiev State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Chernihiv, Ukraine

Mariia Dorosh Chernihiv Polytechnic National University, Chernihiv, Ukraine

Anatoliy Doroshenko National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Mykhailo Dubok Institute for Information Recording, National Academy of Sciences of Ukraine, Kyiv, Ukraine

Oleksandr Dveirin Antonov Company, Kyiv, Ukraine

Shamal Faily Poole, Bournemouth, Dorset, Great Britain, UK

Eugene Fedorov Donetsk National University Named Vasyl Stus, Vinnytsia, Ukraine;
Cherkasy State Technological University, Cherkasy, Ukraine

Serhii Florov Dnipro University of Technology, Dnipro, Ukraine

Angel Genchev Defence Institute "Prof. Tsvetan Lazarov", Bul. "Prof. Tsvetan Lazarov" 2, Sofia, Bulgaria

Ievgenii Gorbatyuk Department of Construction Machinery, Kyiv National University of Construction and Architecture, Kyiv, Ukraine

Viktor Grechaninov Institute of Mathematical Machines and Systems Problems of the Ukraine National, Academy of Science, Kyiv, Ukraine

Oleksandr Haidachuk Ningbo University of Technology, Jiangbei Qu, Ningbo Shi, Zhejiang Sheng, China

Maksym Herashchenko State Science and Research Institute of Armament and Military Technics Testing and Certification of the Ukraine Armed Forces, Chernihiv, Ukraine

Serhii Holub Cherkasy State Technological University, Cherkasy, Ukraine

Tetyana Honcharenko Department of Information Technologies of Design and Applied Mathematics, Kyiv National University of Construction and Architecture, Kyiv, Ukraine

Oleksandr Isachenko State Science and Research Institute of Armament and Military Technics Testing and Certification of the Ukraine Armed Forces, Chernihiv, Ukraine

Dmytro Kamak State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Chernihiv, Ukraine

Yurii Kamak State Science and Research Institute of Armament and Military Technics Testing and Certification of the Ukraine Armed Forces, Chernihiv, Ukraine

Liudmyla Kapitanova National Aerospace University «KhAI», Kharkiv, Ukraine

Ihor Karpachev Datascope Systems Ltd., Chester, UK

Oleh Karpenko Dnipro University of Technology, Dnipro, Ukraine

Vasilis Katos Poole, Bournemouth, Dorset, Great Britain, UK

Volodymyr Kazymyr Chernihiv Polytechnic National University, Chernihiv, Ukraine

Larysa Khalanchuk Dmytro Motornyi Tavria State Agrotechnological University, Melitopol, Ukraine

Iulia Khlevna Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Oleksandr Khoshaba Institute of Mathematical Machines and Systems Problems of the Ukraine National, Academy of Science, Kyiv, Ukraine

Duncan Ki-Aries Poole, Bournemouth, Dorset, Great Britain, UK

Anastasiia Kiian V. N. Karazin, Kharkiv National University, Kharkiv, Ukraine

Dmitriy Klyushin Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Andrii Kondratiev O.M. Beketov National University of Urban Economy in Kharkiv, Kharkiv, Ukraine

Ihor Korniienko Chernihiv Polytechnic National University, Chernihiv, Ukraine

Svitlana Korniienko Chernihiv Polytechnic National University, Chernihiv, Ukraine

Hryhoriy Kosenyuk Bohdan Khmelnytsky National University of Cherkasy, Cherkasy, Ukraine

Yevgen Kotukh University of Customs and Finance, Dnipro, Ukraine

Kateryna Kovalets Institute of Mathematical Machine and System Problems, Kyiv, Ukraine

Kyrylo S. Krasnikov Dniprovskyi State Technical University, Kamianske, Ukraine

Bernd Kratz Castalytics GmbH, Kaufungen, Germany

Sergey Krivenko National Aerospace University, Kharkov, Ukraine

Anatoliy S. Kulik National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine

Svitlana Kunytska Cherkasy State Technological University, Cherkasy, Ukraine

Alexandr Kuznetsov V. N. Karazin, Kharkiv National University, Kharkiv, Ukraine;
JSC "Institute of Information Technologies", Kharkiv, Ukraine

Tetiana Kuznetsova V. N. Karazin, Kharkiv National University, Kharkiv, Ukraine

Maryna Kyrylenko   National Aerospace University «KhAI», Kharkiv, Ukraine

Dmytro Lande   Institute for Information Recording, National Academy of Sciences of Ukraine, Kyiv, Ukraine;
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Fangfang Li   National Aerospace University, Kharkov, Ukraine;
Nanchang Hangkong University, Nanchang, China

Anatoliy Lopushanskyi   Institute of Mathematical Machines and Systems Problems of the Ukraine National, Academy of Science, Kyiv, Ukraine

Vladimir Lukin   National Aerospace University, Kharkov, Ukraine

Sergey Lyashko   Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Svitlana Lytvyn   Chernihiv Polytechnic National University, Chernihiv, Ukraine

Vladimir Maderich   Institute of Mathematical Machine and System Problems, Kyiv, Ukraine

Serhii Morhun   Admiral Makarov National University of Shipbuilding, Mykolaiv, Ukraine

Tetyana Nabokina   National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine

Valentyn V. Nekhai   Chernihiv Polytechnic National University, Chernihiv, Ukraine

Tetiana Neskorodieva   Donetsk National University Named Vasyl Stus, Vinnytsia, Ukraine

Serhii Nesterenko   State Science and Research Institute of Armament and Military Technics Testing and Certification of the Ukraine Armed Forces, Chernihiv, Ukraine

Sergey N. Pasichnik   National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine

Anatolii Pavlenko   State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Chernihiv, Ukraine

Volodymyr Petrivskyi   Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Oleksii Pokotylo   National Defense University of Ukraine Named After Ivan Cherniakhovskyi, Kyiv, Ukraine

Svitlana Popereshnyak   Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Yurii Posternak   Chernihiv Polytechnic National University, Chernihiv, Ukraine

Dmytro Rahozin Software Systems Institute of NASU, Kyiv, Ukraine

Viktor Riabkov National Aerospace University «KhAI», Kharkiv, Ukraine

Andrii Rohovenko Chernihiv Polytechnic National University, Chernihiv, Ukraine

Inna Rozlomii Bohdan Khmelnytsky National University of Cherkasy, Cherkasy, Ukraine

Serhii Rudnichenko State Science and Research Institute of Armament and Military Technics Testing and Certification of the Ukraine Armed Forces, Chernihiv, Ukraine

Maxim Saveliev Institute for Safety Problems of Nuclear Power Plants of NAS of Ukraine, Chernobyl, Kyiv region, Ukraine;
Institute of Mathematical Machines and Systems Problems of NAS of Ukraine, Kyiv, Ukraine

Anton Sazonov Kyiv National University of Construction and Architecture, Kyiv, Ukraine

Viktor Shevchenko Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Volodymyr Shevchenko Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Igor Sinitsyn Institute of Software Systems, National Academy of Sciences of Ukraine, Kyiv, Ukraine

Igor Skiter Institute for Safety Problems of Nuclear Power Plants National Academy of Science of Ukraine, Kyiv, Ukraine

Dmytro V. Sokol National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine

Iryna Solomakha Chernihiv Polytechnic National University, Chernihiv, Ukraine

Nikolai Stoianov Defence Institute "Prof. Tsvetan Lazarov", Bul. "Prof. Tsvetan Lazarov" 2, Sofia, Bulgaria

Oleksandr Terentyev Department of Information Technologies of Design and Applied Mathematics, Kyiv National University of Construction and Architecture, Kyiv, Ukraine

Yuliia Tkach Chernihiv Polytechnic National University, Chernihiv, Ukraine

Elena Trunova Chernihiv Polytechnic National University, Chernihiv, Ukraine

Anton Tsaritsynskyi National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine

Ruslan Tsukanov National Aerospace University «KhAI», Kharkiv, Ukraine

Vitaliy Tsyganok Institute for Information Recording, National Academy of Sciences of Ukraine, Kyiv, Ukraine;
Faculty of Information Technology, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Olha Tsyhanok Department of Foreign Philology and Translation, Kyiv National University of Trade and Economics, Kyiv, Ukraine

Andrii Tymoshenko Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Vladyslav Vasylenko Institute of Telecommunications and Global Information Space of NAS of Ukraine, Kyiv, Ukraine

Anastasiya Vecherkovskaya Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Mariya Verovko AgileVision sp. z o.o., Krakow, Poland

Oleksandr Verovko AgileVision sp. z o.o., Krakow, Poland

Mariia Voitsekhovska Chernihiv Polytechnic National University, Chernihiv, Ukraine

Florian Wieduwilt Castalytics GmbH, Kaufungen, Germany;
Institut für Nanophotonik Göttingen e.V., Göttingen, Germany

Andrii Yarmilko Bohdan Khmelnytsky National University of Cherkasy, Cherkasy, Ukraine

Oleksii Yehorchenkov Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Nataliia Yehorchenkova Kyiv National University of Construction and Architecture, Kyiv, Ukraine

Sergei Zaitsev Chernihiv Polytechnic National University, Chernihiv, Ukraine

Alexandr Zaslavskiy Dnipro University of Technology, Dnipro, Ukraine

Kostiantyn Zavertailo Institute of Mathematical Machines and Systems Problems of the Ukraine National, Academy of Science, Kyiv, Ukraine

Dmytro Zhovtukhin Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

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

Mathematical Modeling of Information System Designing Master Plan of the Building Territory Based on OLAP Technology

Tetyana Honcharenko ⓘ**, Oleksandr Terentyev** ⓘ**, and Ievgenii Gorbatyuk** ⓘ

Abstract This study is devoted to the problem of the ability to represent multidimensional design of master plan of building territory based on OLAP technology. Building territory is viewed as complex information system which relate with other information systems of surrounding areas. Mathematical modeling of such information system for multidimensional design of master plan based on OLAP technology is proposed to solve this problem. It is proposed to use a multidimensional data model as a hypercube, the edges of which are sequences of values of the analyzed parameters for multidimensional analysis and design of the two subject areas. The first subject area contains background information about the topography, geodesy and geology of the building territory. The second subject area contains information about the surrounding areas to the building territory. OLAP cubes represent multidimensional information objects with associated properties. Each OLTP system is an implementation of a domain model and is designed by a multidimensional matrix, which is represented as an OLAP cube. The application of this concept in a Web-based environment is covered.

Keywords Mathematical modeling · Master plan · Building territory · OLAP

1 Introduction

Complex construction projects are becoming an important direction in the formation of sound economic decisions when assessing the possibilities of introducing innovative information technologies. Many types of products are complex systems, based on the interaction of a set of managerial and technical actions, including technical means, software and the human factor. Their production is carried out using

T. Honcharenko (✉) · O. Terentyev
Department of Information Technologies of Design and Applied Mathematics, Kyiv National University of Construction and Architecture, 31, Povitroflotsky Avenue, Kyiv 03037, Ukraine
e-mail: geocad@ukr.net

I. Gorbatyuk
Department of Construction Machinery, Kyiv National University of Construction and Architecture, 31, Povitroflotsky Avenue, Kyiv 03037, Ukraine

processes, having a variety of technical and control "inputs" and "outputs". At the same time, the role of standards used in at all stages of management, primarily because the standards provide for the interaction of various components with each other.

Author's works [1–3] describe modeling of spatial data on the construction site based on multidimensional information objects and multidimensional space structure for adaptable data model.

The articles [4–9] give a detailed review of innovative tools for management the lifecycle of construction objectives of the enterprise-stakeholder.

In the work [10], the authors propose multidimensional design of master plans of construction site based on multidimensional information objects.

The articles [4, 11–14] are devoted to the structural information management of production systems in construction and models for conversion of mortgage applications by the method of multiple regression and neural networks.

In the works [15–18], the authors provide the reengineering of the construction processes based on BIM-technology.

The authors of the works [19–21] formalize a concept of parametric modeling on BIM technology application in the whole life cycle construction objects.

The authors of [22–24] examine construction of membership functions in fuzzy modeling tasks using the analytic hierarchy process.

The articles [25] and [26] are devoted to a BIM-based approach for communicating and implementing a construction site safety plan.

Authors' work [27] is devoted to integrate processing of spatial information based on multidimensional data models for general planning tasks.

This research discusses the problem of the ability to represent multidimensional design of master plans based on OLAP Technology. Building territory is viewed as complex information system for multidimensional design of master plans. To solve this problem mathematical modeling of such information system based on OLAP technology is proposed.

2 The Main Research

The information system consists of objects that differ in the level of complexity, with the corresponding connections and relationships. Complexity is characterized by internal information, which quantitatively depends on the number of incoming objects and the connections of various levels established between them. The division of the system into information objects and functional modules and the description of all of its interaction interfaces make it possible to declare the relative completeness of the set of considered relations between the elements of the system. These elements determine its behavior and are the subject of functional stability analysis.

Building territory should be considered as a complex open system. For a system to operate and interact with the external environment, it must consume information from the environment and transmit information about its state to the environment
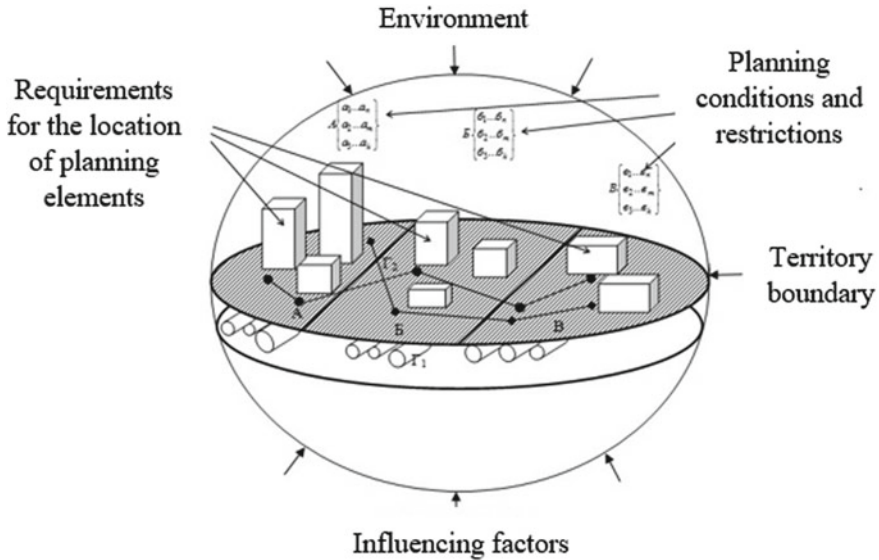
Fig. 1 Scheme of interaction of the building territory with the external environment

to increase its information potential. The territory of the building, depending on its geographical location, climate and other factors, requires additional research to ensure its safe construction and operation. The analysis of literature sources [4–7] allowed classifying the requirements for the layout of the elements of general planning and the main environmental factors that affect their location.

Figure 1 provides a scheme of the interaction of the building territory with the external environment.

The mathematical model of an information system can be represented as a set G, which contains information on each type of system object. On the other hand, the information system can be represented in the form of a functional system. That is in the form of a set of functions F. Let's introduce the notation:

Q_G is a set of properties determined by the relations between elements of the set G.

Q_F is a set of properties determined by the relationships between elements of the set F.

Q_{FG} is a set of properties determined by connections between elements of the sets F and G.

Then the relations within the sets F and G can be respectively determined by relations on the Cartesian product:

$$Q_F \times F = \left\{ \begin{array}{c} z_i^F = \left(q_i^F, f_i \right) : q_i^F \in Q_F, \\ f_i \in F, i = 1, \ldots, n \end{array} \right\} \tag{1}$$

and

$$Q_G \times G = \left\{ \begin{array}{l} z_i^G = \left(q_{i,}^G g_i \right) : g_i^G \in Q_G, \\ g_i \in G, i = 1, \dots, n \end{array} \right\} \tag{2}$$

The relation between them can be determined by the relation on the Cartesian product

$$Q_{FG} \times G \times F = \left\{ \begin{array}{l} z_i^{FG} = \left(q_i^{FG}, g_i, f_i \right) : q_i^{FG} \in Q_{FG}, \\ g_i \in G, f_i \in F, i = 1, \dots, n \end{array} \right\} \tag{3}$$

of the element z_i^{FG} to this relation can be interpreted as follows: "the element g_i of the information system contains information on the property q_i^{FG} of the functional part of the information system f_i". Searching for information corresponding to a particular element f_i in g_i is reduced to determining the ratio $R \subseteq G \times F$.

Thus, for any pair $(g_i, f_i) \in R$: $g_i \in G$, $f_i \in F$, $i = 1, \dots, n$, we can say that f_i is relevant to g_i, and the solution to the problem of determining the relevance of elements of the sets G and F is reduced to the definition of the relation $R \subseteq G \times F$. Moreover, $\forall g_i \in G$, $f_i \in F$, $g_i \in G$, $f_i \in F$, $i, j = 1, \dots, n$, it is true that if $f_i \subseteq f_j$ and $g_i \subseteq g_i$, that is, all elements g_i are contained in g_i and all elements f_i are contained in f_i and $(g_i, f_i) \in R$, then $(g_i, f_i) \in R$.

Except in the extreme case, when the relation R is the Cartesian product $G \times F$ itself, the relation does not include all possible tuples from the Cartesian product. This means that for each relation there is a criterion for determining which tuples are included in the relation and which are not. Thus, each relation R can be associated with a logical expression (predicate) Q_{FG}, which depends on a certain number of parameters (n-ary predicate) and determines whether the tuple (g_i, f_j) belongs to the relation R.

Thus, the membership of a tuple in the relation is equivalent to the truth of the predicate:

$$\left(g_j, f_j \right) \in R \Leftrightarrow \{ Q_{FG} \} = \{ G, F, R \} \tag{4}$$

However, in any case, with the information approach to formalize the subject area, the categories of objects and the relations between them are primary, i.e., formally, the system of relations can be represented by a set of objects in the subject area and a variety of relations between them.

The characteristic of the process can be represented as a set of pairs:

$$\{ \langle A_i, D_i \rangle, i = 1, \dots n \}, \tag{5}$$

where A_i is a non-empty set of property names (attributes), D_i is a set of values of the corresponding attributes.

Values are broken down into object classes that interact with each other based on rules. Let π be the set of these rules. Relationships $G = \{\overline{G}, \tilde{G}\}$ can be established on a set of attributes, which are divided into quantitative $\overline{G}$ and qualitative $\tilde{G}$, for which many types of assessment are defined, for example $T = \{$"projects are moving towards achieving their goals", "projects are being conducted in accordance with the relevant directives", "projects are being implemented as planned", "projects remain viable"$\}$. Then any estimation rule can be represented by a tuple $\pi = \langle G, T \rangle$.

Thus, the set of information characteristics of the process $\{\langle A_i, D_i \rangle, i = 1, n\}$, the established relations $G = \{\overline{G}, \tilde{G}\}$, and the rules for establishing relations $\pi = \langle G, T \rangle$ can be used to formally define the process in the form the following tuple of components:

$$Z = \left\{ A_i, D_i, \left\{\overline{G}, \tilde{G}\right\}, \left\{\overline{G}, \tilde{G}\right\}, T \right\}, i \in N \tag{6}$$

Attribute values may not be numeric. In particular, the linguistic form of data presentation is widely used in macroeconomic, sociological, marketing, medical, legal data warehouses. To assess the characteristics that are of a qualitative nature, ordinal scales can be used, the point elements of which correspond to the gradations of the verbal scales.

We can assign the values of linguistic variables to the levels of ordinal scales and perform all further operations with their membership functions. At the same time, their adequacy cannot be verified by the means of theory, and each existing method for constructing a membership function formulates its own requirements and justifications for the choice of just such a construction.

Consider N objects for which the intensity of manifestation of characteristics of attributes with the names $A_j, j = 1, \ldots, k$ is estimated, the values of which $X_j, j = 1, \ldots, k$ are used to assess the qualitative characteristic Y.

From the standpoint of the apparatus of the theory of fuzzy sets, the models of expert evaluation of features are full orthogonal semantic spaces, where normal triangular numbers and T-numbers are used as membership functions.

This method allows going from diverse qualitative information to a single abstract value of the membership function. We use the method of working with fuzzy information, considered in [9].

Let $X_{ij}, i = 1, \ldots, m_j$ are levels of verbal scales used to assess attributes with the names $A_j, j = 1, \ldots, k$ and arranged in ascending order of intensity of their manifestation.

Let's denote by $a_i^j, i = 1, \ldots, m_j, j = 1, \ldots, k$ − the relative numbers of objects referred during the assessment of attributes with the names $A_j, j = 1, \ldots, k$ to the level $X_{ij}, i = 1, \ldots, m_j, j = 1, \ldots, k$:

$$\sum_{i=1}^{m} a_i^j = 1, j = 1, \ldots, k. \tag{7}$$

We can construct k sets of fuzzy numbers corresponding to attributes named A_j, $j = 1, k$.

Let $\mu_{ij}(x)$ denote the membership function of the fuzzy number $\tilde{X}_{ij}$ corresponding to the i-th level of the j-th attribute $i = 1,\ldots, m_j, j = 1,\ldots, k$.

Fuzzy numbers can be taken as an estimate of the object, $\tilde{X}_{ij}$, $i = 1,\ldots, m_j, j = 1,\ldots, k$ with the corresponding membership functions $\mu_{ij}(x)$, $i = 1,\ldots, m_j, j = 1,\ldots, k$.

Then the estimate of the n-th object by the attribute X_j, $j = 1, \ldots, k$ can be represented by the membership function μ_j^n.

Using the so-called L-R membership function, the assessment of each object can be written in the form:

$$\{\mu_j^n(x) = (a_{j1}^n, a_{j2}^n, a_{jL}^n, a_{jR}^n)\}, n = 1, \ldots, N, j = 1, \ldots, k \tag{8}$$

In the general case, the characteristic of each object X_i can be described by the corresponding linguistic variable:

$$X_i = \langle A_j, T_j, D_j \rangle \tag{9}$$

where $T_j = \left\{ T_1^j, T_2^j, \ldots, T_{m_j}^j \right\}$ is the term set of the linguistic variable, A_j is a set of linguistic values of the attribute, m_j is number of attribute values; D_j is subject scale base set of the attribute A_j.

To describe the terms T_k^j, $k = 1,\ldots, m_j$ corresponding to the values of the attribute A_j, fuzzy variables T_k^j, D_j, $\tilde{C}_k^j$ can be used. The value T_k^j is described by the fuzzy set $\tilde{C}_k^j$ base on set D_j:

$$\tilde{c}_k^j = \left\{ \left\langle \mu_{c_k^j}(d)|d \right\rangle \right\}, d \in D_j, k = 1, \ldots, m_j \tag{10}$$

Then the fuzzy set of the second level can be taken as a fuzzy characteristic of the object x_i, that is described as follows:

$$\tilde{x}_i = \left\{ \mu_{x_i}(a_j)|a_j \right\},$$
$$\mu_{x_i}(a_j) = \bigcup_{k=1}^{m_j} \left\{ \left\langle \mu_{\mu_{x_i}}\left(T_k^j\right) \middle| T_k^j \right\rangle \right\}, T_k^j \in T_j, a_j \in A_i \tag{11}$$

All this makes it possible to move from diverse qualitative information to one abstract quantity—the value of the membership function. To assess the qualitative characteristics of objects, we use the following representation of their elements:

the membership function for the extreme term-set corresponding to the minimum intensity of the feature manifestation can be represented as:

$$\mu_{x_1}(x) = \begin{cases} 1, & 0 \le x \le a_1 - \frac{1}{2}\min_a(a_1, a_2) \\[2mm] 1 - \dfrac{x - \left(a_1 - \frac{1}{2}\min\limits_a(a_1, a_2)\right)}{\min\limits_a(a_1, a_2)}, \\[3mm] & a_1 - \frac{1}{2}\min_a(a_1, a_2) < x \le a_1 + \frac{1}{2}\min_a(a_1, a_2) \\[2mm] 0, & a_1 + \frac{1}{2}\min_a(a_1, a_2) < x \le 1 \end{cases} \tag{12}$$

the membership function for the extreme term-set corresponding to the maximum intensity of the manifestation of the feature can be represented as:

$$\mu_{x_m}(x) = \begin{cases} 0, & 0 \le x \le 1 - a_m - \frac{1}{2}\min_a(a_{m-1}, a_m) \\[2mm] 1 + \dfrac{x - \left(1 - a_m - \frac{1}{2}\min\limits_a(a_{m-1}, a_m)\right)}{\min\limits_a(a_{m-1}, a_m)} \\[3mm] & 1 - a_m - \frac{1}{2}\min_a(a_{m-1}, a_m) < x \le 1 - a_m + \frac{1}{2}\min_a(a_{m-1}, a_m) \\[2mm] 1, & 1 - a_m + \frac{1}{2}\min_a(a_{m-1}, a_m) < x \le 1 \end{cases}$$

$$\tag{13}$$

the membership function for average term sets of a qualitative feature can be represented as:

$$\mu_{x_l}(x) = \begin{cases} 0, \ 0 \le x \le \sum_{i=1}^{l-1} a_i - \frac{1}{2} \min_a(a_{l-2}, a_{l-1}, a_l) \\[2mm] 1 + \dfrac{x - \left(\sum_{i=1}^{l-1} a_i - \frac{1}{2} \min_a(a_{l-2}, a_{l-1}, a_l)\right)}{\min(a_{l-2}, a_{l-1}, a_l)}, \\[2mm] \quad \sum_{i=1}^{l-1} a_i - \frac{1}{2} \min_a(a_{l-2}, a_{l-1}, a_l) < x \le \sum_{i=1}^{l-1} a_i \\[2mm] \quad + \frac{1}{2} \min_a(a_{l-2}, a_{l-1}, a_l) \\[2mm] 1, \ \sum_{i=1}^{l-1} a_i - \frac{1}{2} \min(a_{l-2}, a_{l-1}, a_l) < x \le \sum_{i=1}^{l} a_i \\[2mm] \quad + \frac{1}{2} \min_a(a_{l-1}, a_l, a_{l+1}) \\[2mm] 1 + \dfrac{x - \left(\sum_{i=1}^{l} a_i - \frac{1}{2} \min_a(a_{l-1}, a_l, a_{l+1})\right)}{\min_a(a_{l-1}, a_l, a_{l+1})}, \\[2mm] \quad \sum_{i=1}^{l} a_i - \frac{1}{2} \min_a(a_{l-1}, a_l, a_{l+1}) < x \le \sum_{i=1}^{l} a_i \\[2mm] \quad + \frac{1}{2} \min_a(a_{l-1}, a_l, a_{l+1}) \\[2mm] 0, \ \sum_{i=1}^{l} a_i + \frac{1}{2} \min_a(a_{l-1}, a_l, a_{l+1}) < x \le 1 \end{cases}$$

$\mu_{x_l}(x)$ Based on the above, the subject area can be represented in the form of a multi-level environment consisting of many elements of the subject area, a variety of functions and methods working on these elements and a variety of properties of elements and relationships between elements, i.e., in the form of an ontology that includes description of the properties of the domain and the interaction of objects in some formal language that has logical semantics. If the system is complex and the number of factors is large, then taking into account all its characteristics (components) leads to extreme complexity. Therefore, only a limited number has to be entered into the model, and the remaining components must be taken into account without explicitly entering into the model, but taking into account their influence as a fuzzy reaction of the model to one or another choice of an alternative. Obviously, algebraic comparison of components is impossible and can be performed using fuzzy logic methods.

3 Results and Discussion

In practice, this concept finds its application in a Web-based environment, which requires high performance when processing large amounts of data, for example, a
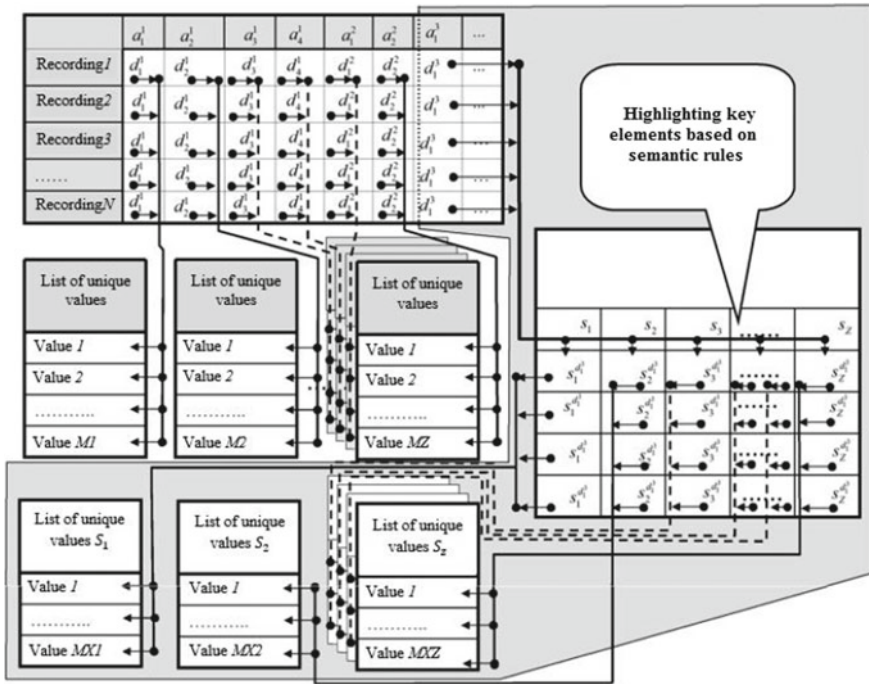
Fig. 2 OLAP cube formation model

Web portal, which, in turn, is a set of database information transmission channels connected into a single system data, information storages, knowledge bases, as well as information technologies that support the processing, analysis and transfer of information at various levels of integration.

Figure 2 shows OLAP cube formation model. At the same time, when solving problems of one system, knowledge of other systems is used. Similarly, the same objects in different systems can be described by different properties and, accordingly, have a different structure. One of the challenges is determining how to integrate the knowledge of different topics within a single Web portal.

The main operation for a larger number of users is the operation of searching and obtaining information. Moreover, the data itself, once formed, is no longer subject to modification. Dynamic databases in such a situation are ineffective, and in the world practice today, to solve such problems, a transition to information storages is carried out. Let D_j be a collection of documents, M_j a set of queries:

$$d_k : M_j \rightarrow 2^{D_j}$$

where d_k is a mapping that associates a set of documents with each query.

Each information subsystem of the portal can be represented by a tuple:

$$S_j = \left(D_j, M_j, d_j\right),$$

where $j = 1, \ldots, m$ are information subsystems used in the Web-portal.

Then the information system of the portal can be represented in the form of a distributed information system based on the global thesaurus T and defined by the tuple:

$$S = (T, D, M, d),$$

where $D = \bigcup_j D_j$ and $d = \bigcup_j d_j$ are the local components.

Any database can be viewed as an object archive designed to store attributes:

(1)   passive objects that do not have state models
(2)   active objects subject to the life cycle.

Figure 3 shows the process of creating new OLAP cube based on the two subject areas relations. The first subject area contains background information about the topography, geodesy and geology of the building territory. The second subject area contains information about the surrounding areas to the building territory.

There is a multi-tempo change in attributes. For example, the identity of a student does not change, and a number of specific attributes, such as faculty, group, current performance, are dynamic and modified. At the end of training, the data becomes permanent and can never be changed.
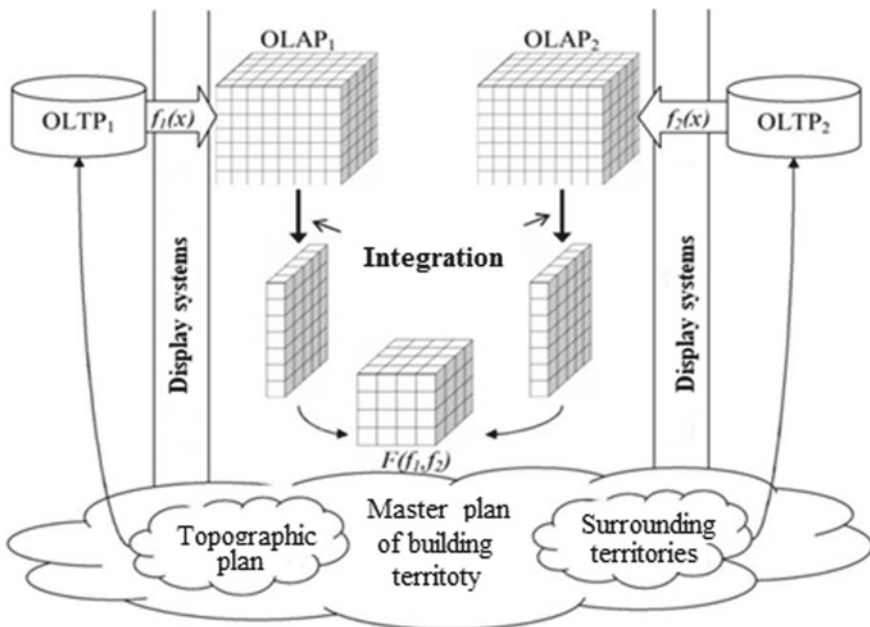
Fig. 3 The process of designing master plan of building territory based on OLAP technology

Thus, it is advisable to divide all the considered information into different levels, both in terms of lifetime and access, and the development of the database consists in adding and removing tuples corresponding to object instances. In this case, the database automatically goes into the category of information storages completely when the following conditions are met: the frequency of deleting tuples from the storage is commensurate with the time of its existence, and the OLAP cube is supplemented with a new dimension— fragmented time. This is the moment of transferring information from the database to the storage, in accordance with the changed value of the attributes.

If the information in dynamic databases after modification is not of interest from any of the points of view, then the organization of an information storage based on such databases does not make sense. The OLAP cubes themselves, corresponding to OLTP systems, can be represented as multidimensional information objects with corresponding properties.

Since each OLTP system, being an implementation of a domain model, is mapped to a multidimensional matrix (OLAP cube), there are functions for this mapping. Considering the subject areas as part of the extended subject area, we can conclude that it is possible to establish a relation function between display functions, and, as a consequence, build a new system or OLAP cube based on this relation. It should be noted that the efficiency of the Web portal can be significantly improved if the features of the measurement scales are more accurately taken into account in modeling.

4 Conclusions

This study is devoted to the problem of the ability to represent design of master plan of building territory based on OLAP technology. Building territory is viewed as complex information system which relate with other information systems of surrounding areas. Mathematical modeling of such information system for multidimensional design of master plan based on OLAP technology is proposed to solve this problem.

The digital model covering the information system of building territory is presented in the form of a metabase, which contains information about each type of spatial objects. A mathematical description of the model of the master plan of building territory from the standpoint of OLAP-tools is given. It is proposed to use a multidimensional data model or a hypercube, the edges of which are sequences of values of the analyzed parameters for multidimensional analysis of the two subject area. The first subject area contains background information about the topography, geodesy and geology of the building territory. The second subject area contains information about the surrounding areas to the building territory. Each OLTP system is an implementation of the subject area model and is mapped to a multidimensional matrix as OLAP-cube. The application of this concept in a Web-based environment is covered.

References

1. Mihaylenko V, Honcharenko T, Chupryna K, Andrashko Yu, Budnik S (2019) Modeling of spatial data on the construction site based on multidimensional information objects. Int J Eng Adv Technol 8(6):3934–3940. https://www.ijeat.org/wp-content/uploads/papers/v8i6/F9057088619.pdf
2. Sacks R, Eastman C, Lee G, Teicholz P (2018) BIM handbook: a guide to building information modeling for owners, designers, engineers, contractors, and facility managers, 3rd edn. Wiley, Hoboken, NJ, USA, p 688
3. Pepe M, Costantino D, Restuccia Garofalo A (2020) An Efficient pipeline to obtain 3D model for HBIM and structural analysis purposes from 3D point clouds. Appl Sci 10:12–35
4. Riabchun Y, Honcharenko T, Honta V, Chupryna K, Fedusenko O (2019) Methods and means of evaluation and development for prospective students' spatial awareness. Int J Innovative Technol Exploring Eng 8(11):4050–4058
5. Bushuyev S, Kozyr B, Rusan N (2019) Modeling of empathy, emotional intelligence and transformational leadership to the project success. Mathematical modeling and simulation of systems. In: Selected Papers of 14th international scientific-practical conference, MODS, June 24–26, Chernihiv, Ukraine, vol 1019. Springer Nature Switzerland AG, pp 209–223
6. Trach R, Bushuyev S (2020) Analysis of communication network of the construction project participants. Sci Rev Eng Environ Sci 29(3):388–396
7. Ryzhakov D, Dikiy O, Druzhynin M, Petrenko H, Savchuk T (2020) Innovative tools for management the lifecycle of strategic objectives of the enterprise-stakeholder in construction. Int J Emerg Trends Eng Res 8(8):4526–4532. https://doi.org/10.30534/ijeter/2020/78882020
8. Voitushenko A, Bushuyev S (2020) Development of project managers' creative potential: determination of components and results of research. In: Advances in intelligent systems and computing, 1080 AISC, pp 283–292
9. Chernyshev D, Ryzhakov D, Dikiy O, Khomenko O, Petrukha S (2020) Innovative technology for management tools of commercial real estate in construction. Int J Emerg Trends Eng Res 8(9):4967–4973
10. Honcharenko T, Chupryna Y, Ivakhnenko I, Zinchenco M, Tsyfra T (2020) Reengineering of the construction companies based on BIM-technology. Int J Emerg Trends Eng Res 9(5):8670–8676
11. De Santana SA (2013) Modeling urban landscape: new paradigms and challenges in territorial representation. Disegnare Con 6(11):161–174
12. Dyomin M, Dmytrenko A, Chernyshev D, Ivashko O (2020) Big cities industrial territories revitalization problems and ways of their solution. Lect Notes Civil Eng 73:365–373
13. Biancardo SA, Viscione N, Cerbone A, Dessì E (2020) BIM-based design for road infrastructure: a critical focus on modeling guardrails and retaining walls. Infrastructures 5:59
14. Cheng JCP, Lu Q, Deng Y (2016) Analytical review and evaluation of civil information modeling. Autom Constr 67:31–47
15. Azhar S, Behringer A (2013) A BIM-based approach for communicating and implementing a construction site safety plan. In: 49th ASC annual international conference proceedings. Associated Schools of Construction. https://ascpro0.ascweb.org/archives/cd/2013/paper/CPRT43002013.pdf
16. Tan Y, Fang Y, Zhou T, Gan VJL, Cheng JCP (2019) BIM-supported 4D acoustics simulation approach to mitigating noise impact on maintenance workers on offshore oil and gas platforms. Autom Constr 100:1–10. https://doi.org/10.1016/j.autcon.2018.12.019
17. Shkuro M, Bushuyev S (2019) Development of proactive method of communications for projects of ensuring the energy efficiency of municipal infrastructure. Eureka Phys Eng 1:3–12. https://doi.org/10.21303/2461-4262.2019.00826
18. Kolbe TH, Lee J, Zlatanova S (2009) Representing and exchanging 3D city models with CityGML. In: 3D geo-information sciences. Springer, Berlin, Heidelberg, pp 15–31
19. Koshak N, Flemming U (2002) Object-oriented data modeling and warehousing to support urban design. In: Proceedings of DDSS 2002, Ellecom, The Netherlands, p 18

20. Montenegro N, Duarte JP (2009) Computational ontology of urban design: towards a city information model. In: Computation: the new realm on architectural design. Istanbul Technical University, Turkey, pp 253–260
21. Pandit P (2009) Using postGIS/postgreSQL for managing CAD and GIS data, autodesk
22. Stadler A, Nagel C, König G, Kolbe TH (2009) Making interoperability persistent: a 3D geo database based on cityGML. In: 3D geo-information sciences. Springer, Berlin Heidelberg, pp 175–192
23. Duarte JP, Montenegro N, Beirão JN, Gil J (2011) City induction: a model for formulating, generating, and evaluating urban designs. In: Digital urban modelling and simulation. Communications in computer and information science series. Springer, Berlin
24. Gil J, Beirão JN, Montenegro N, Duarte JP (2010) Assessing computational tools for urban design: towards a city information model. In: Future Cities, Proceedings of the eCAADe conference. ETH Zurich, Switzerland, pp 361–369
25. Gil J, Duarte JP (2010) Towards an urban design evaluation framework. In: Architecture in Computro, Proceedings of the eCAADe conference, Antwerpen, Belgium, pp 257–264
26. Gil J, Duarte JP (2010) A review of urban design sustainability evaluation tools. In: HJP Timmermans and B de Vries 10th international conference on design & decision support systems in architecture and urban planning. Eindhoven University of Technology
27. Mihaylenko V, Honcharenko T, Chupryna K, Liazschenko T (2021) Integrated processing of spatial information based on multidimensional data models for general planning tasks. Int J Comput 20(1):55–62. https://doi.org/10.47839/ijc.20.1.2092

Models and Information Technologies of Coverage of the Territory by Sensors with Energy Consumption Optimization

Volodymyr Petrivskyi ⬤**, Viktor Shevchenko** ⬤**, Oleksii Bychkov** ⬤**, and Oleksii Pokotylo** ⬤

Abstract According to great number of advantages sensors and sensor networks become widespread and take a great part in all human activities. One of the key problems is to reduce sensor networks' energy consumption while achieving maximum territory coverage. In the article a mathematical model that allows to choose the optimal combination of sensors of different types for territory coverage with minimum total energy consumption based on knapsack problem is presented. The type of sensor means different coverage radius and as a consequence different energy consumption value. For minimizing energy consumption with maximizing area coverage in case of coverage areas intersection existing, a mathematical model in multicriteria optimization problem form is formulated. A modified algorithm for constructing the optimal motion trajectory using the Bellman principle in the case of motion of several sensors is proposed. The possibility of the existence of an already covered area which may be an obstacle to the movement of sensors is considered. Also, the example of appropriate software architecture is presented. Computer simulating results confirms the effectiveness of the proposed models and approaches.

Keywords Sensors · Sensor networks · Territory covering · Energy efficiency · Optimization · Mathematical modeling

1 Introduction

Nowadays, sensors have gained significant application in all areas of human use due to a number of advantages: size, autonomy, accessibility, mobility [1, 2]. Sensors are also used to perform a number of specific tasks related to the collection and processing of information [3], among which the common task is monitoring. This task means optimal territory coverage by a given number of sensors with appropriate

V. Petrivskyi (✉) · V. Shevchenko · O. Bychkov
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

O. Pokotylo
National Defense University of Ukraine Named After Ivan Cherniakhovskyi, Kyiv, Ukraine

characteristics finding [4]. In the article [5] the consideration of ecological moni-
toring technology is presented, as well as the solution of the problem of placing
sensors in a polygonal observation zone with the presence of obstacles. In papers
[6, 7] algorithms and methods for solving the problem of monitoring and optimal
location of sensors are described. The algorithm of optimal location of sensors for
determining structural damage of different types of equipment is presented in [8].
Authors of works [9, 10] present the optimal strategy for placing sensors to monitor
the environment using wireless sensor networks. Optimization of sensor placement
using gradient descent and probabilistic coverage is presented in paper [11]. In turn,
in manuscript [12] the approach of reduction of energy consumption of the sensor
network by regulation of sensor coverage areas is presented. It is also necessary to
take into account the possibility of intersection of the sensor coverage areas. Sensors'
movability should also be taken into account [13]. In turn moveable sensors can be
uncontrolled and, as a result, such movement will be arbitrarily or be controlled and
move a certain trajectory. Controlled motion of the sensor causes optimal trajec-
tory the best route choosing problems. In the manuscript [14] optimal trajectory is
defined as a solution of the problem of choosing the optimal route using methods of
partially integer linear programming. The optimal trajectory for avoiding dangerous
zones is determined by the Pontryagin maximum principle [15, 16]. In the article
[17] information technology of one sensor's optimal trajectory definition by using
dynamically programming approach is presented.

2   Area Covering by Stationary Sensors Mathematical Models

Suppose we have territory that must be covered with sensors. We consider territory
in the two-dimensional line space. It obviously that given area has square, let mark
it T_s. According to the shapes of the territory the value of S can be calculated in
different ways. Let consider three cases of the territory's square S calculation:

1.   Rectangle shaped territory:

$$T_s = ab, \tag{1}$$

where a and b—length and width of the territory.

2.   Any-shaped territory T and the equation of the curve that limits the area is
known:

$$T_s = \iint\limits_{T} f(x, y)dxdy, \tag{2}$$

where $f(x, y)$—equation of the curve that limits the territory.

3. Any-shaped territory T and the equation of the curve that limits the area is unknown. Pick's theorem use [18]:

$$T_s = P_i + \frac{P_b}{2} - 1,$$ (3)

where P_i—number of interior points, P_b—number of boundary points.

For the Pick's theorem use we build the grid of equal-distanced points on the territory T. To make square approximation with Pick's theorem more accurate the distance between points of the net should be reduced.

Also, any-shaped territory can be approximated with a rectangle. Examples of territory approximation and Pick's theorem use presented in Fig. 1.

Suppose we have a set of sensors $S = \{s_1, s_2, \ldots, s_n\}$, n—number of sensors. Each sensor has following parameters:

$$s_i = s_i(x_i, y_i, r_i, Bc_i),$$ (4)

where x_i, y_i—sensor's coordinates, r_i—sensor's coverage radius, Bc_i—volume of the sensor's battery.

Each sensor has own battery consumption. According to [19] we will calculate sensor's energy consumption, mark it as E_i, per unit of time in the next way:

$$E_i = r_i^2,$$ (5)

where r_i—i-sensors coverage radius, $i = \overline{1, n}$, n—number of sensors.
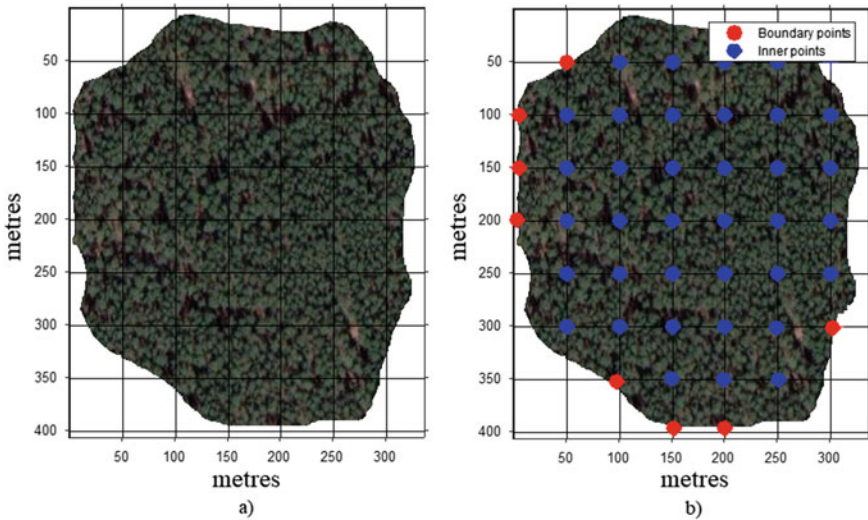
Fig. 1 a Territory approximation by rectangle and b Pick's theorem use examples

The problem is to cover given territory by sensors with minimum energy consumption. For solving described problem let modify well known knapsack optimization problem. Classical knapsack problem has next form [20]:

$$\sum_{i=1}^{n} x_i p_i \rightarrow \max,$$ (6)

$$\sum_{i=1}^{n} x_i w_i \leq W,$$ (7)

$$x_i \in \{0, 1\} \text{ or } x_i \in \{0, 1, 2, \ldots, m_i\},$$ (8)

where p_i—price of the thing, w_i—weight of the thing, m_i—maximum i-thing count.

According to (1)–(3) and (5) problem (6)–(8) can be modified in the next form:

$$\sum_{i=1}^{n} x_i r_i^2 \rightarrow \min,$$ (9)

$$\sum_{i=1}^{n} x_i \pi r_i^2 \geq T_s,$$ (10)

$$x_i \in \{0, 1\} \text{ or } x_i \in \{0, 1, 2, \ldots, m_i\},$$ (11)

where r_i—i-sensors coverage radius, m_i—available count of i-sensors.

In case when constraints (8) and (11) has $x_i \in \{0, 1\}$ form we get 0–1 optimization problem. In other case we get bounded optimization problem.

The solution of the optimization problem (9)–(11) is vector $x = (x_1, x_2, \ldots, x_n)$ elements of which are count of sensors each type that should be taken for territory coverage.

Let consider case when it is necessary for sensors to exchange information. For information exchange process a necessary condition is the intersection of coverage areas. Let mark minimum available coverage areas intersection level when information exchange exists as c and schematically be described as in Fig. 2.

Sensor's intersection should be taken into account while solving the problem of minimization sensors' energy consumption during maximizing territory coverage. In general minimization sensors' energy consumption during maximizing territory coverage problem can be presented in the next form:

$$\begin{cases} E(r, c) \rightarrow \min \\ Z(r, c) \rightarrow \max \end{cases},$$ (12)

where $E(r, c)$—energy consumption, $Z(r, c)$—covered territory, r—sensor's coverage radius, c –minimum available intersection level.
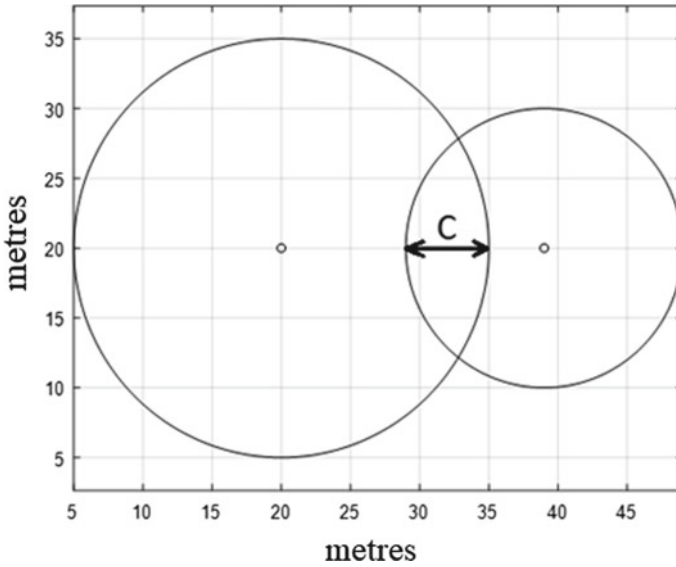
Fig. 2 Value c meaning

With taken into account sensors intersection value $Z(r, c)$ in the Eq. (12) should be presented as the difference between covered value and intersected value:

$$Z(r, c) = Z_{covered}(r, c) - Z_{intersected}(r, c). \tag{13}$$

Given the assumption that the coverage area of the sensor is a circle centered at a certain point (x_i, y_i) and radius r_i value of covered area as follows:

$$Z_{covered} = \sum_{i=1}^{N} \pi r_i^2. \tag{14}$$

In the Eq. (14) value N—number of sensors and can be calculated in the next way:

$$N = \left[\frac{T_s}{\pi r^2}\right], \tag{15}$$

where T_s—square of territory (1)–(3).

Value of intersected area according to [21] can be presented in the next form:

$$Z_{intersected}(r, c) = m\frac{r^2}{2}(K(r, c) - \sin(K(r, c))), \tag{16}$$

where $K(r, c) = 2 \arccos\left(\frac{r - \frac{c}{2}}{r}\right)$, m—number of intersections.

Summary problem (12)–(16) is nonlinear multicriterial optimization problem. Let change area coverage equation in the next way:

$$Z(r, c) = T_s - Z_{covered}(r, c) + Z_{intersected}(r, c), \tag{17}$$

where T_s—square of territory (1)–(3).

According to the new form of total coverage equation problem (12) will have next form:

$$\begin{cases} E(r, c) \to \min \\ Z(r, c) \to \min \end{cases}. \tag{18}$$

By using parameter convolution method present minimization sensors' energy consumption during maximizing territory coverage problem in the next way:

$$F(r, c) = \alpha_1 Z(r, c) + \alpha_2 E(r, c) \to \min, \tag{19}$$

$$0 \le r \le r^{\max}, \tag{20}$$

$$c_{\min} \le c \le r^{\max}, \tag{21}$$

$$0 \le \alpha_1 \le 1, \alpha_2 = 1 - \alpha_1, \tag{22}$$

where r—coverage radius, c—intersection level, $c_{\min}$—minimum available intersection value, α_1 and α_2—expert estimates of parameters.

The solution of the (19)–(22) problem will be pair of the optimal coverage radius r and intersection level c with using of which minimization energy consumption with coverage maximization is achieved.

3 Area Covering by Dynamical Sensors Information Technology

Suppose there are territory for covering with already covered part and a set of moveable sensors. Each sensor has next parameters:

$$s_i(t) = s_i(x_i(t), y_i(t), v_i(t), r_i(t), Bc_i), \tag{23}$$

where $x_i(t)$, $y_i(t)_i$—sensor's position at the time moment t, $v_i(t)$, $r_i(t)$—movement speed and coverage radius at the time moment t.

In the paper [17] information technology for territory covering with using dynamic programming approach and Bellman's principle is presented. But presented technology is only for one sensor case. According to the [17] we take movement equation in the differential equation form and get system of differential equations:

$$\begin{cases} \frac{dx_1}{dt} = v_1(t)\cos(\alpha_1(t)) \\ \frac{dy_1}{dt} = v_1(t)\sin(\alpha_1(t)) \\ \qquad\vdots \\ \frac{dx_n}{dt} = v_n(t)\cos(\alpha_n(t)) \\ \frac{dy_n}{dt} = v_n(t)\sin(\alpha_n(t)) \end{cases} \tag{24}$$

Principle of solving subproblems which are parts of the main problem is used [22], in our case the main problem is to find the optimal trajectory with maximum coverage, subproblems are to select the direction of motion of the sensor at each time moment t. Equation (24) by the finite difference approximation can be presented in the next form:

$$\begin{cases} x_{i+1}^1 = x_i^1 + v_{i+1}^1 \cos\alpha_{i+1}^1 \\ y_{i+1}^1 = y_i^1 + v_{i+1}^1 \sin\alpha_{i+1}^1 \\ \qquad\vdots \\ x_{i+1}^n = x_i^n + v_{i+1}^n \cos\alpha_{i+1}^n \\ y_{i+1}^n = y_i^n + v_{i+1}^n \sin\alpha_{i+1}^n \end{cases} \tag{25}$$

Let introduce "unitary motion" as motion per unit of time Δt on the unitary step value Δ. Also, make next assumptions:

1. Each sensors' speed is constant and denote $v_i = 1$;
2. Each sensor can move in 8 directions (Fig. 3) on the unitary step size Δ.

Diagonal motion was introduced to reduce energy consumption by avoiding two movements.

For the defining unitary movement direction let calculate the estimation for each sensors' endpoint described above and denote it c_i^k, $i = \overline{1,8}$, $k = \overline{1,n}$, n—count of sensors. This estimation will be equal:

$$c_i^k = A_{io}^k + A_{iz}^k + A_{is}^k, \tag{26}$$

$$A_{io}^k = \frac{1}{2}r_k^2(\theta - \sin\theta), \theta = 2\arccos\left(\frac{r_k - h}{r_k}\right), \tag{27}$$

where A_{oi}^k—area coverage outside the zone, A_{zi}^k—intersection value with already covered zone, A_{is}^k—intersection value with other sensors (Fig. 4), $i = \overline{1,8}$, $k = \overline{1,n}$, n—count of sensors.

Fig. 3 Available sensors'
movement directions
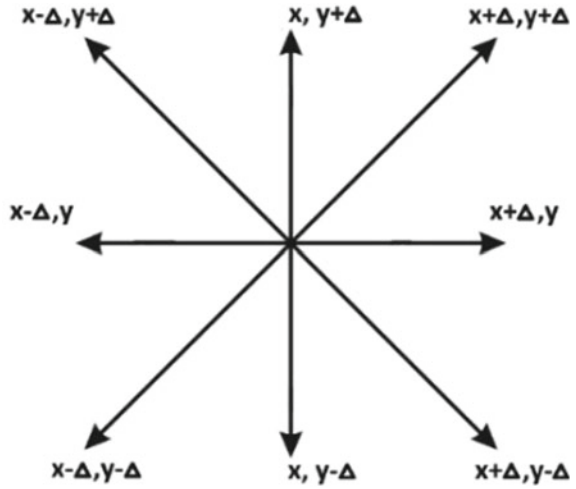
Fig. 4 Coverage area
intersection zones
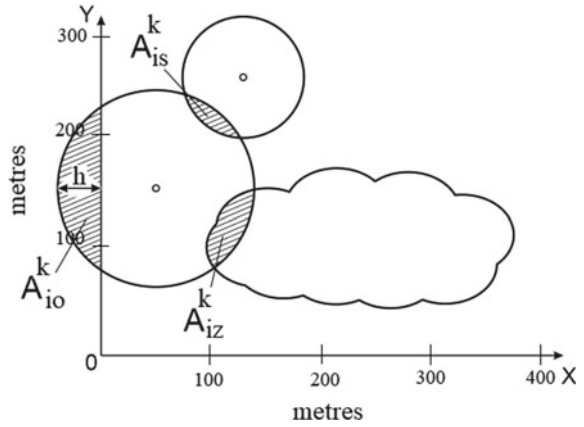
Value of the A_{iz}^k and A_{is}^k can be calculating by using formulas (1)–(3) or approach presented in the paper [23]. Schematically coverage area intersection zone can be described on the following figure:

For each sensor among all estimations $c_i^k, i = \overline{1, 8}, k = \overline{1, n}, n$—count of sensors, choose the minimum one, which index will be each sensors' direction for the next step:

$$d_k = \min_i c_i^k. \tag{28}$$

Let introduce for each sensor priority direction pw_1^k and second priority direction pw_2^k in order to avoid possible cases of same estimations existing. In such cases choose direction that is equal to pw_1^k or pw_2^k, or the nearest to them that has minimum

Table 1 Request-response example

Request	Response

{
"id": 1;
"x": 54.8;
"y": 36;
"pw1": 2;
"pw2": 8;
"pd": 5;
}
	{
	"id": 1;
	"direction": 2;
	"stop": false;
	}

estimation value. also add the prohibition movement in the opposite direction pd_k:

$$d_k \neq pd_k. \tag{29}$$

The algorithm terminates when all direction estimations are greater than the maximum permissible value of intersected area $C_{\max}$:

$$\min c_i^k \geq C_{\max}. \tag{30}$$

Appropriate software has client–server architecture. Every time period $t_i \, mod \, period = 0$ sensors send requests to the server that consists of sensor's id, coordinates, priority and previous directions. The server send response that has sensor's id, movement direction and stop flag. All messages have json format. In case when there is no request or response during limited time the sensor remains in the same position. The request-response example presented in Table 1.

4 Computer Simulation Results

For the first computer experiment suppose we have territory for covering with square value $T_s = 126080$ m^2 and sensors with coverage radiuses 50, 25 and 20 m. With using of approach (9)–(11) optimal territory achieves with the set of sensors: 3 sensors with coverage radius 50, 24 sensors with coverage radius 25 and 29 sensors with coverage radius 20. Results presented on Figs. 5 and 6.

For the next computer test we have a territory with square value $T_s = 117983$ m^2 and sensors with maximum coverage radius $r_{\max} = 17$ m. Minimum intersection level $c_{\min} = 1$ m. According to (18)–(22) optimal coverage radius $r_{optimal} = 8.75$ m with optimal intersection value $c_{optimal} = 7.43$ m. Schematically territory coverage presented on Fig. 7.

During next simulation rectangle-shaped territory $T_s = 10000$ m^2 with already covered area $T_{covered} = 2524.9$ m^2 get covered with sensor that has coverage radius $r = 10$ m and priority directions $pw_1 = 2$, $pw_2 = 3$. Sensor movement trajectory presented on Fig. 8.
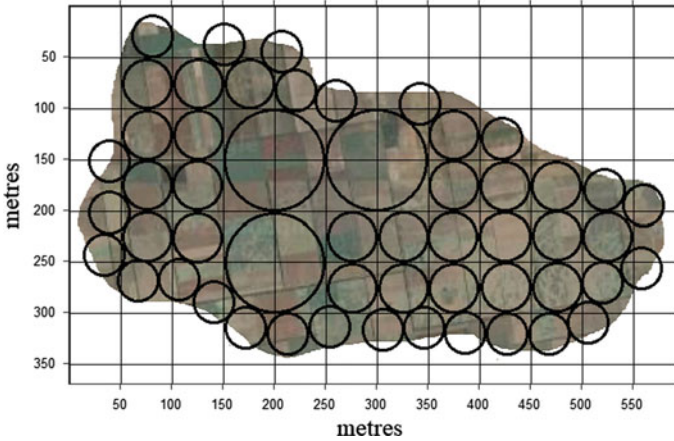
Fig. 5 Optimal territory coverage

With adding another sensor with parameters $r_2 = 10$ m and priority directions $pw_1^2 = 1$, $pw_2^2 = 8$ sensors' trajectories can be presented as in Fig. 9.

5 Conclusions

Thus, in the paper mathematical models and information technologies for covering territory by sensors are presented. In the first case approach that allows to choose the optimal combination of sensors of different types for territory coverage with minimum total energy consumption is described. Given approach based on knapsack optimization problem with some goal function and constraints modifications. It allows to cover territory with sensors that have different coverage radiuses and energy consumption and achieves minimum total energy consumption. Second presented model can be used in case when sensors' coverage radiuses intersection is needed. It allows minimize energy consumption with maximizing area coverage by solving nonlinear multicriterial optimization problem. The solution of the formulated problem is pair of the optimal coverage radius r and intersection level c with using of which minimization energy consumption with territory coverage maximization is achieved. Also, case of moveable sensors is described. A modified algorithm for constructing the optimal motion trajectory using the Bellman principle in the case of motion of several sensors is proposed. The possibility of the existence of an already covered area which may be an obstacle to the movement of sensors is considered. For described information technology software architecture and request-response examples are described. Also, computer simulation results which confirms the effectiveness of the proposed models and approaches are presented.
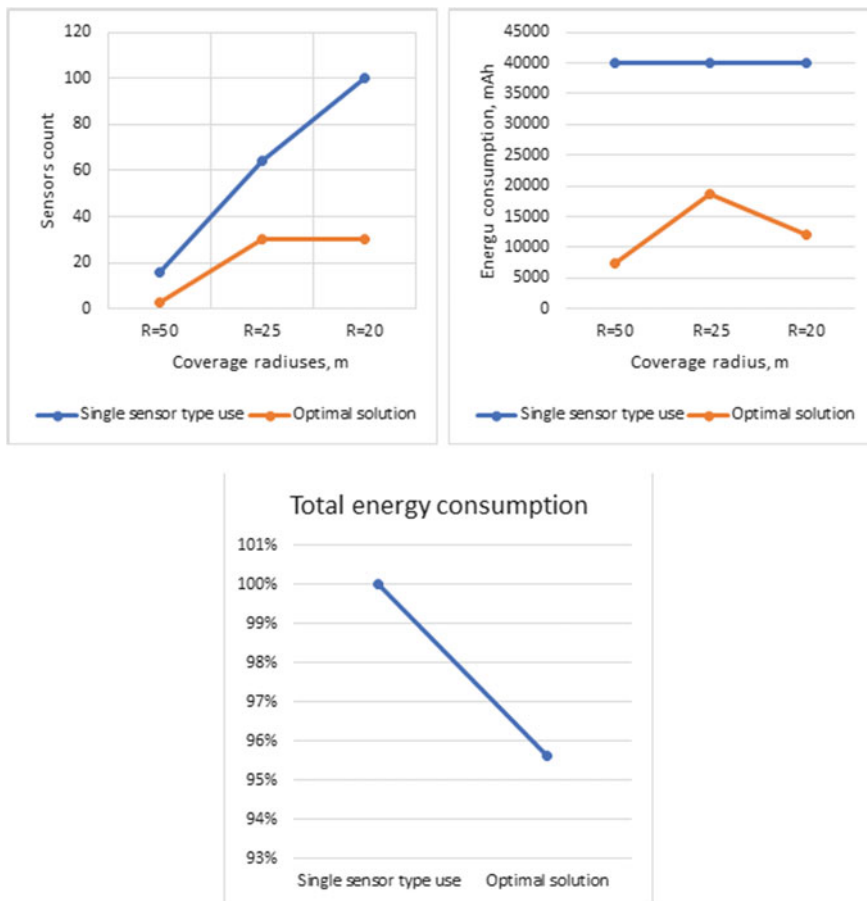
Fig. 6 Territory coverage and energy consumption comparison
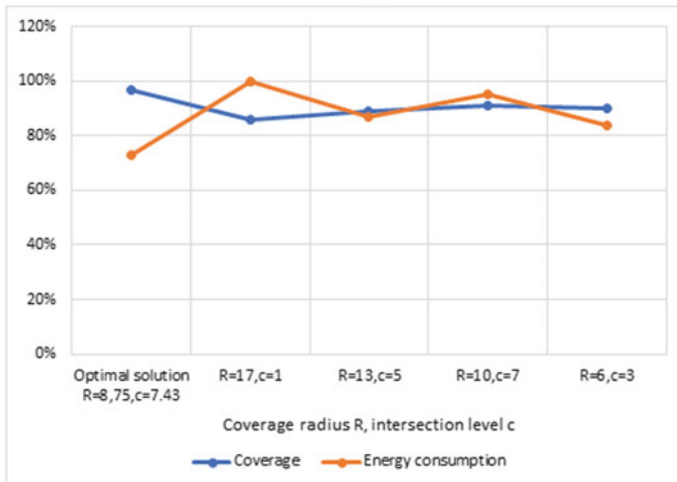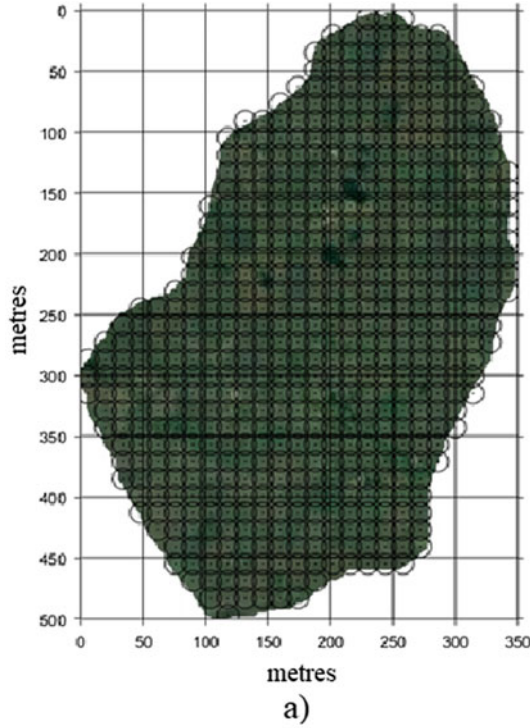
Fig. 7 **a** Optimal territory coverage result; **b** Coverage and energy consumption

Fig. 8 Sensor's movement trajectory with $r = 10$ m, $pw_1 = 2$, $pw_2 = 3$
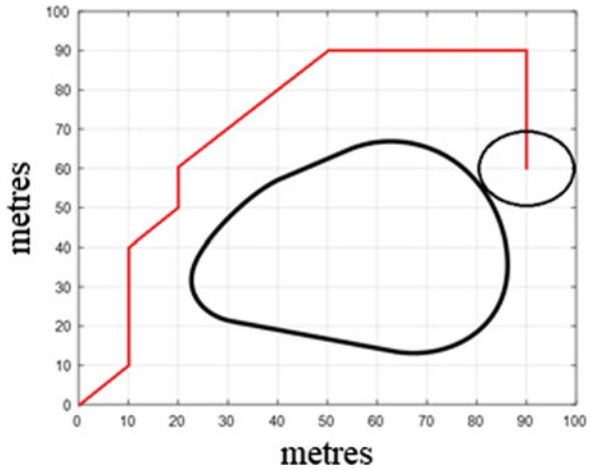
Fig. 9 Territory coverage with sensors $r_1 = 10$m, $pw_1^1 = 2$, $pw_2^1 = 3$; $r_2 = 10$ m, $pw_1^2 = 1$, $pw_2^2 = 8$
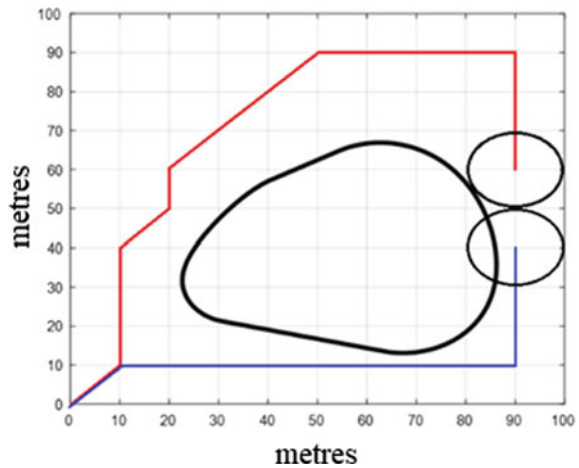
References

1. Pandey M, Mishra G (2019) Types of sensor and their applications, advantages, and disadvantages, emerging technologies in data mining and information security. In: Advances in intelligent systems and computing, pp 791–804. Springer, Singapore
2. Michalaki P, Quddus M, Pitfield D, Mageean M, Huetson A (2016) A sensor-based system for monitoring hard-shoulder incursions: review of technologies and selection criteria. In: MATEC web of conferences, vol 81, pp 1–8
3. Argyriou A (2015) Data collection from resource-limited wireless sensors for cloud-based applications. In: 2015 IEEE global communications conference
4. Dorozhynskyi O (2016) Geomantyka v monitoryngu dovkillya ta ocinci zagrozlyvyh ssytuacii: monografiia [Geomatics in environmental monitoring and threat assessment]. Lviv Polytechnic National University, Lviv

5. Danyliuk S (2016) Konceptyalni pidhody do vyrishennya zadachi optymalnogo rozmishchennya sensoriv v oblasti ekologichnogo monitoryngu [Conceptual approaches to solving the problem of optimal placement of sensors in the field of environmental monitoring]. In: Modern information technologies in the field of security and defense, vol 26, pp 45–48

6. Kocharov A, Yackin D (2016) Algoritm poiska optimalnogo razpolozeniya sensorov dla resheniya zadachi monitoringa prostranstva [Algorithm for finding the optimal location of sensors for solving the problem of monitoring space]. Softw Prod Syst 3(115)

7. Krishnamurthy P, Khorrami F (2017) Optimal sensor placement for monitoring of spatial networks. IEEE Trans Autom Sci Eng 1(15):33–44

8. Li C, Yang Q (2020) Optimal sensor placement algorithm for structural damage identification. Recent Patents Eng 14(69):69–81

9. Castello C, Fan J, Davari A, Chen R-X (2010) Optimal sensor placement strategy for environmental monitoring using wireless sensor networks. In: Proceedings of the annual southeastern symposium on system theory, pp 275–279

10. Krishnamurthy P, Khorrami F (2018) Optimal sensor placement for monitoring of spatial networks. IEEE Trans Autom Sci Eng 1(15):33–44

11. Akbarzadeh V, Levesque J, Gagne C, Parize M (2014) Efficient sensor placement optimization using gradient descent and probabilistic coverage. Sensors (Basel) 8(14)

12. Stoyan Y, Yaskov G (2010) Packing identical spheres into a cylinder. Int Trans Oper Res 17:51–70

13. Wang Y (2014) Survey on mobile social networking in proximity (MSNP): approaches, challenges and architecture. In: Wireless networks, vol 6, pp 1295–1311

14. Liguo T, Fomichev A (2016) The spatial flight route planning of unmanned aerial vehicles using the methods of mixed-integer linear programming. In: Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, vol 2, pp 53–66

15. Bunakov A, Grevtsov N, Melz N, Sazonov I, Tegin A (2012) Calculation of extremals and the search for a globally optimal trajectory in the problem of flying around obstacles. In: TsAGI scientific notes, vol 4, pp 84–92

16. Aseev S, Kryazhimskiy A (2015) The Pontryagin maximum principle and transversality conditions for a class of optimal control problems with infinite time horizons. SIAM J Control Optim, 1094–1119

17. Petrivskyi V, Shevchenko V, Bychkov O, Brazhenenko M, Petrov P (2021) CAD technology for optimal territory covering. In: CADSM 2021, 16th international conference on the experience of designing and application of CAD systems (CADSM). Lviv Polytechnic National University, Lviv

18. Ram M, Thain N (2010) Pick's theorem via Minkowski's theorem. Amer Math Monthly 8:732–736

19. Bouguera T, Diouris J-F, Chaillout J-J, Jaouadi R, Andrieux G (2018) Energy consumption model for sensor nodes based on LoRa and LoRaWAN. Sensors 18(2104):64–87

20. Jianhui L, Xingwei W, Min H, Hui CH, Fuliang L (2015) Solving 0–1 knapsack problem by greedy degree and expectation efficiency. Appl Soft Comput 41:94–103

21. Larson R, Boswell L, Stiff L (2004) Geometry. McDougal Littell

22. Bhowmik B (2013) Dynamic programming—its principles, applications, strengths, and limitations. Int J Eng Sci Technol 2

23. Petrivskyi V, Shevchenko V, Bychkov O, Brazhenenko M (2020) Information technology of the increasing sensors term of use considering their movement. In: IEEE 16-th international conference perspective technologies and methods in MEMS design (MEMSTECH). Lviv Polytechnic National University, Lviv, pp 86–89

Transport of Reactive Tracer in Compacting Multi-Fraction Bottom Sediments

Vladimir Maderich, **Igor Brovchenko**, and **Kateryna Kovalets**

Abstract The one-dimensional transport of reactive tracer (radionuclide) in the multi-fractional aquatic sediments is considered. In the model a transfer of tracer in sediment is governed by the molecular diffusion of dissolved phase in pore water, biodiffusion, phase exchange between dissolved and solid phases of tracers, exchange between fast and slow reversible phases of tracer, and compaction under gravity force resulting in a decrease of porosity with sediment layer depth. The transport equations for 2-step reactions were completed by an equation for fractions of sediments assuming that mixing in the bottom sediment is intraphase. The compaction dynamics was parameterized using an approximation of porosity by travelling wave profile. The results of an idealized numerical experiment on the constant deposition of contaminated sediments show that in the case of 2-step reactions the exchange processes in multifraction sediments were far from equilibrium.

Keywords Reactive tracer · Multifraction sediments · Compaction

1 Introduction

An exchange between the dissolved and solid phases is a significant process in the transport of reactive tracers in the marine environment. Settling of contaminated suspended sediments and resuspension by waves and currents results in tracer exchanges between the bottom and suspended sediment. The transfer of activity between the water column and the pore water in the bottom sediment is governed by the bottom boundary layer turbulence, which regulated diffusional processes [1]. The migration of tracer in the sediments is due to molecular diffusion, transport driven by bioturbation, bioirrigation, and also due to advection driven by surface waves and by subsurface groundwater flow [2]. Bottom sediments are important for the burial of particle-reactive radionuclides. After the Fukushima Dai-ichi accident, they are a

V. Maderich · I. Brovchenko (✉) · K. Kovalets
Institute of Mathematical Machine and System Problems, Glushkova av., 42, Kyiv 03187, Ukraine

K. Kovalets
e-mail: katkov0912@ukr.net

long-term source of radionuclide contamination through resuspension, bioturbation, and losses of soluble radionuclides from the pore water.

An important factor in the transport of tracer in the sediments is compaction processes caused by gravity forces resulting in a decrease of porosity with the depth of the sediment layer and corresponding fluxes of pore water and sediment particles. In general, the dynamics of this process is described by the nonlinear Gibson equation [3]. However, for a slow varied compaction process, the profile can be parameterized using an approximation of porosity by travelling wave profile [4]. Most of the studies of migration of reactive tracer were carried out for a single fraction of sediments. At the same time, the kinetics of fast reversible exchange between water and solids depends on the particle size. Therefore, it is necessary to extend transport models of reactive transport on compacting multifraction bottom sediments.

In the paper, the one-dimensional transport of reactive tracer (radionuclide) in the multi-fractional compacting aquatic sediments is considered. The model is described in Sect. 2. The results of an idealized numerical experiment are given in Sect. 3. Conclusions are drawn in Sect. 4.

2   Model Formulation

Consider one-dimensional transport of reactive tracer (radionuclide) in the multi-fractional aquatic sediments. It is assumed that reversible phase exchange is described by reactions of the first order. The bottom sediment layer is characterized by porosity $\varepsilon^w(z, t)$, solid volume fraction $\varepsilon^s(z, t)$, $(\varepsilon^w + \varepsilon^s = 1)$, a fraction of sediment particles of i-th class $\phi_i(z, t)$ $\left(\sum_{i=1}^{n} \phi_i = 1\right)$, and density of sediment fraction $\rho_{s,i}$. In the model, a transfer of tracer in sediment is governed by (i) molecular diffusion of dissolved phase in porewater, (ii) bioturbation (mixing of solid sediment particles by bottom organisms), (iii) phase exchange between dissolved and solid phases of tracers, (iv) exchange between fast and slow reversible phases of tracer, and (v) compaction under gravity force resulting in a decrease of porosity with sediment layer depth. The transport equations for dissolved tracer concentration in pore water C_d^b (Bq m^{-3}), fast particulate concentration in i-th fraction of the bottom sediments $C_{s,i}^b$ (Bq kg^{-1}), and slow particulate concentration in i-th fraction of the bottom sediments $\tilde{C}_{s,i}^b$ are written following [2] as

$$\frac{\partial \varepsilon^w C_d^b}{\partial t} = -\frac{\partial \varepsilon^w w_f C_d^b}{\partial z} + \frac{\partial}{\partial z}\left(\nu_B + \frac{\nu_D}{\psi^2}\right)\varepsilon^w \frac{\partial C_d^b}{\partial z}$$
$$- a_{ds}\theta\varepsilon^s\left(C_d^b \hat{K}_d^b - \hat{C}_s^b\right) - \lambda\varepsilon^w C_d^b, \tag{1}$$

$$\frac{\partial \varepsilon^s \phi_i C_{s,i}^b}{\partial t} = -\frac{\partial \varepsilon^s w_s \phi_i C_{s,i}^b}{\partial z} + \frac{\partial}{\partial z}\nu_B \varepsilon^s \frac{\partial \phi_i C_{s,i}^b}{\partial z} + a_{ds}\theta\varepsilon^s\phi_i\left(C_d^b K_{d,i}^b - C_{s,i}^b\right)$$
$$- a_{fs}\varepsilon^s\phi_i C_{s,i}^b + a_{sf}\varepsilon^s\phi_i \tilde{C}_{s,i}^b - \lambda\varepsilon^s\phi_i C_{s,i}^b, \tag{2}$$

$$\frac{\partial \varepsilon^s \phi_i \tilde{C}^b_{s,i}}{\partial t} = -\frac{\partial \varepsilon^s w_s \phi_i \tilde{C}^b_{s,i}}{\partial z} + \frac{\partial}{\partial z} \nu_B \varepsilon^s \frac{\partial \phi_i \tilde{C}^b_{s,i}}{\partial z} + a_{sf} \varepsilon^s \phi_i C^b_{s,i}$$
$$- a_{fs} \varepsilon^s \phi_i \tilde{C}^b_{s,i} - \lambda \varepsilon^s \phi_i C^b_{s,i} \tag{3}$$

where t is time (s), z is vertical coordinate directed upward (m), w_f is a vertical pore water velocity (m s^{-1}), w_s is a particle vertical velocity (m s^{-1}), λ (s^{-1}) is the tracer decay constant. The phase exchange between dissolved and reversible tracer phase is written in terms of desorption rate a_{ds} (s^{-1}) and distribution coefficient $K_{d,i}$ (m^{3}kg^{-1}) [5], whereas a_{fs} and a_{sf}(s^{-1}) are the direct and reverse exchange rates between fast and slow reversible phases of tracer, θ is a correction factor for desorption rate. The free solution diffusion coefficient ν_D was corrected for tortuosity ψ following [1] as $\psi^2 = 1 - 2 \ln \varepsilon^w$. The bioturbation is parameterized by the bioturbation coefficient $\nu_B(z)$. The total concentration of tracer $\hat{C}^b_s$ (Bq m^{-3}) is defined as

$$\hat{C}^b_s = \sum_{i=0}^{n} \rho_{s,i} \phi_i C^b_{s,i}, \tag{4}$$

The total distribution coefficient $\hat{K}^b_d$ is defined as

$$\hat{K}^b_d = \sum_{i=1}^{n} \rho_{s,i} \phi_i K_{d,i}, \tag{5}$$

The dependence of $K_{d,i}$ on sediment particle diameter d_i is written following [5] as.

$$K_{d,i} = \frac{\chi}{a_{ds} \rho_{s,i}} \frac{6}{d_i}, \tag{6}$$

where χ is an exchange velocity (m s^{-1}), d_i (m) is the sediment particle diameter.

The transport equation for fractions of bottom sediments at $\rho_{s,i} = \rho_s$ is described as

$$\frac{\partial \varepsilon^s \phi_i}{\partial t} = -\frac{\partial \varepsilon^s w_s \phi_i}{\partial z} + \frac{\partial}{\partial z} \nu_B \varepsilon^s \frac{\partial \phi_i}{\partial z} \tag{7}$$

Summing this equation on all fractions we obtain the transport equation for ε^s.

$$\frac{\partial \varepsilon^s}{\partial t} = -\frac{\partial \varepsilon^s w_s}{\partial z} \tag{8}$$

It was assumed in the derivation of (1)–(3), (7) that mixing in the bottom sediment is intraphase (i.e. porosity is not affected by biodiffusivity [5]). Therefore, an equation

for porosity can be written as

$$\frac{\partial \varepsilon^w}{\partial t} = -\frac{\partial \varepsilon^w w_f}{\partial z} \tag{9}$$

Combining (8) and (9) and assuming that $w_f \to 0$, $w_s \to 0$ at $z \to -\infty$ yields

$$\varepsilon^w w_f + \varepsilon^s w_s = 0. \tag{10}$$

Introducing moving coordinate system

$$\varsigma = H(t) - z, \tag{11}$$

where $H(t)$ is the coordinate of the water–sediment interface. Assume that in moving system coordinate profile of porosity was a profile of travelling wave, which does not depend on time. Following [5] this profile was approximated by exponent

$$\varepsilon^w(\varsigma) = \left(\varepsilon_0^w - \varepsilon_\infty^w\right) \exp(-k_\varepsilon \varsigma) + \varepsilon_\infty^w, \tag{12}$$

where k_ε is an attenuation coefficient, $\varepsilon_0^s = \varepsilon^s(0)$, ε_∞^s is an asymptotic value at large depth.

$$\frac{\partial}{\partial \varsigma}\left[\varepsilon^s(\omega + w_s)\right] = 0, \tag{13}$$

where $\omega = \partial H/\partial t$ is displacement velocity of the water-bottom interface described as

$$\omega = \frac{1}{\varepsilon_0^s \rho_s} \sum_{i=0}^{n} W_{p,i} S_{p,i} - w_s(0), \tag{14}$$

where W_p (m s^{-1}) is the settling velocity of suspended sediment, $S_{p,i}$ (kg m^{-3}) is suspended sediment concentration, and $w_s(0)$ is consolidation velocity at sediment surface.

Integration of (13) and determination of integration constant from asymptotic conditions: $w_s \to 0$, $\varepsilon^s \to 0$ at $\varsigma \to \infty$ leads to the relations between vertical velocities of particles and pore water and displacement velocity of the water sediment interface

$$w_f = -\frac{\varepsilon^w - \varepsilon_\infty^w}{\varepsilon^s}\omega, \quad w_s = \frac{\varepsilon_\infty^s - \varepsilon^s}{\varepsilon^s}\omega. \tag{15}$$

As seen in Eq. (15) the consolidation velocity $w_s(0)$ can be comparable with sedimentation velocity. Taking into account (15) the displacement velocity was rewritten as

$$\omega = \frac{1}{\varepsilon_\infty^s \rho_s} \sum_{i=0}^{n} W_{p,i} S_{p,i} \qquad (16)$$

Finally, governing system of equations describing transport of reactive transport of tracer in compacting sediments

$$\varepsilon^w \frac{\partial C_d^b}{\partial t} = -\varepsilon_\infty^w \omega \frac{\partial C_d^b}{\partial \varsigma} + \frac{\partial}{\partial \varsigma}\left(v_B + \frac{v_D}{\psi^2}\right)\varepsilon^w \frac{\partial C_d^b}{\partial \varsigma}$$
$$- a_{ds}\theta\varepsilon^s\left(C_d^b \hat{K}_d^b - \hat{C}_s^b\right) - \lambda\varepsilon^w C_d^b, \qquad (17)$$

$$\varepsilon^s \frac{\partial \phi_i C_{s,i}^b}{\partial t} = -\varepsilon_\infty^s \omega \frac{\partial \phi_i C_{s,i}^b}{\partial \varsigma} + \frac{\partial}{\partial \varsigma} v_B \varepsilon^s \frac{\partial \phi_i C_{s,i}^b}{\partial \varsigma}$$
$$+ a_{ds}\theta\varepsilon^s \phi_i\left(C_d^b K_{d,i}^b - C_{s,i}^b\right) - \lambda\varepsilon^s \phi_i C_{s,i}^b, \qquad (18)$$

$$\varepsilon^s \frac{\partial \phi_i \tilde{C}_{s,i}^b}{\partial t} = -\varepsilon_\infty^s \omega \frac{\partial \phi_i \tilde{C}_{s,i}^b}{\partial \varsigma} + \frac{\partial}{\partial \varsigma} v_B \varepsilon^s \frac{\partial \phi_i \tilde{C}_{s,i}^b}{\partial \varsigma} + a_{fs}\varepsilon^s \phi_i C_{s,i}^b$$
$$- a_{sf}\varepsilon^s \phi_i \tilde{C}_{s,i}^b - \lambda\varepsilon^s \phi_i \tilde{C}_{s,i}^b, \qquad (19)$$

$$\varepsilon^s \frac{\partial \phi_i}{\partial t} = -\varepsilon_\infty^s \omega \frac{\partial \phi_i}{\partial \varsigma} + \frac{\partial}{\partial z}\varepsilon^s v_B \frac{\partial \phi_i}{\partial z}. \qquad (20)$$

The boundary conditions at interface water–sediment $\zeta = 0$ are

$$\left(v_B + \frac{v_D}{\psi^2}\right)\frac{\partial C_d^b}{\partial \zeta} = -W_{PW}\left(C_d^w - C_d^b\right), \qquad (21)$$

$$v_B \frac{\partial C_{s,i}^b}{\partial \zeta} = -\frac{\omega}{\varepsilon_\infty^s}\left(C_{p,i}^w - C_{s,i}^b\right), \qquad (22)$$

$$v_B \frac{\partial \tilde{C}_{s,i}^b}{\partial \zeta} = -\frac{\omega}{\varepsilon_\infty^s}\left(\tilde{C}_{p,i}^w - \tilde{C}_{s,i}^b\right), \qquad (23)$$

$$v_B \frac{\partial \phi_i}{\partial \zeta} = -\frac{\omega}{\varepsilon_\infty^s}\left(\phi_i(0) - \phi_i\right), \qquad (24)$$

where C_d^w (Bq m^{-3}) is tracer concentration in the water C_d^w (Bq m^{-3}) is tracer concentration in the water, $C_{p,i}^w$ is the particulate concentration in i-th fraction of the deposited from water column sediments. The exchange rate W_{pw} is estimated from

corrected for surface roughness relation [6] as

$$W_{PW} = 0.1778 u_* \text{Re}^{-0.2} \text{Sc}^{-0.604}. \tag{25}$$

Here u_* is friction velocity, $\text{Sc} = \nu_M/\nu_D$ is the Schmidt number, ν_M is kinematic viscosity, ν_D is free solution diffusion coefficient, $\text{Re} = u_* \delta_* \nu_M^{-1}$ is the Reynolds number, δ_* is an average height of roughness elements, ν_M is kinematic viscosity. Alternatively, the values of $C_{s,i}^b$, $\tilde{C}_{s,i}^b$ and ϕ_i can be prescribed [7]. At the lower boundary of the sediment layer, the zero diffusion fluxes of whole variables are prescribed. The system of Eqns. (17)–(20) is solved using the finite-difference implicit scheme of second order.

3 Results of Modelling

Consider results of an idealized numerical experiment where the constant flux of contaminated by ^{137}Cs sediments of two fractions caused the formation of contaminated layer and redistribution of activity in different forms. The sediment layer thickness is 1 m, density of sediment particles $\rho_s = 2600$ kg m^{-3}, surface porosity $\varepsilon_0^w = 0.8$, bottom porosity $\varepsilon_\infty^w = 0.4$, $\omega = 6.4 \cdot 10^{-10}$ m s^{-1}, $\nu_D = 1.45 \cdot 10^{-9}$ m^{2}s^{-1}, $\nu_M = 1.6 \cdot 10^{-6}$ m^{2}s^{-1}, $\phi_1 = 0.2$, $d_1 = 10^{-5}$ m, $d_2 = 4 \cdot 10^{-5}$ m, $u_* = 10^{-2}$ m s^{-1}, $\delta_* = 10^{-3}$ m. The biodiffusivity is described by the relation

$$\nu_B = \nu_{B0} \exp(-\varsigma^2/\varsigma_{eff}^2), \tag{26}$$

where $\nu_{B0} = 10^{-12}$ m^2 s^{-1}, $\varsigma_{eff} = 0.02$ m. Kinetic parameters were chosen for ^{137}Cs [2, 5] as: $a_{ds} = 1.16 \cdot 10^{-5}$ s^{-1}, $\chi = 3.8 \cdot 10^{-6}$ m s^{-1}, $a_{fs} = 0.25 \cdot 10^{-7}$ s^{-1}, $a_{sf} = 0.25 \cdot 10^{-8}$ s^{-1}, $\theta = 0.1$, $\lambda = 1.06 \cdot 10^{-8}$ s^{-1}. Initially, the sediment layer was not contaminated.

Two scenarios were considered. In the first scenario, the 1-step reactions were taken into account. It was assumed that the concentration of ^{137}Cs in the settled to the bottom sediment particles was in equilibrium with a concentration in water. The corresponding interface values of variables were $C_d^b(0) = 23.4$ Bq m^{-3}, $C_{s,1}^b(0) = 372.9$ Bq kg^{-1}, $C_{s,2}^b(0) = 31.7$ Bq kg^{-1}, $\phi_1 = 0.2$. In the second scenario, the 2-step reactions were taken into account where corresponding interface values of variables were $C_d^b(0) = 23.3$ Bq m^{-3}, $C_{s,1}^b(0) = 371.2$ Bq kg^{-1}, $C_{s,2}^b(0) = 31.5$ Bq kg^{-1}, $\tilde{C}_{s,1}^b(0) = \tilde{C}_{s,2}^b(0) = 0$, $\phi_1 = 0.2$.

The simulation results for the concentration of ^{137}Cs in solid and dissolved phases calculated using 1-step and 2-step reaction models after 10 years of deposition are shown in Figs. 1 and 2. In the case of a 1-step reaction (Fig. 1) the exchange processes between dissolved and particulate phases of radionuclide were close to equilibrium excluding a thin layer near the interface. However, accounting of 2-step reactions (Fig. 2) results in the slow evolution of both slow and fast phases and
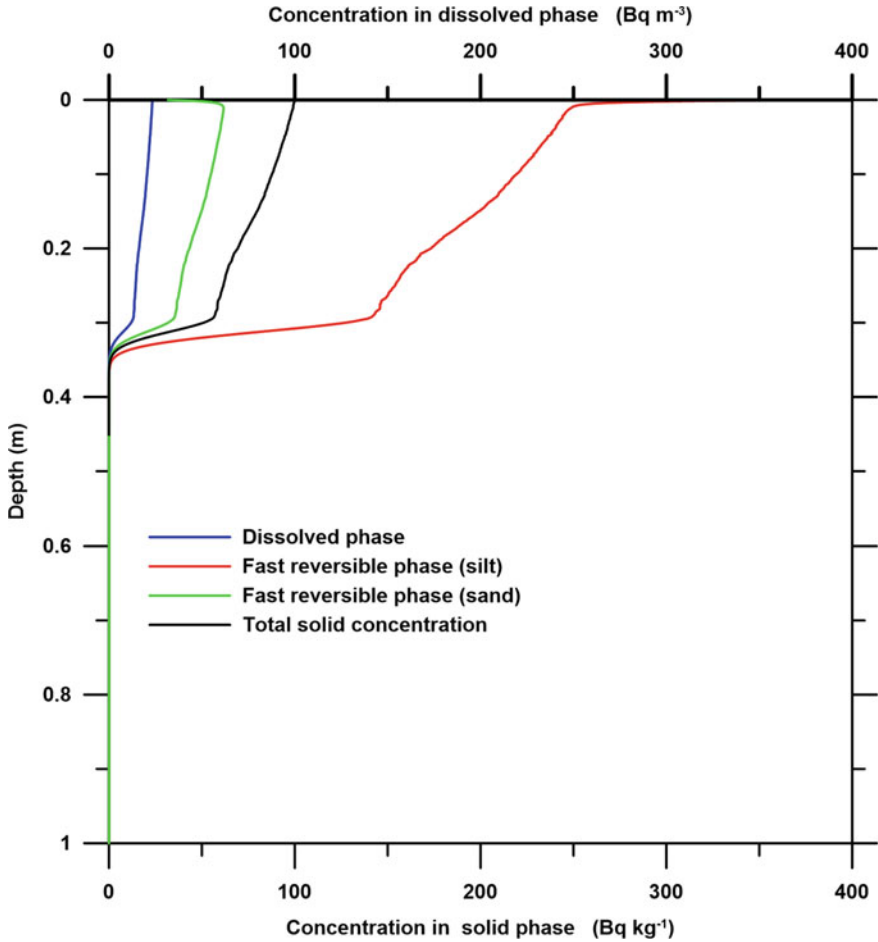
Fig. 1 The concentration of ^{137}Cs in solid and dissolved phases calculated using 1-step reaction models after 10 years of deposition

exchange between fractions of sediments through the pore water. In that case, the exchange processes in multifraction sediments were far from equilibrium. A decrease of porosity with depth caused by the compaction process changes the profile of the concentration of ^{137}Cs in solid and dissolved phases.

The effect of seasonal variations of deposition of sediment fraction was simulated using the periodic representation of fractions. The corresponding relation for silt fraction was approximated as

$$\phi_1 = \phi_{01} + \phi_{01} \sin\left(\frac{2\pi t}{T}\right), \tag{27}$$

Concentration in dissolved phase (Bq m⁻³)
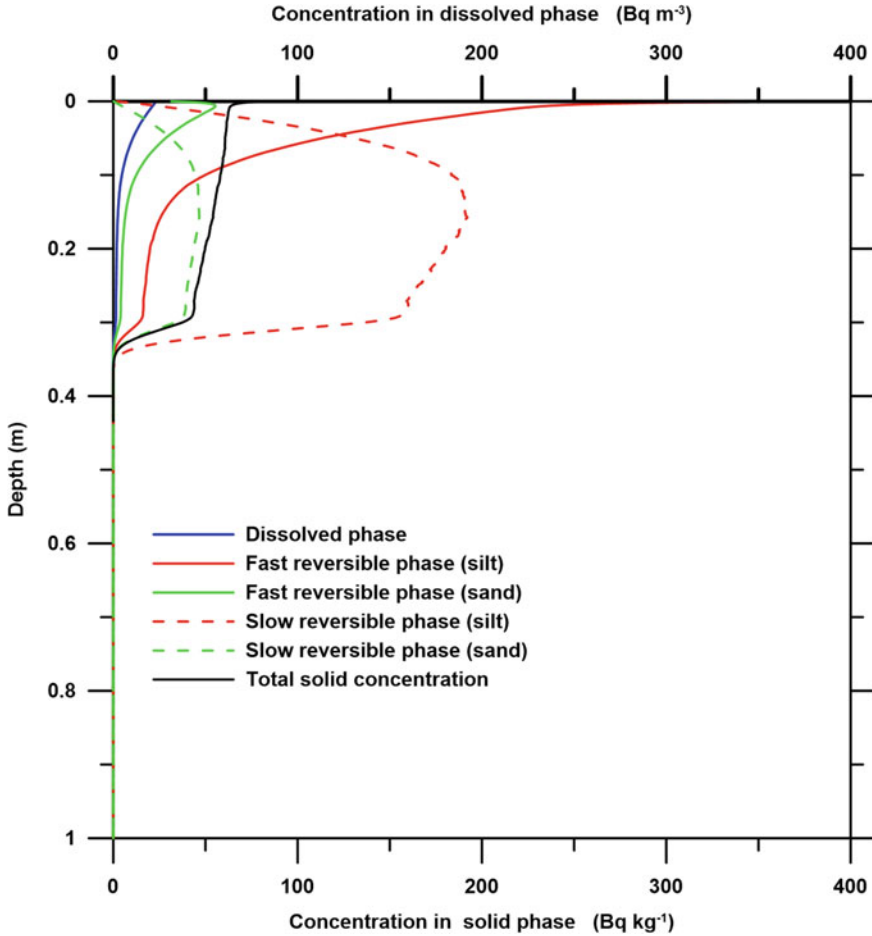
Fig. 2 The concentration of ^{137}Cs in solid and dissolved phases calculated using 2-step (b) reaction models after 10 years of deposition

where $\phi_{01} = 0.5, \phi_{02} = 0.3, T = 3.1536 \cdot 10^7$ s. As seen in Fig. 3 accounting variation in depositing sediments results in more complicated profiles of the concentration of ^{137}Cs for different fractions, although the profile of total solid concentration is more smooth due to mutual compensation by different fractions. An accounting of 2-step reactions (Fig. 4) results is similar variations of profiles of the concentration of ^{137}Cs for different fractions with decaying seasonal effects with depth due to diffusivity and slow kinetics.
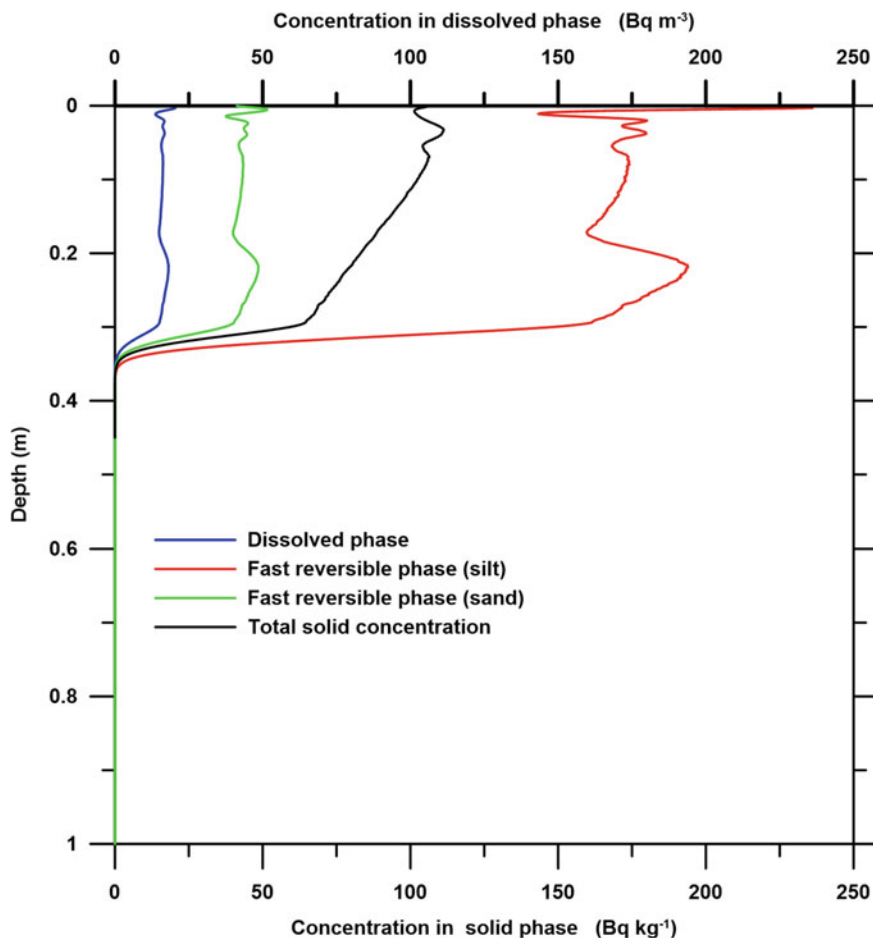
Fig. 3 The concentration of ^{137}Cs in solid and dissolved phases calculated using 2-step reaction models after 10 years of seasonally varying deposition of fractions

4 Conclusions

In the paper, the one-dimensional transport of reactive tracer (radionuclide) in the multi-fractional aquatic sediments is considered. In the model, a transfer of tracer in sediment is governed by the molecular diffusion of dissolved phase in pore water and biodiffusion. The phase exchange of tracer is described in the frame of 2-step reactions of first-order between dissolved phase and fast reversible phase of tracer on particle and exchange between fast and slow reversible phases of tracer on the particles. The transport equations for 2-step reactions were completed by the equation for sediment fraction transport assuming that mixing in the bottom sediment is
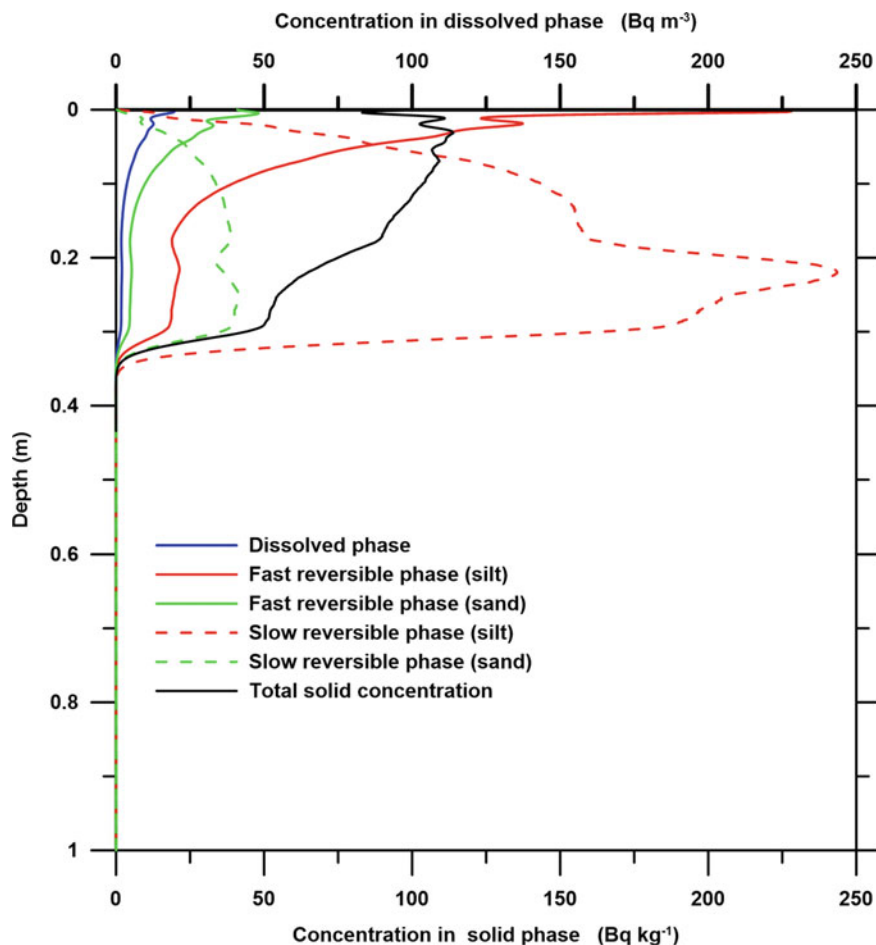
Fig. 4 The concentration of ^{137}Cs in solid and dissolved phases calculated using 2-step reaction models after 10 years of seasonally varying deposition of fractions

intraphase. It was found that the consolidation velocity due to gravity can be comparable with sedimentation velocity. An accounting of 2-step reactions results in the slow evolution of both slow and fast phases and exchange between fractions of sediments through the pore water. Therefore, the exchange processes in the multifraction sediments were far from equilibrium. Time variations in depositing fractions make profiles of reactive tracer result in more complicated profiles.

References

1. Boudreau BP (1997) Diagenetic models and their implementation. Springer
2. Maderich V, Jung KT, Brovchenko I, Kim KO (2017) Migration of radioactivity in multi-fraction sediments. Environ Fluid Mech 17(6):1207–1231
3. Winterwerp JC, Van Kesteren WG (2004) Introduction to the physics of cohesive sediment in the marine environment, vol 56. Developments in Sedimentology Series. Elsevier, Amsterdam
4. Boudreau BP (1996) A method-of-lines code for carbon and nutrient diagenesis in aquatic sediments. Comput Geosci 22(5):479–496
5. Meysman FJR, Boudreau BP, Middelburg JJ (2005) Modeling reactive transport in sediments subject to bioturbation and compaction. Geochim Cosmochim Acta 69:3601–3617
6. Shaw DA, Hanratty TJ (1977) Turbulent mass transfer rates to a wall for large Schmidt numbers. Amer Inst Chem Eng J 23:28–37
7. Massoudieh A, Bombardelli F, Ginn T (2010) A biogeochemical model of contaminant fate and transport in river waters and sediments. J Contaminant Hydr 112:103–117

Pillars for Establishing a Durable and Future-Proof IT Architecture Maturing Along with the NSC: Approaches from Continuous Integration to Service Mesh

Bernd Kratz [ID]**, Florian Wieduwilt** [ID]**, and Maxim Saveliev** [ID]

Abstract The development of future-proof and durable IT landscapes is becoming increasingly important, especially for large-scale projects spanning decades, where the requirements for many individual building blocks change over time and have to be continuously modified. An example for such an ambitious project is the construction of the New Safe Confinement (NSC) for the Chernobyl Nuclear Power Plant. It has been designed to provide multiple protections for a period of 100 years. A gigantic amount of diverse data is needed to realize this purpose, so IT structures must be developed and built in a foreseeing manner in order to achieve these ambitious goals. The concepts of "Continuous Development", "Continuous Integration", and "Continuous Deployment", as well as the technologies "Kubernetes" and "Service Mesh" are introduced as important pillars for establishing such a future-proof IT system. Finally, a self-developed concept for an early detection forecast system of the locations and concentrations of radioactive aerosols, implemented within the NSC and based on the described pillars, is presented and discussed.

Keywords Aerosol early detection forecast system · Future-proof IT architecture · Continuously quality assurance · New safe confinement

B. Kratz (✉) · F. Wieduwilt
Castalytics GmbH, Stiftshof 4, 34260 Kaufungen, Germany
e-mail: b.kratz@castalytics.de

F. Wieduwilt
e-mail: f.wieduwilt@castalytics.de; florian.wieduwilt@ifnano.de

F. Wieduwilt
Institut für Nanophotonik Göttingen e.V., Hans-Adolf-Krebs-Weg 1, 37077 Göttingen, Germany

M. Saveliev
Institute for Safety Problems of Nuclear Power Plants of NAS of Ukraine, str. Kirova, 36-a, Chernobyl, Kyiv region 07270, Ukraine
e-mail: m.saveliev@ispnpp.kiev.ua

Institute of Mathematical Machines and Systems Problems of NAS of Ukraine, Akademika Glushkova Ave., 42, Kyiv 03187, Ukraine

1 Introduction

The building of the NSC is one of the fundamental transformation stages of the 4th reactor of the Chernobyl Nuclear Power Plant towards a safer future. The NSC is designed to prevent the release of contaminated material as well as water intrusions, to protect the reactor from external influences during the dismantling works, and to guarantee this protection over a 100-year time span.

During the construction and run time of a such a complex generation-spanning operation, it is of uttermost importance to set up and establish a durable and future-proof IT architecture, which has the ability to mature along with the NSC.

Selected essential pillars that strongly contribute to the overall aspired IT landscape will be introduced and discussed by means of this publication. The focus will subsequently be on the following scope of functions:

(I) Receive and store sensor and measurement data from different sensor and measurement networks
(II) Processing and integration of scientific historical data
(III) Provide an IT landscape and corresponding methods to develop, test and operate different software models to mitigate risks for working personal in the NSC (e.g. through inhalation of radioactive aerosols in the working area)
(IV) Operation and maintenance of the emerging IT landscape over a long time period.

Out of this scope is the control of hardware related IT infrasystems such as Siemens SIEMATIC or SPS systems.

The listed scope of functions is an area in which the Castalytics GmbH offers its expertise.

2 Requirements for a Future-Proof IT System

The requirements for software solutions to be labeled as future-proof are very clear. A strategy must be developed, which facilitates the software to be viable and reliable for future developments in upcoming decades as an essential core element.

The main tasks are managing the complexity, change, and uncertainty during the construction and evolution of the software-system. However, it should be noted that all software-systems released into real-world applications generate certain risks when they are in operation. These risks vary from data loss or malfunctions to security vulnerabilities or misusage. The consequences can be severe, leading to possible damage or loss of property or, in case of working in dangerous environments like the NSC, even of life. Thus, the software-system must have an acceptable level of risk. A zero-risk software-system is not accomplishable so there will always remain a residual risk. For this reason, it is of utmost importance that comprehensive quality standards are met by the software and that a comprehensive risk assessment is carried out before the release.
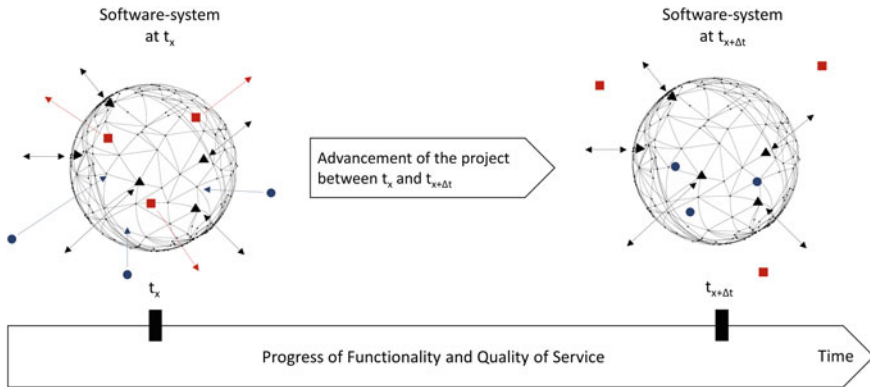
Fig. 1 The software-system modification process

According to to F. J. Furrer a future-proof software-system is defined as follows: "*A future-proof software-system is a structure that enables the management of complexity, change, and uncertainty with the least effort, with acceptable risk and with specified quality properties.*" [3].

Working on an IT system over a long period of time involves many changes that are made to the software. These essentially include changes to functionalities, which can be extended, added, replaced or removed. This transformation process is dynamic as shown in Fig. 1. It consists of many individual modifications or processes that are handled independently from each other. After completion of the desired adjustments the software-system has changed its functionality and modified its quality attributes.

In a large-scale software-system, thousands of projects may run in parallel. Their software modifications are often continuously brought into production [7].

The sequence "change requirements → project execution → deployment into production" repeats itself for the whole lifetime of the software-system. It is called an evolution cycle (Software Development Life Cycle, SDLC) and illustrated in Fig. 2. An evolution cycle may contain one or many projects.

Evolution cycles for durable and therefore future-proof software-systems must fulfil some characteristic criteria [3]:

(I) Modify or extend the functionality according to the stakeholder's requirements
(II) Change the level of the relevant quality of service properties (Improve or maintain)
(III) Avoid and reduce technical debt
(IV) Avoid or rectify architecture erosion
(V) Adapt to new or changed demands of the operating environment.

The future-proofing of current software is largely possible due to the fact that it is based on a flexible, versatile and adaptable architecture.

Fig. 2 The software development life cycle
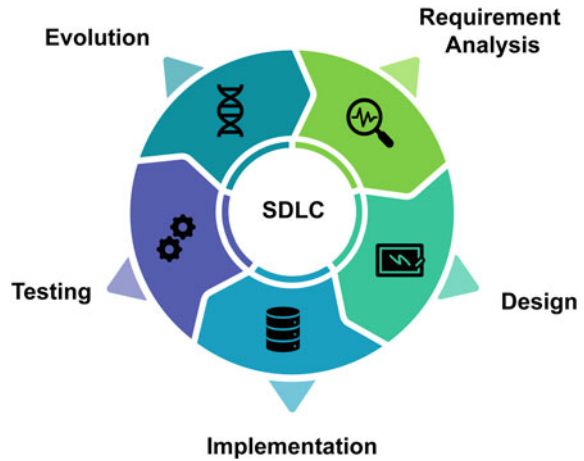
The software architecture becomes more important as a system becomes larger and more complex. Architecture is the strongest means to manage complexity, absorb change, and deal with uncertainty. It does so by building and maintaining a sound foundation for the construction and integration of software.

3 Essential Pillars for Establishing a Durable and Future-Proof IT Architecture

This chapter will introduce a design of the Castalytics GmbH for an early detection forecast system of the locations and concentrations of radioactive aerosols, implemented within the NSC. It is based on essential technical pillars used to establish a durable and future-proof IT architecture. These pillars are "Continuous Integration, Continuous Delivery, and Continuous Deployment", "Kubernetes", and "Service Mesh".

3.1 Continuous Integration, Continuous Delivery, and Continuous Deployment

Continuous Integration/Continuous Delivery (CI/CD) is a procedure which enables software delivery to a project or customers on a regular basis through certain automation techniques during the software development. CI/CD is also often accompanied by Continuous Deployment. This tool set can provide solutions when facing problems when introducing new software code to a system.

The term "Continuous Integration" originates from "Extreme Programming (XP)" [10]. It is a description of changes being made to the source code of a software. These changes are continuously integrated into the main branch. For successful integration, the modified source code is compiled and controlled with the help of automated tests. These checks take place before the changes are incorporated into the main branch as well as in the main branch itself. Not only the changes made by the developer are integrated, but also all changes made up to that point. Embedding the main branch in a dedicated integration environment ensures that the software also runs correctly in a production-like environment. This guarantees that conflicting changes introduced by different developers are quickly detected and fixed. Automated testing of the main branch is performed by a CI system. These tests are the basis for most of the current CD systems.

The expression "Continuous Delivery" was significantly shaped by Humble and Farley and merges different techniques and procedures for facilitating the software delivery process. In their book called "Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation" Humble and Farley draw the conclusion that previous considerations in this field of study included the relevant (sub-) processes, but not the interactions between them [7]. According to this approach, the essential parts of the value streams are missing, which makes it difficult to implement the processes efficiently and reliably. The application of Continuous Delivery into the workflow offers versatile advantages. An important vehicle is the reduction of the so-called cycle time. The cycle time depicts how long it takes for a certain software feature to be released or delivered to the customer.

In CD, every development stage involves test automation and code release automation. Working according to CD enables the possibility to react more promptly to errors or changes regarding the constructed software system. Detecting and correcting software errors early or steadily during software development reduces costs and increases the reliability of new software versions. New software code can be implemented with minimal effort in a fast processing time.

The final step in a CI/CD sequence, shown in Fig. 3, is Continuous Deployment. It can be considered as an extension of CD, which automates the release of a production-ready software build to a code repository. Continuous Deployment then automates the release of software to production.

3.2 Kubernetes

Kubernetes is a platform technology for creating, managing and deploying distributed applications. These distributed applications can be very diverse regarding their structure. Despite their differences, one thing they have in common is that they consist of single or multiple individual programs, which are executed on separated computers. These applications accept data input as well as data change and then return the results. A very prominent example working with this kind of system infrastructure, often referred to as "Container Ship", is Netflix [2].
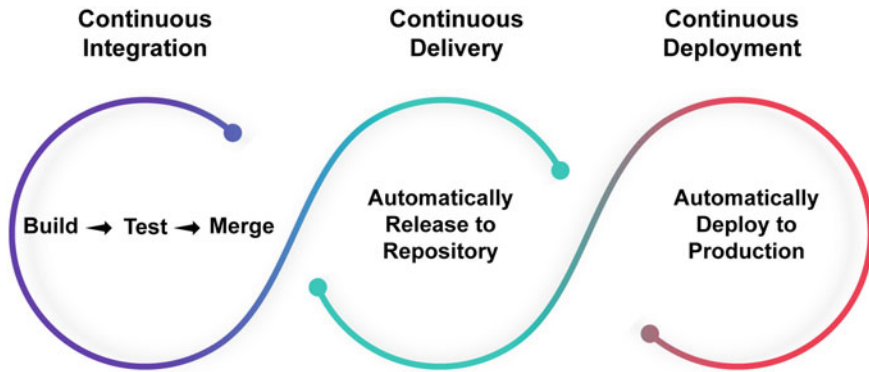
Fig. 3 Sub-steps in a CI/CD sequence

The "Container Ship"-approach therefore provides an ideal basis for building up a long-lasting distributed IT architecture. However, here the starting point of the system design is the consideration of how to create the application container images that contain these programs and form the individual fragments of the distributed system. Summarized Kubernetes is an open source orchestrator for deploying applications in the form of software containers. The technology was originally developed by Google as a result of the long and profound work and extensive experience gained in deploying reliable and scalable systems in form of containers by means of application-oriented APIs (Application Programming Interface). Originally introduced in 2014 [9], Kubernetes has become one of the most comprehensive and lauded global open source projects. It has become the standard API for building cloud-native applications.

In 2015 Kubernetes was committed to GitHub. As of this point, developers could start to make their own contributions to the system. It did not take long for large IT companies like Microsoft, IBM, Docker, Red Hat, and others to show their interest and they started contributing to the Kubernetes themselves. As a result, various modules were developed for Kubernetes, ensuring the quality of the basic orchestrator and securing the ongoing development. In 2017 and 2018, in cloud services specialized companies like Amazon Web Services (AWS) and DigitalOcean opened up their systems for integration of Kubernetes [14]. Today, Kubernetes is a very popular open source platform for managing containerized applications. It has the advantages of being extensible and portable and letting micro components taking care of the basic features of the orchestrator. Because of its proven infrastructure for distributed systems that is suitable for cloud-native developers of all scales, varying from a cluster of tiny single-board computers like Raspberry Pis to a cluster of high-end premium computers. It provides the necessary software to successfully build and deploy reliable and scalable distributed systems. Figure 4 shows the basic building blocks of the Kubernetes architecture.
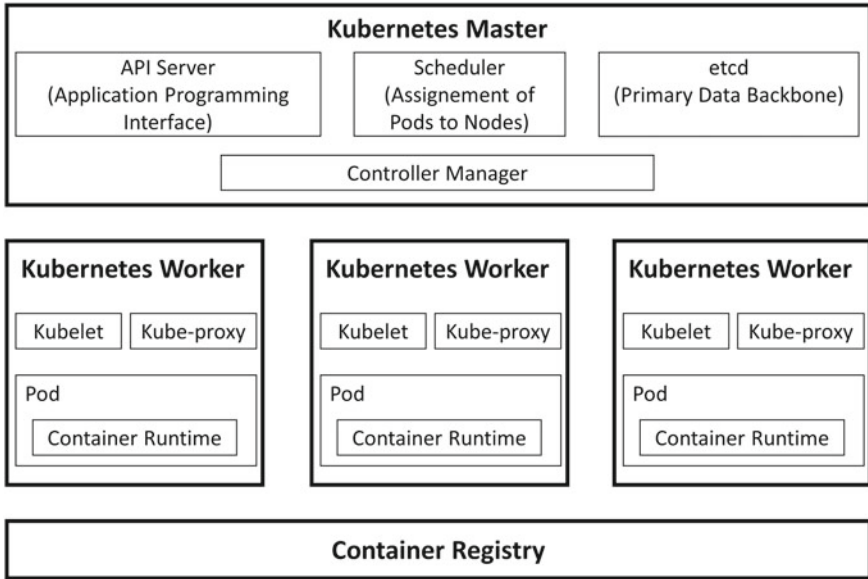
Fig. 4 Schematic overview of the Kubernetes architecture

The Kubernetes architecture follows a client–server architecture where the master is installed on a machine and nodes are distributed across multiple machines accessible via the master. The Kubernetes Master and Kubernetes Workers are part of the Kubernetes control plane, whereas the container registry may lie outside of the control plane.

3.3 Service Mesh

The final pillar presented in this contribution is the Service Mesh. A Service Mesh can be described as an infrastructure layer assigned to control services to make their intercommunication safe and reliable [12]. It usually consists of a control plane and a data plane [1, 12]. The data plane is responsible for various tasks such as service discovery, health checks, traffic control, monitoring, and security functions [1]. The Service Mesh system provides functions that are complementary to Kubernetes.

The aim of deploying a Service Mesh is to have an integrated platform that can actively monitor all service traffic. In doing so, a Service Mesh can address concerns related to fault tolerance, service security, and application monitoring as well as allowing systematically tracking and correlation interactions between various services. The information collected this way will be presented in the form of request timelines. This capability is termed as "Tracing". Tracing enables the debugging of how a request flows from one service to another. A monolith application delivers a

request log that could identify the request as being a service. In a Service Mesh, this situation is fundamentally different. In this case, a service is calling the next one, which makes it fairly challenging to debug a request lifecycle. Here, Tracing enables us to reconstruct such a request flow to analyze performance issues in our software application.

Usually, microservices in a Service Mesh achieve this by forwarding tracing context headers. This procedure allows to visualize interservice dependencies by distributed Tracing. In addition, metrics regarding request volumes and failure rates can be taken up. Another main topic is the handling of service log data. In a conventional software application, usually only one log file is generated, whereby with a microservice software architecture the opposite is the case. Several logging automatisms are active at the same time, so all of them must be analyzed to understand the applicable behavior.

In conclusion, it can be said that a service mesh offers the possibility of centralized logging, with the goal to provide operational visibility within the microservices.

4   Self-Developed Aerosol Early Detection Forecast System

4.1   *Foundation of the Forecast System*

The prerequisites that are necessary for an IT landscape to be designated as a future-proof were discussed in the previous sections. Three IT concepts that meet the criteria were presented. These concepts are also used by the Castalytics GmbH as essential pillars to develop the aspired forecast system.

Basically, it consists of different sensor-based measurement networks and systems to measure a variety of physical parameters such as windspeed and -direction, temperature, humidity, dust, radiation, differential pressure, vibrations on the terrain and others. The different sensors will be distributed over the entire area of the NSC resulting in a multi-dimensional sensor positioning for each spatial direction x, y, z and additionally the corresponding time at which the sensor was located at the described position. The sensors are ideally movable in order to generate more flexibility and variations regarding the measurement locations. The forecast system provides methods and components to store and handle heterogenous data from different sources, such as current measurement data for future measurement networks, data of historical measurement systems or archived data. A unique improvement in this approach is the consolidation of the enumerated heterogeneous data into a main database. Additionally, the forecast system inhabits methods and components to develop and run the different flow models, e.g. rapid Computational Fluid Dynamic (CFD) [8] models and Artificial Intelligence (AI) as well as Machine Learning (ML) algorithms. This highly versatile, huge database is equipped with all the necessary information and will be the source for all future forecast models.

The goal being pursued is to provide constant and increasing quality assurance methods for the different models used and to quantify their accuracy as well as the uncertainty of the resulting forecast system for the distribution of radioactive aerosols in the NSC. The methods and components used are designed for long-term operation and maintenance of the forecast system (~100 years) and for the conservation of the know-how and the knowledge over such a long period of time.

4.2 Basis of the Forecast Models and Related Works

The approach to increase the forecast quality is based on ensemble forecast methods [16]. This will lead to an automatic and permanent quality measurement between the future orientated forecast and the historical measurement data of each algorithm. The evaluation of different forecasting models is the basis for the choice of the appropriate model for a prediction task.

In past projects, as former part of the enercast GmbH, experience has already been gained with regard to deterministic error scores. They were applied for model comparison of power forecasting algorithms, based on AI and ML algorithms, in the field of weather forecasting [4, 5]. This experience will also be used for the introduced forecast system for the distribution of radioactive aerosols and will be developed further.

A major goal to be achieved is to increase the accuracy and reduce the uncertainty of the predictions. The time intervals in which forecasts for the upcoming 24–48-h period are going to be made are constant frames between 15 and 60 min each.

The continuous ranked probability score (CRPS) is one of the most prominent approaches to describe the quality of a predictive distribution, which was first introduced by Matheson and Winkler in 1976 [6, 11]. The CRPS enables the analysis of the difference in the area between two cumulative distribution functions (CDFs). CDFs, in turn, are a central concept in the examination of probability distributions of real numbers. Every probability distribution and every real-valued random variable can be assigned a distribution function. The CRPS is mathematically described by the following equation:

$$CRPS = \int_{-\infty}^{\infty} \left(\hat{P}(y) - H(y - o) \right)^2 dy \tag{1}$$

In this equation $\hat{P}(y)$ represents the CDF of the predictive distribution at a particular point in time. $H(y - o)$ is the corresponding CDF of the observation, which is a Heaviside step function with a step at the location of the observation. If a deterministic forecast is evaluated with the CRPS, the error function becomes equivalent to the mean absolute error (MAE) because in this case the CDF of the deterministic forecast is a step function. In addition to the probabilistic quality measure described above, deterministic quality measures such as the Nash–Sutcliffe efficiency (NSE), mean

error (ME), mean absolute error (MAE) or root mean square error (RMSE) can be used to verify the forecasts [4, 5, 15]. For ensemble forecasts the quality of the ensemble mean is determined and for probabilistic forecasts the quality of the expected value. A continuous improvement and comparison of different forecast models based on ensemble forecast methods in combination with a detailed analysis of possible errors and their distributions summarized in error scores, will be a focal point in future development.

4.3 Conception and Explanation of the Forecast System

The technical options presented in Sects. 3.1, 3.2, and 3.3 form the pillars of our aspired future-proof IT landscape for an early detection forecast system of the locations and concentrations of radioactive aerosols, implemented within the NSC. During the dismantling and other tasks, air-spraying of accumulated radioactive materials and formation of radioactive aerosols are going to take place. A large amount of dust will be re-suspended by the dismantling works and release radioactive aerosols in the NSC. The inhalation of such radioactive aerosols is one of the major radiation-damaging factors of the personnel working at the site. It is a key task to reduce this latent potential danger to the personnel. This can be realized by setting up a versatile early detection forecast system. A schematic overview of the system is shown in Fig. 5. It operates according to the following principle: The Sensors 1 to n collect measurement data from the ambient air of the NSC. The measurement data from the sensors are then fed into the *New Dust Measurement System.* In parallel, there are two additional system sources that can also provide measurement data. On the one hand, measurement data from the different *Measurement Systems* are supplied and on the other hand, *Historical Data* are also examined. These measurement data are then transmitted to two different *Data Adapters.* The function of the *Data Adapter* is to receive data from a measurement network, like the *New Dust Measurement System* and deliver it to the *Data Gatherer.* The *Data Adapter* consists of two components: The *Data Driver* and the *Data Connector.* The *Data Driver* is specific to the measurement network and contains the interfaces to receive data from the measurement network. The *Data Adapter* is a universal component. This architecture is long-term support (LTS)-aware. Conveniently, to import data from another measurement system or historical data into the system, only the data driver has to be modified.

The *Data Gatherer* provides different interfaces for third party systems or components to deliver data. Data can be received or gathered by different protocols, e.g. REST, FTP, File Sharing, NFS, SMB/CIFS, HTTP, and more. The *Data Gatherer* receives the original data and enriches it with meta data, like the originator-ID, the interface, and the timestamp. This meta data will be useful for future data analytics.

In addition, the *Data Gatherer* always stores the original data as well, which provides extensive error resistance. In the event of an error, data processed in the *Abstraction Layer* is discarded and can be recalculated from the original data.
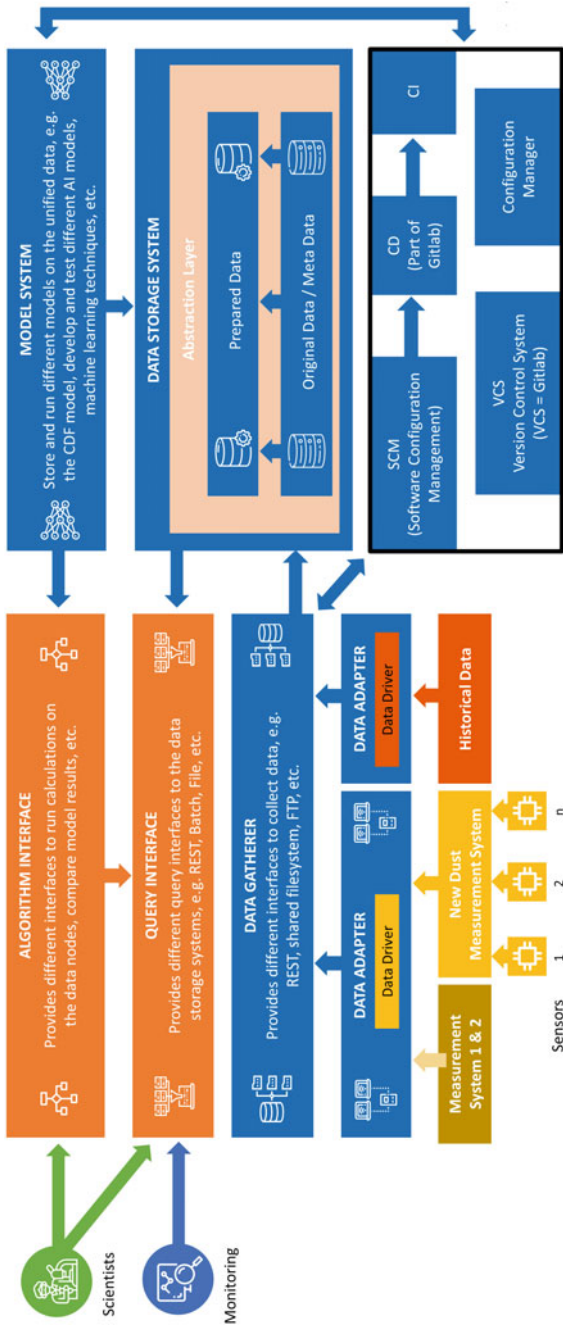
Fig. 5 Layout of the forecast system

The enriched data is sent to the *Data Preparer*. This architecture is likewise LTS-aware. Multiple instances of the *Data Preparer* can be deployed to increase the fault tolerance. The *Data Preparer* receives data in different formats, normalizes the data to one common format and stores the data in the *Data Storage*. The received data will be stored twice. First in the original format along with the meta data. With this approach, the data can be later recovered in the original format, for debugging or special analytics. The original data will be converted to a common, unified format, which then allows common operations on the data. The *Data Preparer* takes also care about details, such as character encoding (ISO-8859-1, LATIN, Cyrillic encoding) and transforms the data to UTF-8. UTF-8 supports all worldwide character encodings, including Japanese and Chinese. Each data record receives two timestamps, the original timestamp and one common unified timestamp in UTC.

The *Data Storage* will be a clustered, fault-tolerant big data system, such as Apache Cassandra. The *Data Storage* should provide features such as cluster awareness or data center awareness and an automatic synchronization mechanism. The storage and access of the data will be realized by means of an *Abstraction Layer*, such as Spring Boot. This enables the flexibility to exchange the underlying data system with less changes in the program code. The *Abstraction Layer* provides CRUD (Create, Read, Update and Delete) access to the data. It will be part of the *Query Interface* or the *Data Preparer*.

The *Query Interface* provides a unified read-only access to the stored data. Scientists or other users can run queries or access the data via different interfaces, e.g. REST, batch, file, etc. The data of different systems can be queried and comparand in a unified way by a domain specific language (DSL). The *Query Interface* supports also simple statistics (min, max, sum, avg., mean, std. deviation, count, etc.) and time-based query operations such a sliding time windows or fixed time periods. Threshold queries, or data produced by the *Algorithm Interface*, such as a trend analysis or forecasts, can be stored and queried. This allows the integration to existing monitoring or reporting systems.

The *Algorithm Interface* controls the computations of the data and scientists can develop and test different models and computations with it. Typical big data algorithms such as "map-and-reduce" or other distributed computations, such as Apache SPARK are supported. The *Algorithm Interface* allows the implementation of algorithms in different programming languages, like Python, Scala, Java and others. The result of the calculations can be stored and used by the *Query Interface* for integrating it into existing reporting or monitoring systems. The *Algorithm Interface* permits the development and the execution of named or scheduled calculations, like daily, weekly or monthly reports.

It obtains its fundamentals from the connected *Model System*, which stores developed models, such as the CFD model, new A.I. models, machine learning algorithms. It supports a versioning and release system for the models by tagging and stagging the different models. The versioning system is providing a continuous build system (CBS) such as GitHub or GitLab. Each committed change in the model is tracked. After a change, the model is automatically built by the model system and tested against pre-defined test scenarios. It is a distributed version control system for

tracking changes in any set of files, e.g. program code. The staging system is providing a continuous integrations system. This allows the automatically deployment of new well tested models in the cluster.

The *Model System* can be further developed from the *Software Configuration Management* (SCM) and the *Version Control System* (VCS). The procedure for the versioned development steps comes from the CI/CD concept described in Sect. 3.1. The key element is that data from different measurement networks are obtained and integrated into the development of the *Model System.* The CD system continuously checks the quality of the selected model. An action called commit or push (publication of changes) initializes automatics tests in the CD system. Automatic quality and assurance checks of the software and the model(s) are then performed. After successful testing, the new model can be released and automatically installed with the help of the CI. The release generates a unique stamp ("TAG"), which equals a version number. All components with a certain version number can now be deployed automatically. This means the software has a traceable status. The *Model System* can thus be improved iteratively. Each iteration of the model system receives its own version number. These different model systems can then be evaluated for their accuracy by processing historical data. The result is the quality of the model used.

The application of the diverse possibilities of Kubernetes makes it possible to subsequently evaluate and check different versions of the models.

4.4 Résumé of the Forecast System

Summarized the forecast system delivers the following value propositions:

First it provides a continuously quality assurance. Scientists can develop and update the aerosol flow model in the NSC and the forecast algorithm according to the ongoing dismantling work. The system should provide methods and components to measure and assure the quality of the aerosol flow models and the forecast algorithms. The quality of the model is continuously monitored and the forecast algorithms raise alarm in case of unexpected deviations. In addition, the quality of the predictions is continuously improved in a self-learning manner with each new data package collected, so the forecasts will get subsequently more precise with time.

Secondly it addresses the challenges of long-term support. The innovation cycles in the IT sector are very fast and standards can change quickly. The lifetimes of IT components and standards are very heterogeneous. The components of an IT system, such as operating systems, programming languages, runtime systems (Java, C#, etc.), and databases each have different lifetimes, which reduces the total lifetime of the system. Another important factor that needs to be considered here is the transfer of the built-up knowledge. It is tightly coupled to the people who are working with it. A striking example of such a linkage is the COBOL crisis in the insurance IT systems [13].

5   Conclusion and Outlook

The construction of future-proof IT landscapes is a challenging task in the present time. Several mentioned multifaceted requirements must be fulfilled for a software-system to be acknowledged as future-proof.

The structure of an early detection forecast system for the locations and concentrations of radioactive aerosols within the NSC designed by the Castalytics GmbH was presented. With this software architecture, a very applicable solution for the given problem was developed. The scope of functions required in the introduction from (I) to (IV), based on the criteria for evolution cycles for durable and therefore future-proof software-systems, was completely fulfilled.

This means the forecast system has to be adaptive and must be updated continuously, because the dismantling works continuously changes the aerosol flow model in the NSC. This is why flow measurement data must be continuously recorded and evaluated. At this point the introduced concept of CI/CD can be applied in a sustainable and future-proof manner.

By using these and other techniques (Kubernetes, Service Mesh) it is possible to provide methods and components to store and handle heterogenous data from different sources, such as existing data for future measurement networks, data of historical measurement systems or archived data. Additionally, methods and components to develop and run the different flow models, e.g. rapid computational fluid dynamic models and artificial intelligence as well as machine learning algorithms are provided.

This infrastructure is supported by quality assurance methods of the different models and the forecast systems to measure their accuracy and the uncertainty of the forecast system. Ensemble forecasting algorithms and error scoring methods are applied. Every IT component, such as the operating system, the database systems, the data storage systems, the programming languages, the runtime systems, the developed IT components (data adapter, data gatherer, etc.) is designed to be exchangeable with less effort or low costs. A crucial part will be the costs of the different sensors. The objective is to reduce the costs of the individual sensor systems as much as possible, which could be achieved, for example, through the use of inexpensive single-board computers. With a suitable price, the sensors could be regularly replaced with new ones over a period of a few years, due to the accessible infrastructure of the entire system.

With this toolset and preserved human knowledge, the foundation for a long-term operation and maintenance of the forecast system is established.

References

1. Calcote L (2018) The enterprise path to service mesh architecture: decoupling at layer 5 (1st edn). O'Reilly Media, Sebastopol, USA

2. CIO Homepage. https://www.cio.de/a/devops-in-einer-two-speed-architektur,3251208,2. Last Accessed 30 March 2021
3. Furrer FJ (2019) Future-proof software-systems: a sustainable evolution strategy (1st edn). Springer Vieweg, Wiesbaden, Germany
4. Gensler A (2019) Wind power ensemble forecasting—performance measures and ensemble architectures for deterministic and probabilistic forecasts, intelligent embedded systems (12). Kassel University Press GmbH, Kassel
5. Gensler A, Sick B, Vogt S (2016) A review of deterministic error scores and normalization techniques for power forecasting algorithms. In: IEEE symposium series on computational intelligence (SSCI) 2016, pp 1–9
6. Hersbach H (2000) Decomposition of the continuous ranked probability score for ensemble predictionsystems. Weather Forecast 15(5):559–570
7. Humble J, Farley D (2015) Continuous delivery: reliable software releases through build, test, and deployment automation. Addison-Wesley, Upper Saddle River, USA
8. Krukovskyi PG, Diadiushko YV, Garin VO, Tryfonov OV, Kabanov YY (2020) CFD model as a digital twin of the radiation state of the new safe confinement of the Chernobyl NPP. Voprosy Atomnoj Nauki i Tekhniki 4–128:54–62
9. Kubernetes Homepage. https://kubernetes.io/blog/2018/07/20/the-history-of-kubernetes-the-community-behind-it/. Last Accessed 30 March 2021
10. Martin Fowler Homepage, Continuous Integration. https://www.martinfowler.com/articles/continuousIntegration.html. Last Accessed on 29 March 2021
11. Matheson JE, Winkler RL (1976) Scoring rules for continuous probability distributions. Manag Sci 22(10):1087–1096
12. Miranda G (2018) The service mesh: resilient service-to-service communication for cloud native applications (1st edn). O'Reilly Media, Sebastopol, USA
13. SCRY Analytics Homepage. https://scryanalytics.ai/resolving-the-cobol-crisis-with-artificial-intelligence-full-article/. Last Accessed 6 April 2021
14. Techrepublic Homepage. https://www.techrepublic.com/article/digitalocean-simplifies-container-deployment-for-developers-with-kubernetes-support/. Last Accessed 1 April 2021
15. Wang WC, Chau KW, Cheng CT, Qiu L (2009) A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. J Hydrol 374(3–4):294–306
16. Zamo M, Naveau P (2018) Estimation of the continuous ranked probability score with limited information and applications to ensemble weather forecasts. Math Geosci 50:209–234

Optimal Control of Buried Point Sources in a Two-Dimensional Richards-Klute Equation

Sergey Lyashko ⓘ **, Dmitriy Klyushin** ⓘ **, and Andrii Tymoshenko** ⓘ

Abstract The article is dedicated to a two-dimensional humidification problem, described by Richards-Klute equation. Several simplifications are made: water is assumed incompressible, external pressure and temperature are constant. The initial state and desired function are known, while the optimal source power should be calculated. The initial equation is transformed using the Kirchhoff transformation. New scaled values are introduced to get linear dimensionless equation, which can be easily discretized over space and time. Then implicit (or explicit) methods and numerical methods are applied to solve the system. In general, the problem can also take into account heat transfer, chemical reactions or surface heterogeneity leading to more complicated equations, so the research will continue.

Keywords Mathematical simulation · Control · Optimization · Richards-Klute equation

1 Introduction

Richards-Klute equation is a general equation describing moisture transfer in porous medium. It can be quasilinear or non-linear depending on moisture conductivity and diffusion [1, 2], so for some cases it is possible to apply Kirchhoff transformation and introduce new dimensionless values to make transition to a linear problem [3]. As an alternative, one can use iterative methods and their hybrid combinations for the Richards equation [4].

When linear equation is formulated, variation algorithm, consisting of direct problem, conjugate problem and new source power approximation can be applied [5]. The optimization criteria can be formulated as minimization of modular difference between humidity according to current source power approximation and the desired state at the last time step. Existence and uniqueness of the optimal source power and a polynomial approach to get the initial approximation, were discussed for similar problems in [6, 7].

S. Lyashko · D. Klyushin (✉) · A. Tymoshenko
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

To solve the linear system for direct and conjugate equations explicit and implicit methods were used [8–10]. In case of explicit method, the amount of time steps is larger to gain convergence, but calculations are relatively simple. In case of implicit scheme, Gaussian Elimination Method was used to get exact solutions and Jacobi method as an alternative. All three approaches showed alike results.

To sum up, the current work represents a combined approach to build a logical connection between the initial problem, its simplifications and numerical approach resulting into optimization of source power.

2   Process Description and Its Model

Consider a two-dimensional problem, described by the following equation [1]:

$$
\frac{\partial \omega}{\partial t} = \frac{\partial}{\partial x}\left[K_x(\omega)\frac{\partial H}{\partial x} \right] + \frac{\partial}{\partial y}\left[K_y(\omega)\frac{\partial H}{\partial y} \right]
$$

$$
+ \sum_{m=1}^{M} Q_m(t)\delta(x - x_m, y - y_m), \quad (x, y, t) \in \Omega_0 \times (0, T]. \tag{1}
$$

In this formula H stands for hydraulic head, K is a moisture conductivity function, ω describes humidity, t refers to time, y represents vertical space coordinate taken positive downwards and x represents horizontal space coordinate. This equation can be applied to homogeneous and heterogeneous medium. Our approach is based on S. N. Novoselskiy's explorations [3] on humidity changes because of concentrated sources of water, but in the current research boundary conditions are different:

$$
\begin{aligned}
&\omega|_{x=0} = \omega_0 = const; \quad \omega|_{x=L_1} = \omega_0; \\
&\omega|_{y=0} = \omega_0; \quad \omega|_{y=L_2} = \omega_0; \\
&\omega(x, y, 0) = \omega_0, \quad (x, y) \in \overline{\Omega}_0.
\end{aligned} \tag{2}
$$

Following [3], $K_x(\omega) = k_1 k(\omega)$, $K_y(\omega) = k_2 k(\omega)$ where k_1, k_2 are filtration coefficients around axes Ox, Oy, $k(\omega)$ stands for humidity function for the ground. To make a transition to an equivalent dimensionless equation, some more scaling variables are used:

$$
\beta_2 = 0.5\ell, \quad \beta_1 = \sqrt{\frac{k_2}{k_1}}\beta_2, \quad \alpha = \frac{\langle D_y \rangle \beta_2^2}{T},
$$

$$
\xi = \frac{\beta_1}{L_1}x, \quad \zeta = \frac{\beta_2}{L_2}y, \quad \tau = \alpha t.
$$

Here $\langle D_y \rangle$ is the average value of D_y.

The Kirchhoff's potential, where $D_y(\omega) = K_y(\omega)\frac{d\psi}{d\omega}$ represents diffusion along y axis, will be:

$$\Theta = \frac{4\pi k_1}{Q^* k_2 \beta_2} \int\limits_{\omega_0}^{\omega} D_y(\omega) d\omega,$$

where Q^* is a vector parametric scale multiplier, Ω, Γ are dimensionless analogues to Ω_0, Γ_0.

As a result, we can make a transition to linearized problem

$$\frac{\partial \Theta}{\partial \tau} = \frac{\partial^2 \Theta}{\partial \xi^2} + \frac{\partial^2 \Theta}{\partial \zeta^2} - 2\frac{\partial \Theta}{\partial \zeta}$$

$$+ 4\pi \sum_{m=1}^{M} q_m(\tau)\delta(\xi - \xi_m, \zeta - \zeta_m), \quad (\xi, \zeta, \tau) \in \Omega \times (0, 1]. \qquad (3)$$

The boundary conditions over new axes can be written as:

$$\Theta|_{\xi=0} = 0; \quad \Theta|_{\xi=1} = 0;$$
$$\Theta|_{\zeta=0} = 0; \quad \Theta|_{\zeta=1} = 0; \quad (\xi, \zeta, \tau) \in \Gamma \times [0, 1].$$

At the start of simulation, the initial condition will be:

$$\Theta(\xi, \zeta, 0) = 0, (\xi, \zeta) \in \overline{\Omega}. \qquad (4)$$

To make such transition, several conditions are required:

- The relation between $\theta(\omega)$ and $K_y(\omega)$ is linear:

$$\frac{1}{D_y(\omega)}\frac{dK_y(\omega)}{d\omega} = l = const;$$

- Linearization is achieved by assuming

$$\frac{\partial \omega}{\partial t} = \frac{k_2 \beta_2 Q^*}{4\pi k_1}\frac{1}{D_y(\omega)}\frac{\partial \Theta}{\partial t} \simeq \frac{k_2 \beta_2^3 Q^*}{4\pi k_1}\frac{\partial \Theta}{\partial \tau}.$$

3 Control and Optimization

For the control problem we will use an approach offered in [5] so that we can apply results on existence and uniqueness of the solution [6]. Consider the optimal control belongs to Hilbert space.

Measurements of the target state at the last time step φ are given as a known function. The task is to minimize the deviation between the reached state according to current approximation and the target state.

By determining $0 < \alpha < \min(h^2, \tau)$ as a regularization parameter according to possible measurement errors, the smoothing functional will be:

$$J(q) = \int\limits_{\tilde{\Omega}} (\Theta(\zeta, \xi; q) - \varphi(\zeta, \xi))^2 d\Omega + \alpha q^2 \to \min_q \qquad (4)$$

The iterative variation algorithm proposed in [5] consists of three stages. As for the conjugated operator, it is tested and explained by P. N. Vabishevich [8].

1. Solve the state problem

$$\frac{\partial \Theta^n}{\partial \tau} - \frac{\partial^2 \Theta^n}{\partial \xi^2} - \frac{\partial^2 \Theta^n}{\partial \zeta^2} + 2\frac{\partial \Theta^n}{\partial \zeta} = 4\pi \sum_{m=1}^{M} q_m(\tau)\delta(\xi - \xi_m, \zeta - \zeta_m),$$

$$0 < \tau \le 1, \ \Theta^0 = \Theta(\xi, \zeta, 0) = 0.$$

2. Solve the conjugate problem

$$-\frac{\partial \Psi^n}{\partial \tau} - \frac{\partial^2 \Psi^n}{\partial \xi^2} - \frac{\partial^2 \Psi^n}{\partial \zeta^2} - 2\frac{\partial \Psi^n}{\partial \zeta} = 2\big(\Theta^n - \varphi^n(\tau)\big),$$

$$0 \le \tau < 1, \ \Psi^N = \Psi(\xi, \zeta, 1) = 0.$$

3. Define the new approximation for the optimal control

$$\frac{q_m^{p+1} - q_m^p}{\hat{\tau}_{p+1}} + \Psi_m^0 + \hat{\alpha} q_m^p = 0, \quad p = 0, 1, \dots.$$

4 Solving the Example Task

In the research, a two-dimensional linearized according to [3] problem is solved. The target function values are equal to results of modeling phase with q equal to 1. The initial approximation for source power can be taken near zero or achieved through approximation polynomial from the previous part. In case of the polynomial approximation, all values can be found from initial and boundary conditions, but they depend on current source power approximation.

In case of implicit method and one source, it is possible to gain functional dependence between the source power and the resulting Kirchhoff potential distribution by solving the linear system with source power equal to 1. As a result, the initial

approximation for the source power can be calculated by minimizing the poly from the previous part.

For the direct and conjugate problems explicit scheme leads to the following equation:

$$\frac{\Theta_{ij}^{n+1} - \Theta_{ij}^n}{\tilde{\tau}} = \frac{\Theta_{ij+1}^n - 2\Theta_{ij}^n + \Theta_{ij-1}^n}{h^2} - 2\frac{\Theta_{ij+1}^n - \Theta_{ij-1}^n}{2h}$$

$$+ \frac{\Theta_{i+1j}^n - 2\Theta_{ij}^n + \Theta_{i-1j}^n}{h^2} + 4\pi \tilde{Q}_{ij};$$

$$-\frac{\Psi_{ij}^{n+1} - \Psi_{ij}^n}{\tilde{\tau}} = \frac{\Psi_{ij+1}^{n+1} - 2\Psi_{ij}^{n+1} + \Psi_{ij-1}^{n+1}}{h^2}$$

$$+ 2\frac{\Psi_{ij+1}^{n+1} - \Psi_{ij-1}^{n+1}}{2h} + \frac{\Psi_{i+1j}^{n+1} - 2\Psi_{ij}^{n+1} + \Psi_{i-1j}^{n+1}}{h^2} + 2(\Theta_{ij}^N - \varphi_{ij});$$

$$\Theta_{ij}^n, \Psi_{ij}^n : n = 0, \ldots, N; \quad N = \left[\frac{1}{\tilde{\tau}}\right]; \quad i, j = 0, \ldots, \hat{N}; \quad \hat{N} = \left[\frac{1}{h}\right].$$

Here the first equation starts from the known state at the start $\Theta_{ij}^0 = 0, i, j = 0, \ldots, \hat{N}$. The source power $\tilde{Q}$ depend on the placement of sources, it is constant over time and equal to zero if the source is not there.

The second equation is calculated backwards from the known end state $\Psi_{ij}^N = 0, i, j = 0, \ldots, \hat{N}$.

According to the boundary conditions, humidity (and the Kirhoff potential) values are equal to zero over all the period.

As a result, the implicit scheme leads to defined step-by-step values

$$\Theta_{ij}^{n+1} = \Theta_{ij}^n + \frac{\tilde{\tau}}{h^2}(\Theta_{ij+1}^n + \Theta_{ij-1}^n + \Theta_{i+1j}^n + \Theta_{i-1j}^n - 4\Theta_{ij}^n)$$

$$- \frac{\tilde{\tau}}{h}(\Theta_{ij+1}^n - \Theta_{ij-1}^n) + 4\pi\tilde{\tau}\tilde{Q};$$

$$\Psi_{ij}^n = \Psi_{ij}^{n+1} + \frac{\tilde{\tau}}{h^2}(\Psi_{ij+1}^{n+1} + \Psi_{ij-1}^{n+1} + \Psi_{i+1j}^{n+1} + \Psi_{i-1j}^{n+1} - 4\Psi_{ij}^{n+1})$$

$$+ \frac{\tilde{\tau}}{h}(\Psi_{ij+1}^{n+1} - \Psi_{ij-1}^{n+1}) + 2\tilde{\tau}(\Theta_{ij}^N - \varphi_{ij});$$

$$\Theta_{ij}^n, \Psi_{ij}^n : n = 0, \ldots, N; \quad i, j = 0, \ldots, \hat{N}.$$

For explicit approach, to gain convergence, $\tilde{\tau} = 0.0001$ was chosen.

For the implicit approach, the resulting formulas required solving a system of linear equations. All border values are known, so only central part equations are written here.

$$\frac{\Theta_{ij}^{n+1} - \Theta_{ij}^n}{\tilde{\tau}} = \frac{\Theta_{ij+1}^{n+1} - 2\Theta_{ij}^{n+1} + \Theta_{ij-1}^{n+1}}{h^2} - 2\frac{\Theta_{ij+1}^{n+1} - \Theta_{ij-1}^{n+1}}{2h}$$

$$+ \frac{\Theta_{i+1j}^{n+1} - 2\Theta_{ij}^{n+1} + \Theta_{i-1j}^{n+1}}{h^2} + 4\pi \tilde{Q}_{ij};$$

$$- \frac{\Psi_{ij}^{n+1} - \Psi_{ij}^{n}}{\tilde{\tau}} = \frac{\Psi_{ij+1}^{n} - 2\Psi_{ij}^{n} + \Psi_{ij-1}^{n}}{h^2} + 2\frac{\Psi_{ij+1}^{n} - \Psi_{ij-1}^{n}}{2h}$$

$$+ \frac{\Psi_{i+1j}^{n} - 2\Psi_{ij}^{n} + \Psi_{i-1j}^{n}}{h^2} + 2(\Theta_{ij}^{N} - \varphi_{ij});$$

$$\Theta_{ij}^{n}, \Psi_{ij}^{n} : n = 0..N, \quad i, j = 0..\hat{N}.$$

As a result, the implicit scheme leads to defined step-by-step values

$$\Theta_{ij+1}^{n+1}\left(\frac{1}{h^2} - \frac{1}{h}\right) + \Theta_{ij-1}^{n+1}\left(\frac{1}{h^2} + \frac{1}{h}\right) + \Theta_{i+1j}^{n+1}\left(\frac{1}{h^2}\right) + \Theta_{i-1j}^{n+1}\left(\frac{1}{h^2}\right)$$

$$+ \Theta_{ij}^{n+1}\left(-\frac{4}{h^2} - \frac{1}{\tilde{\tau}}\right) = -\frac{\Theta_{ij}^{n}}{\tilde{\tau}} - 4\pi \tilde{Q}_{ij};$$

$$\Psi_{ij+1}^{n}\left(\frac{1}{h^2} + \frac{1}{h}\right) + \Psi_{ij-1}^{n}\left(\frac{1}{h^2} - \frac{1}{h}\right) + \Psi_{i+1j}^{n}\left(\frac{1}{h^2}\right) + \Psi_{i-1j}^{n}\left(\frac{1}{h^2}\right)$$

$$+ \Psi_{ij}^{n}\left(-\frac{4}{h^2} - \frac{1}{\tilde{\tau}}\right) = -\frac{\Psi_{ij}^{n+1}}{\tilde{\tau}} - 2(\Theta_{ij}^{N} - \varphi_{ij});$$

$$\Theta_{ij}^{n}, \Psi_{ij}^{n} : n = 0\ldots N, i \quad , j = 0, \ldots, \hat{N}.$$

To solve these equations, two approaches were applied—Gaussean direct method and Jacobi iterative method.

According to current results, the deviation is less than 2%, accuracy depends on space and time discretization, possibility to achieve the target function and condition of termination for the three-stage algorithm.

To achieve realistic humidity distribution, the source power was taken relatively low and two cases were modelled—when the sourse is buried between the top border and central part (first two pictures) and central insertion (the following two pictures). The first and third plots represent distribution of Θ achieved by source power approximation in comparison with the desired ones. The second and fourth plots represent isolines of Θ with respect to both dimensionless space coordinates (Figs. 1 and 2).

As we can see, the distribution is symmetric along the horizontal axe, because only the second derivative over the horizontal axe is present in the equation.

For the central source, the following distribution was calculated (Figs. 3 and 4).

According to the last plot, the first derivative by the vertical coordinate results in some asymmetric behavior. The liquid is likely to go down rather than up, similar to real experiments.

For tests with multiple sources let us denote a typical source power as q (tests were made with $q = 10$). The third test demonstrates two sources, buried according to the function:
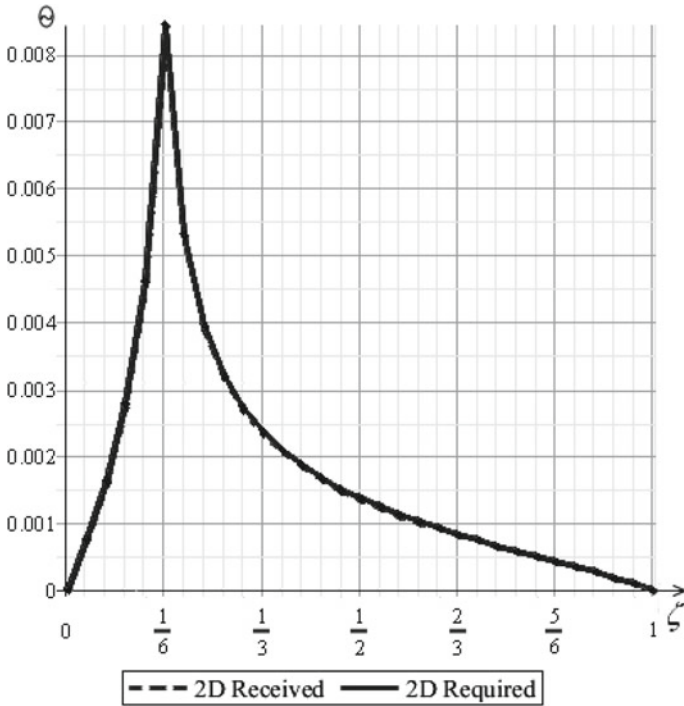
Fig. 1 Kirhhoff potential distribution

$$\tilde{Q}_{ij} = \begin{cases} q, & \left(\xi = \frac{1}{2}, \ \zeta = \frac{7}{30}\right) and \left(\xi = \frac{1}{2}, \ \zeta = \frac{1}{2}\right); \\ 0, & in \ other \ cases, \end{cases}$$

with equal horizontal coordinate. This positioning can be used for humidification with central priority and to increase humidity near the surface (Fig. 5).

Final test was made to demonstrate humidification with three sources, the central source has half of power compared to two others:

$$\tilde{Q}_{ij} = \begin{cases} q, & \left(\xi = \frac{7}{30}, \ \zeta = \frac{7}{30}\right) and \left(\xi = \frac{23}{30}, \ \zeta = \frac{23}{30}\right); \\ \frac{q}{2}, & \left(\xi = \frac{1}{2}, \ \zeta = \frac{1}{2}\right); \\ 0, & else. \end{cases}$$

Results of numerical modelling for this case are demonstrated below (Fig. 6).

Even with different starting optimal source power approximations (zeros for left and central sources and 9 for right source), the iteration process converged with requested accuracy (98%). First, the third source reached power, exceeding optimal, but when powers of two other sources were recalculated many times, all values were back to normal.
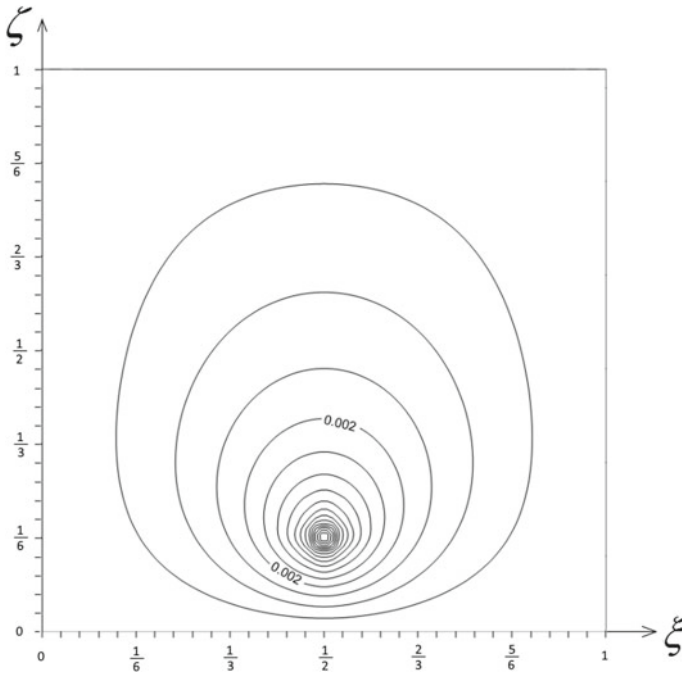
Fig. 2 Isolines for the first test

To sum up, although the method currently has some limitations and can be improved, it results in high accuracy for optimal source power.

5   Programming Details

All calculations were made using C# programming language using Unity 5.6 and Visual Studio (Community edition). The main part of programs consists of three concept parts in a large general cycle.

First of all, all direct and conjugate values are set equal to zero. Then the direct values are calculated from the first time step to the last one. After that, the conjugate equation is solved using received direct values at the last time step from the end of time to its start. After that, new source power value is calculated as well as the average modular difference between the previous reached end state and the current end state for the direct problem.

In the end, calculated optimal source power approximation is shown together with humidity distribution over the area. Received data was used to get visual representation for humidity distribution for humidity in case of one, two and three buried sources.
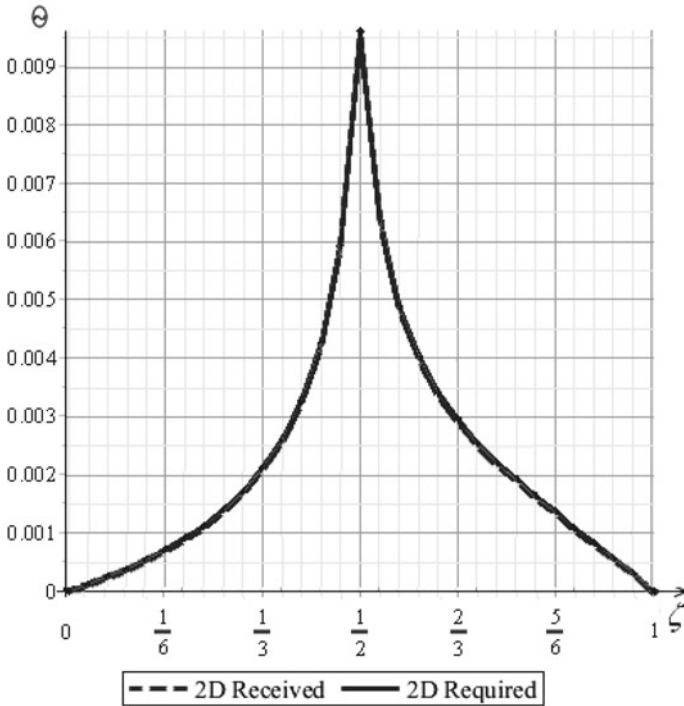
Fig. 3 Kirhhoff potential distribution

6 Connection Between Three-Dimensional and Two-Dimensional Problems

The general linear problem with zero border conditions and depending only on the source power can be formulated the following way:

$$L\Theta = \frac{\partial \Theta}{\partial \tau} - \frac{\partial^2 \Theta}{\partial \xi^2} - \frac{\partial^2 \Theta}{\partial \eta^2} - \frac{\partial^2 \Theta}{\partial \zeta^2} + 2\frac{\partial \Theta}{\partial \zeta}$$

$$= 4\pi \sum_{m=1}^{M} q_m(\tau)\delta(\xi - \xi_m, \eta - \eta_m, \zeta - \zeta_m),$$

$$(\xi, \eta, \zeta, \tau) \in \Omega \times (0, 1].$$

To get rid of the third axe, both sides can be integrated over η two times. As a result, integral equalities are reached:
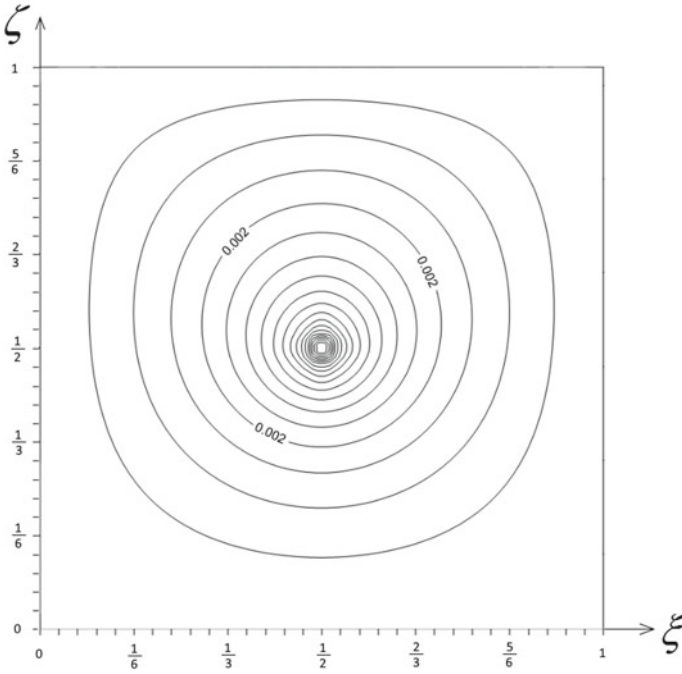
Fig. 4 Isolines for the second test

$$\int_0^1 \left(\frac{\partial\Theta}{\partial\tau} - \frac{\partial^2\Theta}{\partial\xi^2} - \frac{\partial^2\Theta}{\partial\eta^2} - \frac{\partial^2\Theta}{\partial\zeta^2} + 2\frac{\partial\Theta}{\partial\zeta} \right) d\eta$$

$$= \int_0^1 \left(4\pi \sum_{m=1}^M q_m(\tau)\delta(\xi - \xi_m, \eta - \eta_m, \zeta - \zeta_m) \right) d\eta,$$

$$\int_0^1 \left(\frac{\partial\Theta}{\partial\tau} - \frac{\partial^2\Theta}{\partial\xi^2} - \frac{\partial^2\Theta}{\partial\zeta^2} + 2\frac{\partial\Theta}{\partial\zeta} \right) d\eta - \left(\frac{\partial\Theta}{\partial\eta}\bigg|_1 - \frac{\partial\Theta}{\partial\eta}\bigg|_0 \right)$$

$$= 4\pi \sum_{m=1}^M q_m(\tau)\delta(\xi - \xi_m, \zeta - \zeta_m),$$

In case of boundary conditions, described in Discussion the second additive is equal to zero straight from border condition. In our case, we may agree that

$$\Theta|_{\eta=1} = 0, \quad \Theta|_{\eta=0} = 0 \Rightarrow -\left(\frac{\partial\Theta}{\partial\eta}\bigg|_1 - \frac{\partial\Theta}{\partial\eta}\bigg|_0 \right) = 0.$$
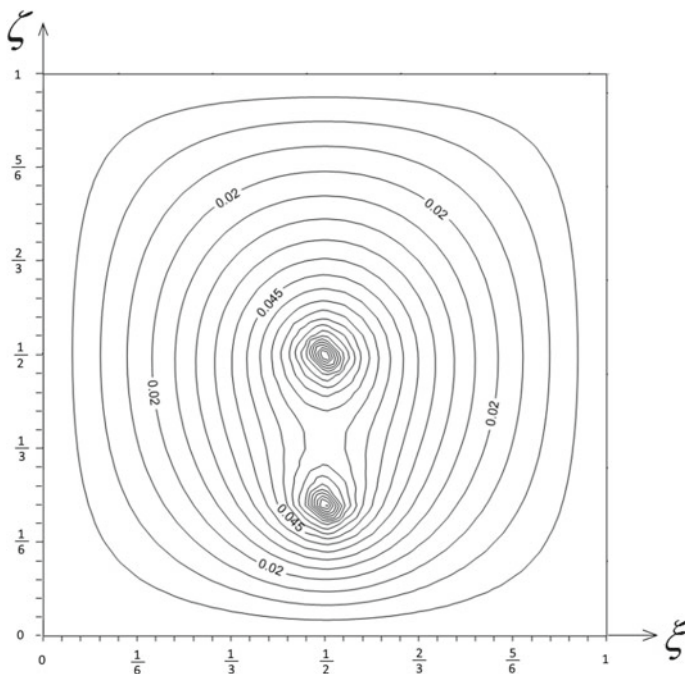
Fig. 5 Isolines for the third test

7 Discussion

The current article is based on cases, when zero values on all borders are achieved or almost achieved, because for this case the existence and uniqueness of optimal control is proven. But in fact, the same accuracy was achieved for more realistic boundary conditions, when derivatives are equal to zero:

$$\xi = 0, \quad \frac{\partial \Theta}{\partial \xi} = 0, \quad \xi = 1, \quad \frac{\partial \Theta}{\partial \xi} = 0;$$

$$\zeta = 0, \quad \frac{\partial \Theta}{\partial \zeta} = 0, \quad \zeta = 1, \quad \frac{\partial \Theta}{\partial \zeta} = 0;$$

$$(\xi, \zeta, \tau) \in \Gamma \times [0, 1].$$

To take into account all the boundary conditions, we refer to [10]. Using integral-interpolation method, the following system of equations, using numeric analogues for derivatives. To get better accuracy, equations for the border lines were taken according to central equation for one coordinate and border condition for another one.
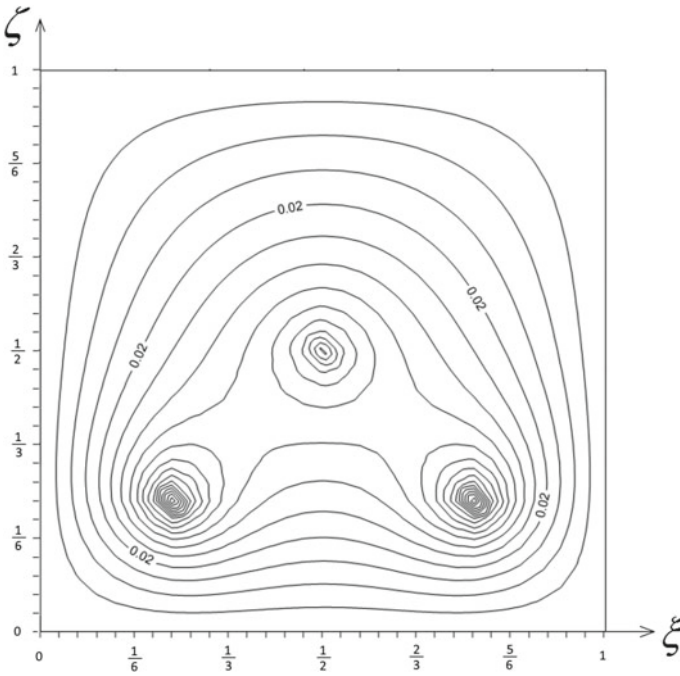
Fig. 6 Isolines for the fourth test

Both explicit and implicit schemes were tested and compared for the direct and conjugate problems. For explicit scheme 10,000 time steps were used to gain convergence, while horizontal and vertical lines were divided into 30 parts (also tested on 20). But for solving the matrix equation Gaussian method was used to get exact values and Jacobi method as an alternative.

Another aspect to discuss is the usage of delta-function representing the source power. As an alternative, linear source with small size can be used. As a result, the right side in case of central source will change into $4\pi F(\xi, \zeta)$,

$$F(\xi, \zeta) = \begin{cases} \frac{1}{l_1} \cdot \frac{1}{l_2} \cdot q, \ \xi = 0, 5, \ \zeta = 0, 5; \\ 0, \qquad\qquad else. \end{cases}$$

To achieve similar result during discrtization, we may choose convenient values for the source size

$$0 < l_1 << L_1, 0 < l_2 << L_2; l_1 l_2 < h^2.$$

As a result, the integrated values in source point can be the same as with delta-function.

At the same time, integrated values over other points are equal to zero.

To sum up, the offered combination allows to solve problems, if boundary and transitional conditions are fulfilled. But the real potential of this approach seems wider, so the current research will continue.

8   Conclusion

The proposed algorithm can be applied for nonlinear problems with respect to transition limitations and buried sources with low or zero humidity values on borders. It is possible to widen the applicability by proving existence and uniqueness of solutions for other boundary conditions.

The general idea of the combination is to apply Kirchhoff transformation and scale the space and time values leading to the new quasilinear equation. Then variation algorithm is applied to get the optimal source power. It consists of direct problem, conjugate problem and the new approximation step.

To solve the algorithm's problems explicit or implicit approach can be applied, but the amount of time steps is usually larger for the explicit one.

As a result, the approximated value of optimal source power had more than 98% accuracy for several tests of buried sources.

References

1. Pullan AJ (1990) the quasilinear approximation for unsaturated porous media flow. Water Resour Res 26(6):1219–1234
2. Van Genuchten MTH (1980) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci Soc Am J 44:892–898
3. Novoselskiy SN (1981) Solution of some boundary value problems of moisture transport with irrigation sources (in Russian). Dissertation for the academic degree of a candidate in Mathematics and Physics, 01.02.05, Kalinin Polytechnic Institute, Kalinin
4. List F, Radu F. A Study on Iterative Methods for Richards' Equation. Math.NA. http://arxiv.org/abs/1507.07837
5. Lyashko S, Klyushin D, Semenov V, Shevchenko K (2006) Identification of point contamination source in ground water. Int J Ecol Develop 5(F06):36–43
6. Lyashko S (1998) Generalized control of linear systems (in Russian). Scientific Thought, Kyiv
7. Kirk DE (2004) Optimal control theory. An introduction. Dover Books on Electrical Engineering
8. Vabishchevich PN (2003) Numerical solution of the problem of the identification of the right-hand side of a parabolic equation (in Russian). Russian Math. (Iz. VUZ) 47:1
9. Samarskiy AA (1989) Theory of difference schemes. Nauka, Moscow (in Russian)
10. Samarsky AA, Nikolaev ES (1978) Methods of solving grid equations. Science, Moscow, p 588 (in Russian)

Mathematical Modeling and Simulation of Systems in Manufacturing

Modeling of Molding of Composite Products on Prepreg Basis

Andrii Kondratiev ⓘ**, Oleksandr Haidachuk** ⓘ**, Anton Tsaritsynskyi** ⓘ**, and Tetyana Nabokina** ⓘ

Abstract The unique set of properties of polymeric composite materials offers many opportunities for improvement of the existing designs and technological processes and development of new ones. The main purpose of the composite product molding process is to establish the optimal relationship between the process parameters, which provide specified quality of the product at minimum energy costs. In this context, there is a need for modeling of the molding process, which would describe physical and chemical processes occurring in the molded polymeric composite material. New method for studying the main parameters of the composite product molding process on prepreg basis has been developed. This method allows calculating theoretically the optimal temperature and time diagram and dependence of the molding pressure on the time for specific binders and reinforcing materials. It is offered to divide the binder viscosity into physical and chemical components. The physical component excludes the presence of chemical transformations and depends on the temperature only. The chemical component is associated with chemical transformations; it is a function of the temperature and time. The method for experimental determination of the minimum total viscosity of the binder, as well as degree of polymerization, from the temperature and time, is proposed. The mathematical model of filling with a binder of specified viscosity through inter-fiber space of the reinforcing material has been developed. Analytical dependence of the molding pressure on the viscosity, geometric parameters of the molded product, prepreg, monolayer of the polymeric composite material and molding time is obtained. The results allow optimizing the main stages of the temperature and time diagram of the composite product molding cycle, which will provide minimization of the energy costs.

A. Kondratiev (✉)
O.M. Beketov National University of Urban Economy in Kharkiv, 17, Marshal Bazhanov St., Kharkiv 61002, Ukraine
e-mail: andrii.kondratiev@kname.edu.ua

O. Haidachuk
Ningbo University of Technology, 201 Fenghua Rd., Jiangbei Qu, Ningbo Shi, Zhejiang Sheng, China

A. Tsaritsynskyi · T. Nabokina
National Aerospace University "Kharkiv Aviation Institute", 17, Chkalova St,, Kharkiv 61070, Ukraine

1 Introduction

Modernization and development of new structures used in aircraft construction, rocket and space technology, power sector, construction industry, mechanical engineering and other industries are associated with the wider use of polymeric composite materials (PCM) [1, 2]. Properties and features of these materials are different compared to those of the traditional structural materials. The unique set of properties of PCM offers many opportunities for improvement of the existing designs and technological processes and development of new ones [3, 4].

As is commonly known [1, 5], technological process of PCM product molding consists in giving it non-reversible shape with the use of shape-generating molding tools through polymerization of the binder at the certain temperature t and pressure p, varying in time τ. Parameters t, p and τ or $t(\tau)$ and $p(\tau)$ are the main external parameters of the PCM product molding process [3, 5]. Internal parameters of the process are dynamic viscosity of the binder $\mu(t, \tau)$, as well as volume content of the binder θ_b and its structural characteristics [5, 6].

Since the internal parameters are predetermined by the chosen binder and reinforcing material based on the operational characteristics of the product, objective of the molding process is to establish a relationship between them and the variable external parameters which provides the specified quality of the product at minimal energy costs and meeting of the requirements of the safe production activity.

Therefore, there is a need for modeling of the molding process, which would describe physical and chemical processes occurring in the molded PCM.

2 Literature Review

In a number of industries, PCM structures made by layup methods prevail. Among them, the highest quality of products with acceptable energy and labor costs is provided by the use of prepregs [6]. However, shaping of products made of PCM by winding-up methods, which is characteristic of the structures of oil and gas industry and space technology, actually does not exclude the shaping described by the ratio of the same external parameters $t(\tau)$ and $p(\tau)$ as well. Temperature and time conditions of shaping depends on many factors, the main of which are listed below [5, 6]:

- binder type and composition;
- method of impregnation of reinforcing material: preliminary preparation of prepregs followed by molding of the product ("dry" molding method), "wet" method which consists in impregnation of the reinforcing material in the process of forming and subsequent molding of the product;

- type of reinforcing material and its preliminary preparation method;
- molding method.

A large number of authors [6, 7] were studying these factors. In most cases, the results obtained represented either a generalized solution scheme, or models taking into account a few factors of the molding process only [1, 8]. For example, in the process of work with prepregs, the issues of optimal operating conditions are solved in order to ensure their specified quality [7, 9]. At the same time, conditions of technological processes are ambiguously harmonized with each other. The works [10, 11] describe the chemical and physical processes occurring in the molded PCM during the curing process. Problems arising in the molding process are considered, chemical transformations in the binder are modeled, and release of volatile products is normalized. The molding process is modeled by only one component (the binder) in [12–14]. In this case, dimensions of the structure, heating conditions and presence of the reinforcing material are not taken into account. The paper [15] deals with molding parameters according to conditions of the safe production activity, which are determined by the volume of emissions of harmful volatiles and are limited by the fire and explosion safety requirements [16]. Experimental studies [17, 18] show that at the stage of PCM heating, along with shrinkage, temperature phenomena are also involved in the occurrence of the stressed state. Temperature stresses become comparable with shrinkage stresses after reaching the material viscosity corresponding to 60–70% conversion in the binder. However, these papers deals with the heating stage only, and the change in properties of the materials from temperature is not taken into account. The paper [19] presents the concept of creation of energy-saving technologies for PCM manufacturing by optimization of the modes of impregnation and heat treatment. The concept is based on obtaining a picture of the change in the PCM physical state: reaching of the minimum viscosity.

Almost all works except [10, 19] do not pay due attention to the determination of the molding pressure, while this parameter is among the main ones. Molding pressure regulates the volume content of the reinforcing material and provides the uniformity, degassing, and possible reduction in the residual stressed state [10, 19].

Models of molding proposed in [20, 21] are of particular interest. Authors of these works propose both solving of the problem of choosing the rational technological parameters of the molding conditions and controlling setup parameters, and, if necessary, making adjustments during the molding process. The model proposed by [21] is further developed in [22] for the problem of two approaching surfaces which squeeze out a liquid with a viscosity. However, this study uses as an external force the force concentrated at the origin point instead of that distributed along the length of the plate. This assumption caused its underestimation by 1.33 times. In addition, the characteristic dimension is unreasonably interpreted in [22].

The paper [23] develops the mathematical model for laying up of prepreg of variable width for the given condition of its molding. The paper [24] includes the experimental study of the operating conditions of PCM molding. The studies allowed increasing the strength of the molded PCM by 25–35%. However, the results are valid only for a narrow class of fluoroplastic materials and quartz fabric. The paper [25]

presents the data of analysis of topical issues of molding for the basic technological cycle of reinforced products molding of the compositions of epoxy polymers. The expediency of using the ultrasonic modification to achieve energy saving and to improve the quality of resulting composite products is substantiated. The paper [26] develops an approach to the application of computer-aided design principles for selection of the optimal design and technology parameters in PCM molding with the use of ultrasonic treatment.

It can be seen from the above review that the phenomenon of occurrence and influence of temperature stresses at the stage of heating of the material and the presence of reinforcing material in the product is not taken into account in most works; process parameters are determined by the physical and mechanical characteristics of the binder only. This approach to determination of process parameters often leads to structural disturbances, occurrence of unacceptable stress–strain behavior in the material and product, and additional costs in the manufacturing of structures of PCM.

The purpose of this work is to develop a method for studying the main parameters of the process of molding of products of PCM on prepreg basis, which would theoretically calculate the optimal temperature and time diagram and dependence of the molding pressure on the time for specific binders and reinforcing materials.

3 Research Methodology

As mentioned above, temperature and time conditions of molding depend on many factors. Each of these factors on which the temperature and time conditions of PCM molding depend is an integrated factor. It includes a number of components, the combination of which forms the integrated factor leading to degeneration of one or another section of the temperature and time diagram. The diagram of the temperature and time conditions of PCM product molding in the general case is of the form shown in Fig. 1.

For example, the use of cold-set binder in the manual layer-by-layer impregnation of the reinforcing material during formation and subsequent free molding without excessive pressure degenerates the diagram in Fig. 1 into straight line $t_0 = $ const at τ

Fig. 1 Typical diagram of the temperature and time conditions of PCM product molding
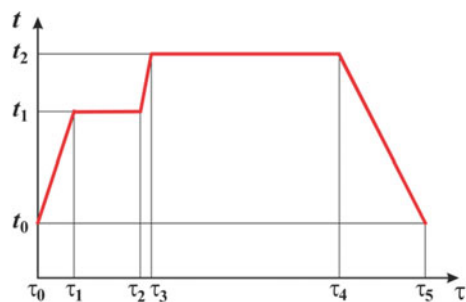
Fig. 2 Typical diagram of pressure supply during molding of PCM products
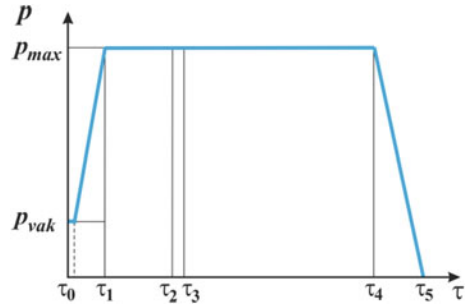
$= \tau_5$. With the use of prepreg technology and a very small amount of gas evolution of the binder components in the polymerization process the area τ_1–τ_2 etc. degenerate [5].

The temperature and time conditions are accompanied by the certain regime of excess pressure p (Fig. 2).

In this case, the mode of supply of excess pressure also depends on the integrated factors mentioned above. For example, with the use of cold-set binder in the manual layer-by-layer impregnation of the reinforcing material during formation and subsequent free molding, the diagram in Fig. 2 degenerates into the line $p = 0$, $\tau = \tau_5$ or $p = p_{vak}$, $\tau = \tau_5$. For molding of the pack, $p = p_{max}$, $\tau = \tau_4$ etc. [5].

Now let's consider the general case characterized by the presence of the diagrams in Figs. 1 and 2 with the most common practical method of molding of PCM product which is preliminarily formed of prepreg [6, 8]. In this case, content of the solvent in "dry" prepreg provided it is made with high quality is 0–3% [5, 27], and it can be neglected. Presence of the area τ_1–t_1 is associated with the ability of the equipment (oven, autoclave) to provide the maximum possible rate of temperature rise in order to "soften" the binder in the prepreg pack to its minimum viscosity μ_{min}, required for its uniform spreading in the volume of the formed pack.

Viscosity of the binder depends on its temperature and degree of chemical transformations (polymerization or poly-condensation) occurring in it. In the absence of chemical transformations, viscosity has a physical nature μ_{ph} [28], decreasing exponentially with increase in the temperature [29].

However, chemical transformations in the binder begin during its preparation, when a hardener is introduced into the resin. The prepreg binder already contains the specified "dry" residue (1–3%), i.e. weight content of the cured binder [28]. When the pack is heated, intensity of chemical transformations rises and, as a result, the viscosity μ_x associated with this process increases (Fig. 3).

True viscosity in the process of molding $\mu(t, \tau)$ is the sum of the physical $\mu_{ph}(t)$ and chemical $\mu_x(t, \tau)$ components:

$$\mu(t, \tau) = \mu_{ph}(t) + \mu_x(t, \tau). \tag{1}$$

Fig. 3 Temperature
dependence of dynamic
viscosity of the binder at the
time point $\tau = \text{const}$ from
the beginning of the molding
process
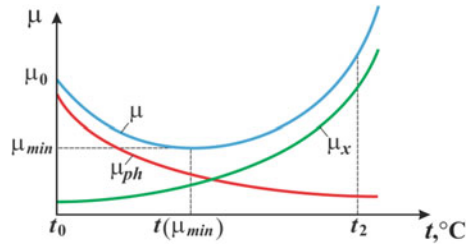
Analytical definition $\mu(t, \tau)$ for each specific compound presents significant difficulties because of complexity of calculating the process of chemical transformations at a certain temperature and also its dependence on the time τ [28].

For this reason, $\mu_{\min}$ should be determined experimentally [27, 28]. As a result of experimental studies of the dependence $\mu(t, \tau_i = \text{const})$ of a specific binder at different rates of temperature rise, we establish the vector of experimental values.

$$\mu^{\text{ex}} = \left\{ \begin{array}{c} \mu_1^{\text{ex}} \\ \ldots \\ \mu_n^{\text{ex}} \end{array} \right\} \text{ for the matrix of arguments } X = \left\{ \begin{array}{cc} t_1 & \tau_1 \\ \ldots & \ldots \\ t_n & \tau_n \end{array} \right\}.$$

We use the linear mathematical model containing m parameters:

$$\mu = \sum_{i=1}^{m} a_i f_i(X), \tag{2}$$

where $f_i(X)$—approximating functions.

Parameters of the model a_i can be determined from the condition of minimum standard deviation of the model function μ from the given experimental values μ^{ex} [30]:

$$S = \sum_{j=1}^{n} \left(\mu_j^{ex} - \sum_{i=1}^{m} a_i f_i(X) \right)^2 \rightarrow \min. \tag{3}$$

The minimum condition (3) is:

$$\frac{\partial S}{\partial a_k} = 2 \left[\sum_{j=1}^{n} \mu_j^{ex} - \sum_{j=1}^{n} \sum_{i=1}^{m} a_i f_i(X) \right] f_k(X) = 0. \tag{4}$$

The condition (4) represents a system of linear equations and can be written as:

$$F^T F\{a\} = F^T \{\mu^{ex}\}, \tag{5}$$

where F—model matrix

$$F = \begin{vmatrix} f_1(X_1) & f_2(X_1) & \ldots & f_m(X_1) \\ f_1(X_2) & f_2(X_2) & \ldots & f_m(X_2) \\ \ldots & \ldots & \ldots & \ldots \\ f_1(X_n) & f_2(X_n) & \ldots & f_m(X_n) \end{vmatrix}. \tag{6}$$

The solution to system (5) is the vector of parameters of the mathematical model (2):

$$\{a\} = \left(F^T F\right)^{-1} F^T \left\{\mu^{ex}\right\}. \tag{7}$$

In addition to the experimental determination of $\mu(t, \tau)$, it is necessary to conduct a series of experiments to determine the degree of polymerization of the binder $C(t, \tau)$.

The technique of approximation of $C(t, \tau)$ using the analytical dependence is similar to that described by formulas (2)–(7). When we have the analytical dependence of $C(t, \tau)$, the degree of polymerization in each section of the diagram can be determined by integrating it over τ, taking into account that:

$$C(t, \tau) = \varphi(t, \tau) = \varphi(\psi(\tau), \tau), \tag{8}$$

where $\psi_1(\tau) = \frac{t_1-t_0}{\tau_1}\tau + t_0$, $\psi_2(\tau) = t_1$, $\psi_3(\tau) = \frac{(t_3-t_1)\tau+(t_1\tau_3-t_3\tau_2)}{\tau_3-\tau_2}$, $\psi_4(\tau) = t_3$.

After that:

$$C_I = \int\limits_0^{\tau_1} \varphi(\psi_1(\tau), \tau)d\tau; \quad C_{II} = \int\limits_{\tau_1}^{\tau_2} \varphi(\psi_2(\tau), \tau)d\tau;$$

$$C_{III} = \int\limits_{\tau_2}^{\tau_{31}} \varphi(\psi_3(\tau), \tau)d\tau; \quad C_{IV} = \int\limits_{\tau_3}^{\tau_4} \varphi(\psi_4(\tau), \tau)d\tau. \tag{9}$$

Since τ_4 is not known, it can be determined from the condition:

$$C_{IV} = 1 - C_I - C_{II} - C_{III}. \tag{10}$$

Thus, the duration of all time intervals of the temperature and time diagram in Fig. 1 is fully defined.

The relationship between the molding pressure p and other process parameters is found using the mathematical model [21, 22] for the problem of two approaching surfaces squeezing out a liquid with the viscosity μ (Fig. 4). The liquid being squeezed is located between plates $-l/2 < x < l/2, y = \pm h/2$, which are compressed by equal and oppositely directed forces parallel to y-axis when there is no slip at the boundaries $y = \pm h/2$ (Fig. 4).

The relevance of interpretation of the considered problem by the model [21, 22] (Fig. 4) follows from the considerations below. Prepreg of which the PCM product
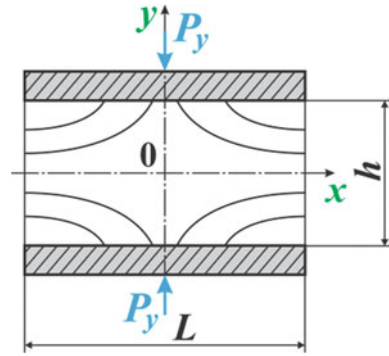
is formed, depending on the type of reinforcing material, is filled with the binder, as
shown in Fig. 5.

In the process of molding, the plates (rigid coupling plate) 5 under the action of
external pressure p will move at a certain speed v_0 until all voids 4 are filled with
binder 3. It corresponds to the change in height h from prepreg thickness h_{pr} to PCM
monolayer thickness h_{CM}.

Strictly speaking, during the molding process the binder in the prepreg repre-
sents the Kelvin-Voigt hard-viscous medium, containing an elastic component in the
physical law [21, 22]:

$$\mu(t, \tau)\dot{\varepsilon} + E_b(t, \tau)\varepsilon = \sigma(t, \tau), \tag{11}$$

where ε and $\dot{\varepsilon}$—deformation and rate of its change from $\tau(t)$; $E_b(t, \tau)$—instantaneous
modulus of elasticity of binder in the prepreg.

Then we estimate the contribution of the elastic component of Eq. (11).

Deformation of the prepreg is:
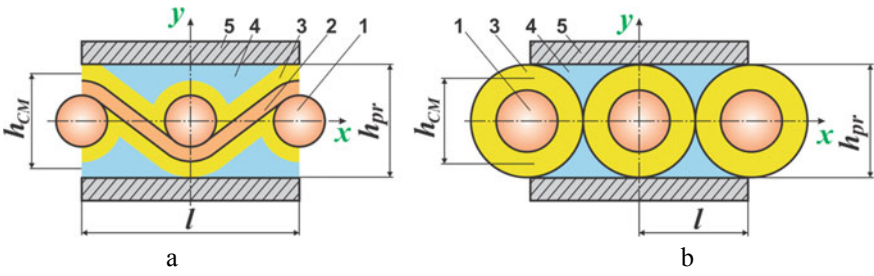
1 – transverse fibers; 2 – longitudinal fibers (pick); 3 – binder layer; 4 – voids;
5 – coupling plate

Fig. 5 Structure of prepreg in PCM pack: **a** prepreg based on woven fabric; **b** prepreg based on
unidirectional tape

$$\varepsilon = \frac{h_{pr} - y}{h_{pr}} = 1 - \frac{y}{h_{pr}}, \tag{12}$$

where y varies with time in the range of $0 \le y \le h_{CM}$.

Since the speed of convergence of the plates v_0 is constant, the deformation ε changes linearly, and then

$$\dot{\varepsilon} = \left(1 - \frac{h_{CM}}{h_{pr}}\right) \frac{1}{\tau_2 - \tau_1}. \tag{13}$$

Substituting (13) into (11), we get:

$$\mu(t, \tau)\left(1 - \frac{h_{CM}}{h_{pr}}\right) \frac{1}{\tau_2 - \tau_1} + E_b(t, \tau)\left(1 - \frac{h_{CM}}{h_{pr}}\right) \frac{1}{\tau_2 - \tau_1} = \sigma(t, \tau). \tag{14}$$

Equation (14) allows estimating the contribution of the second term for the known $E_b(t, \tau)$. To analyze the contribution of $E_b(t, \tau)$ into the Eq. (14), we use the dependences for $E_{pr}(t, \tau)$ and E_{CM}:

$$E_{pr}(t, \tau) = E_f(t, \tau)(1 - \theta_b) + E_{b\,pr}(t, \tau)\theta_b; \tag{15}$$

$$E_{CM}(t, \tau) = E_f(t, \tau)(1 - \theta_b) + E_{bCM}(t, \tau)\theta_b. \tag{16}$$

The modulus of elasticity of the cured PCM within a few hours (i.e. within the duration of the molding cycle) does not depend on τ and decreases slightly at $t \le t_1$.

Since the modulus of elasticity of the reinforcing material $E_f(t, \tau)$ and its volume content θ_b are the same in PCM and prepreg, from (16) it follows that:

$$E_f(t, \tau)(1 - \theta_b) = E_{CM}(t, \tau) - E_{bCM}(t, \tau)\theta_b. \tag{17}$$

Substituting (17) into (16), we get:

$$E_{b\,pr}(t, \tau) = \frac{E_{pr}(t, \tau) + E_{bCM}(t, \tau)\theta_b - E_{CM}(t, \tau)}{\theta_b}. \tag{18}$$

The process of convergence of plates in the model [21, 22] is accompanied by the occurrence of normal stresses in the liquid:

$$\sigma_x = 3\mu \frac{v_0\left[3\left(\frac{h^2}{4} - y^2\right) + x^2 - \frac{L^2}{4}\right]}{2h^3}; \quad \sigma_y = 12\mu \frac{v_0\left[y^2 - \frac{h^2}{4} + x^2 - \frac{L^2}{4}\right]}{h^3}, \tag{19}$$

where L—maximum longitudinal dimension of molded product; h—gap between the plates: $h = (h_{pr} - h_{CM})/2$.

Force per unit length P_{ypr}, applied to the plates at $y = h/2$ is equal to:

$$P_{y\,pr} = -2 \int_{0}^{L/2} \sigma_y dx = \frac{2\mu v_0 L^3}{h^3}. \tag{20}$$

Then the initial pressure acting on the prepreg shall be:

$$p_{pr} = \frac{2\mu v_0 L^2}{h^3}. \tag{21}$$

Speed of convergence of plates included in (17)–(21) can be defined as:

$$v_0 = \frac{h}{\tau}, \tag{22}$$

where τ—time during which the binder fills all the voids in the pack.

When the voids in the pack are completely filled, its thickness h will change from h_{pr} to h_{CM}. Then v_0 is

$$v_0 = \frac{(h_{pr} - h_{CM})}{2\tau}. \tag{23}$$

Taking into account (21), pressure on the pack is equal to:

$$p_{\max} = \frac{8\mu(h_{pr} - h_{CM})L^2}{\tau(h_{pr} - h_{CM})^3} = \frac{8\mu L^2}{\tau(h_{pr} - h_{CM})^2}. \tag{24}$$

Parameter l, characterizing the density of fibers' location in the reinforcing material, can be expressed through the volume content of the binder in PCM θ_b and thickness of its monolayer:

$$\overline{V}_b = lh_{CM} - \frac{\pi h_{CM}^2}{4}; \tag{25}$$

$$\theta_b = \frac{\overline{V}_b}{\overline{V}_{CM}}; \overline{V}_{CM} = lh_{CM}, \tag{26}$$

where $\overline{V}_b$, $\overline{V}_{CM}$—unit volumes of the binder and PCM at length l.

Dividing (25) by $\overline{V}_{CM}$, we obtain, taking into account (26):

$$l = \frac{\pi}{4} \frac{h_{CM}}{(1 - \theta_b)}. \tag{27}$$

The designer developing a product of PCM is focused on setting the minimum possible volume content of the binder $\theta_{b\min}$ in order to obtain high physical and mechanical properties of the material [5]. However, it is also necessary to proceed

Fig. 6 Model of filling inter-fiber space with the binder in the structure of monolayers in the process of molding
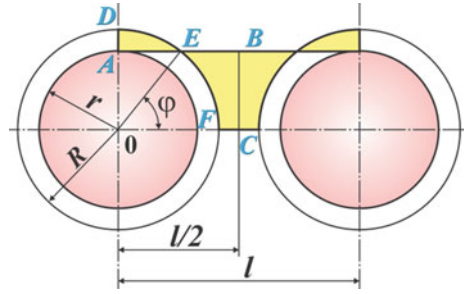
from a specific reinforcing material with certain density of fibers l and their diameter d_f (in case of the tape of $h_{CM} = 2d_f$).

Therefore, formula (27) is the initial one in setting of θ_b for the process engineer, being transformed to the form below:

$$\theta_b = \frac{\pi}{4} \frac{1 - d_f}{2l}. \tag{28}$$

The objective of the process engineer is to provide θ_b by the corresponding amount of the binder applied in the process of prepreg manufacturing.

This amount is determined by the difference between the thicknesses of the prepreg h_{pr} and PCM monolayer h_{CM}.

Consequently, the binder solution (composition and viscosity) should be chosen so that by the time when the product is molded of the finished prepregs the monolayer structure looks like that shown in Fig. 6, where r—radius of the fiber of the reinforcing material, R—radius of the fiber with binder in the prepreg $(r = h_{CM}/2, \ R = h_{pr}/2)$. Radius R should be such that as a result of molding the binder, which is initially retained in the ADE half-segment, overflows and completely fills the unit volume S_{FEBC}.

At $S_{ADE} > S_{FEBC}$ the binder film is formed on the surface of PCM monolayer, i.e. $\theta_b > [\theta_b]$; at $S_{ADE} < S_{FEBC}$ there are voids formed in the PCM structure between monolayers, since $\theta_b < [\theta_b]$.

Therefore, the equation $S_{ADE} = S_{FEBC}$ allows determining the analytical dependence between r and R (or h_{pr} and h_{CM}) at the given l, θ_b and r (i.e. h_{CM}).

From Fig. 6 it follows that:

$$S_{FEBC} = S_{OABC} - S_{OEF} - S_{OAE}; \quad S_{OABC} = \frac{r}{2}\sqrt{R^2 - r^2};$$

$$S_{OEF} = \frac{R^2}{2\pi} \arcsin \frac{r}{R}; \ S_{ADE} = S_{ODE} - S_{OAE};$$

$$S_{ODE} = \frac{\left(\frac{\pi}{2} - \varphi\right)R^2}{2\pi} = \frac{R^2}{4} - \frac{R^2}{4} \arcsin \frac{r}{R};$$

from which

$$S_{ADE} = \frac{R^2}{4} - \frac{R^2}{2\pi} \arcsin \frac{r}{R} - \frac{r}{2}\sqrt{R^2 - r^2}; \tag{29}$$

$$S_{FEBC} = \frac{lr}{2} - \frac{R^2}{2\pi} \arcsin \frac{r}{R} - \frac{r}{2}\sqrt{R^2 - r^2}. \tag{30}$$

Equating (29) and (30), we get:

$$R = \sqrt{2lr} \tag{31}$$

or taking into account (27) and the fact that $2R = h_{pr}$, and $2r = h_{CM}$

$$h_{pr} = h_{CM}\sqrt{\frac{\pi}{1 - \theta_b}}. \tag{32}$$

Now the formula (24) can be transformed taking into account (32) to the final form:

$$p_{\max} = \frac{8\mu L^2}{\tau \left(1 - \sqrt{\frac{\pi}{1-\theta_b}}\right)^2 h_{CM}^2}. \tag{33}$$

4 Results

Thus, the above method for studying the main parameters of the process for molding of PCM products on prepreg basis allowed calculating theoretically the optimal temperature and time diagram (Fig. 1) and dependence of the molding pressure on the time (Fig. 2) for specific binders and reinforcing materials in the cooling mode.

For example, for the prepreg based on T-10-80 glass cloth and EDT 69N binder $E_{CM}(t_0) = 1.8$ GPa; $E_{pr} = 1.5$ GPa; $E_{bCM} = 0.8$ GPa. Then, at $\theta_b = 0.4$ $E_{pr}(t_0) = 8.05$ GPa, i.e. in the entire time interval $0 < \tau < \tau_1$ the second term in (11) prevails over the first one. It indicates the absence of fluidity of the binder at the first stage of molding and, consequently, impossibility of filling the pores in the prepreg. At the second stage $\tau_1 < \tau \leq \tau_2$ (about 30 min or more) $E_{bpr}(t, \tau) \to 0$, and it is relevant to consider the molding process in the framework of the model of the Newtonian ideally viscous medium [21, 22].

Pressure p_x, arising simultaneously with $p_y = p_{\max}$, which acts along the horizontal channels, promotes more complete removal of air bubbles and filling of voids with the binder. This pressure is determined similarly to p_y by integration of σ_x over the height of the compressed pack, comprising n layers of the prepreg, taking into account (19) and (32):

$$p_x = \frac{24\mu(1-\theta_b)}{\pi\tau}\left[\frac{\left(\sqrt{\frac{\pi}{1-\theta_b}}\right)^2}{2} - 4\frac{x^2}{h_{CM}^2} - \frac{L^2}{h_{CM}^2}\right]. \tag{34}$$

At $x = 0$

$$p_{x=0} = \frac{24\mu(1-\theta_b)}{\pi\tau}\left[\frac{\left(\sqrt{\frac{\pi}{1-\theta_b}}\right)^2}{2} - \frac{L^2}{h_{CM}^2}\right]. \tag{35}$$

The largest value of p at $x = L/2$ is equal to

$$p_{x=\frac{L}{2}} = \frac{12\mu(1-\theta_b)}{\pi\tau}\left(\sqrt{\frac{\pi}{1-\theta_b}}\right)^2. \tag{36}$$

The ratio $p_y/p_{x=\frac{L}{2}}$ is equal to:

$$\frac{p_y}{p_{x=\frac{L}{2}}} = \frac{2\pi L^2}{3(1-\theta_b)h_{CM}^2\left(1-\sqrt{\frac{\pi}{1-\theta_b}}\right)^4}. \tag{37}$$

Since $L \gg h_{CM}$, the ratio $p_y/p_{x=\frac{L}{2}}$ is always less than unity. It indicates that at the autoclave pressure p_y determined according to (33) the pressure p_x along the horizontal channels in the reinforcing material is always higher, which ensures high quality of molding.

5 Conclusions

New method for studying the main parameters of the process of molding of PCM products on prepreg basis has been developed. This method allows for the first time calculating theoretically the optimal temperature and time diagram and dependence of the molding pressure on the time for specific binders and reinforcing materials. The method includes key components detailed below:

- division of the binder viscosity into the physical component excluding the presence of chemical transformations and depending on the temperature only and chemical component associated with such transformations and being a function of the temperature and time;
- methods for the experimental determination of the minimum total viscosity of the binder, as well as the degree of polymerization from the temperature and time;
- mathematical model of filling with a binder with the viscosity $\mu(t, \tau)$ through inter-fiber space of the reinforcing material, which is implemented by the analytical

dependence of the molding pressure on the viscosity, geometric parameters of the molded product, prepreg, PCM monolayer, and molding time.

References

1. Kondratiev A, Gaidachuk V, Nabokina T, Kovalenko V (2019) Determination of the influence of deflections in the thickness of a composite material on its physical and mechanical properties with a local damage to its wholeness. Eastern-European J Enterp Technol 4(1(100)):6–13. https://doi.org/10.15587/1729-4061.2019.174025
2. Fomin O, Logvinenko O, Burlutsky O, Rybin A (2018) Scientific substantiation of thermal leveling for deformations in the car structure. Int J Eng Technol 7(4.3):125–129. https://doi.org/10.14419/ijet.v7i4.3.19721
3. Kondratiev, A.V., Kovalenko, V.O.: Optimization of design parameters of the main composite fairing of the launch vehicle under simultaneous force and thermal loading. Space science and technology kosmicna nauka i tehnologia 25 (119):3–21. https://doi.org/10.15407/knit2019.04.003
4. Rodichev YM, Smetankina NV, Shupikov OM, Ugrimov SV (2018) Stress-strain assessment for laminated aircraft cockpit windows at static and dynamic load. Strength Mater 50(6):868–873. https://doi.org/10.1007/s11223-019-00033-4
5. Campbell FC (2004) Manufacturing processes for advanced composites. Elsevier, Amsterdam
6. Budelmann D, Schmidt C, Meiners D (2020) Prepreg tack: a review of mechanisms, measurement, and manufacturing implication. Polym Compos 41(9):3440–3458. https://doi.org/10.1002.pc.25642
7. Baran I, Cinar K, Ersoy N, Akkerman R, Jesper HH (2017) A review on the mechanical modeling of composite manufacturing processes. Arch Comput Methods Eng **24**:365–395. https://doi.org/10.1007/s11831-016-9167-2
8. Kondratiev A (2019) Improving the mass efficiency of a composite launch vehicle head fairing with a sandwich structure. Eastern-European J Enterp Technol 6(7(102)):6–59. https://doi.org/10.15587/1729-4061.2019.184551
9. Deng B, Shi Y, Yu T, Zhao P (2020) Influence mechanism and optimization analysis of technological parameters for the composite prepreg tape winding process. Polymers 12(8):1843. https://doi.org/10.3390/polym12081843
10. Tomashevskii VT, Yakovlev VS (2004) Models in the engineering mechanics of polymer-matrix composite systems. Int Appl Mech 40(6):601–621. https://doi.org/10.1023/B:INAM.0000041391.28104.b7
11. Nikolaev VP, Pichugin VS, Korobeinikov AG (1999) Effect of molding conditions on fracture mechanisms and stiffness of a composite of grid structure. Mech Compos Mater 35:49–54. https://doi.org/10.1007/BF02260811
12. Nemirovskii YV, Yankovskii AP (2002) Effect of the thermal action and thermosensitivity of phase materials on the load-carrrying capacity of momentless shells with an equal-stressed reinforcement. Mech Compos Mater 38:525–538. https://doi.org/10.1023/A:1021778626055
13. Teters G, Kregers A (2000) Optimization of a composite plate buckling under thermal action with account of reliability. Mech Compos Mater 36:453–458. https://doi.org/10.1023/A:1006750431407
14. Verbitskaya NA (2001) Influence of complex compounds of rhenium (V), molybdenum (V) with macrocyclic ligands on processes of structure formation in epoxypolyurethane binder. Plasticheskie massy: sintez svojstva pererabotka primenenie. 7:10–14
15. Nemirovskii YV, Yankovskii AP (2000) Effect of the thermosensitivity and reinforcement structure on the load-carrying capacity of metal-composite gas-turbine disks. Mech Compos Mater 36:481–491. https://doi.org/10.1023/A:1006710717295

16. Kovalov AI, Otrosh YA, Vedula S, Danilin OM, Kovalevska TM (2019) Parameters of fire-retardant coatings of steel constructions under the influence of climatic factors. Sci Bull Nat Mining Univ 3:46–53. https://doi.org/10.29202/nvngu/2019-3/9
17. Jinno M, Sakai S, Osaka K, Fukuda T (2012) Smart autoclave processing of thermoset resin matrix composites based on temperature and internal strain monitoring. Adv Compos Mater 12(1):57–72. https://doi.org/10.1163/156855103322320374
18. Russell JD, Madhu SM, Mohamed SG, Andre YL (2000) A new method to reduce cure-induced stresses in thermoset polymer composites. Part III: Correlating stress history to viscosity, degree of cure, and cure shrinkage. J Compos Mater 34(22):1925–1947
19. Slyvyns'kyy V, Gajdachuk A, Tkachenko G, Kirichenko V, Karpikova O, Verbitskaya N (2009) Creation of energy-saving technologies of forming articles made of polymeric composite materials. In: Proceedings of 60th International Astronautical Congress, Daejeon, South Korea. IAC-09. C2.4.9
20. Blagonadezhin VL, Vorontsov AN, Murzakhanov GK (1988) Technological problems of mechanics of structures made of composite materials. Mech Compos Mater 23:608–625. https://doi.org/10.1007/BF00605687
21. Jaeger JC (1969) Elasticitfracture and flow. Springer. https://doi.org/10.1007/978-94-011-6024-7
22. Gaidachuk VE, Sidorenkova MA (1997) Selection of the optimal pressure when molding structures from polymer composite materials. In: Design and production of aircraft structures, pp 8–12
23. Bitjukov YuI, Kalinin VA (2010) The numerical analysis of the scheme on packing of the tape of variable width on the technological surface in the course of winding of designs from composite materials. Mekhanika kompozitsionnykh materialov i konstruktsii. 16(2):276–290
24. Rodionov VV (2019) Optimization of molding the polymeric composite material with improved characteristics. Plast Massy 3–4:55–58. https://doi.org/10.35164/0554-2901-2019-3-4-55-58
25. Kolosov AE, Sakharov AS, Sivetskii VI, Sidorov DE, Sokolskii AL (2012) Substantiation of the efficiency of using ultrasonic modification as a basis of a production cycle for preparing reinforced objects of epoxy polymer composition. Chem Pet Eng 48:391–397. https://doi.org/10.1007/s10556-012-9629-9
26. Kolosov AE, Virchenko GA, Kolosova EP, Virchenko GI (2015) Structural and technological design of ways for preparing reactoplastic composite fiber materials based on structural parametric modeling. Chem Pet Eng 51:493–500. https://doi.org/10.1007/s10556-015-0075-3
27. Mustafa LM, Ismailov MB, Sanin AF (2020) Study on the effect of plasticizers and thermoplastics on the strength and toughness of epoxy resins. Naukovyi visnyk Natsionalnoho Hirnychoho Universytetu. 4:63–68. https://doi.org/10.33271/nvngu/2020-4/063
28. Johannes KF (2005) Reactive polymers fundamentals and applications. William Andrew Publishing, New York
29. Karandashov O, Avramenko V (2017) Studies of thermal stability of epoxy compound for glass-fiber pipes. Chem Chem Technol 11(1):61–64. https://doi.org/10.23939/chcht11.01.061
30. Ayranci C, Carey JP (2011) Experimental validation of a regression–based predictive model for elastic constants of open mesh tubular diamond–braid composites. Polym Compos 32:243–251. https://doi.org/10.1002/pc.21042

Prognostic Model of a Photovoltaic Power Plant

Alexandr Zaslavskiy and **Oleh Karpenko**

Abstract Economically viable integration of such variable renewable energy (VRE) resources as wind and solar power into local, regional, and national electric energy systems is impossible without active balancing of the whole energy market. In this context, hourly prediction of electric energy generation by means of the renewable sources one day in advance is one of the most topical problems. It should be mentioned that despite the availability of numerous proposed algorithms, programs, and program systems the problem has not any adequate universally recognized solution. Electric energy generation by means of photovoltaic power stations depends directly upon weather changes becoming more and more unpredictable. Strictly speaking, the problem of photovoltaic solar energy conversion amounts to a problem of developing a mathematical model able to reflect the most accurate dependence of energy production upon each influencing factor. The paper analyzes certain approaches to develop the taught prognostic models of photovoltaic power stations.

Keywords Photovoltaic power plant · Machine learning algorithms · Influence function · Reflexive learning of a photovoltaic converter

1 Introduction

Solar energy is almost inexhaustible resource. By 2070, it is expected to become the key electricity source in this world. At the beginning of following century, volumes of solar energy will exceed petroleum industry by 3.5 and nuclear industry by 6 times. Being a buy/sell object in power market, photovoltaic power should be predictable concerning its generation volumes. Accuracy of the prognosis depends upon the necessity of a power market to balance; moreover, it is administrated by regulatory framework. For instance, Ukrainian solar energy supplier should provide hourly prognosis one day in advance where dissimilarity from net actual generation is not more than 5%. However, the process of electric energy generation using photovoltaic power

A. Zaslavskiy (✉) · O. Karpenko
Dnipro University of Technology, Dnipro 49000, Ukraine
e-mail: zaslavskyi.o.m@nmu.one

plants differs in its nonstationarity, seasonal variability, and dependence upon random meteorological factors. The facts complicate heavily the balancing of a photovoltaic power market. In addition, the problem cannot be solved without the use of rather accurate prognostic models of photovoltaic power plants.

Currently, several quite different approaches are available to develop prognostic models of processes transforming solar power into electric one. The approaches can be classified in terms of degree of use of physical equations describing processes of photovoltaic power plants. According to the systematic, three classification model groups are singled out. They may be specified as models of a white box, black box, and grey box.

Group 1 is based on the determined description of the specific procedures transforming solar energy into electric one. For instance, [1] proposes electrothermic macromodel of a photovoltaic power plant accepting environmental temperature, lighting intensity, and wind velocity as the input variables, and generating capacity at the output one. The key problem to develop and apply the models is the necessity of the advanced online knowledge of each physical variables being not available completely in the majority of cases.

Group 2 of the models implements 'black box' strategies ignoring real physical relations between the input variables and output power. A model of 'black box' type is developed using such machine learning methods [2, 3] as neural networks [4]; support vector regression [5]; quantile random forest [6]; gradient boosting [7]; and ensemble models [8–10] incorporating the abovementioned methods as well as similar ones. Generally, accuracy of the models depends upon possibility to have rather vast retrospective database. However, it is common knowledge that they are quite applicable (see, for instance [8, 11]).

Group 3 of the models differs in the fact that, in their context, input variables-output power ratio is specified using several simplified physical correlations (e.g. when high-order dependencies are ignored). Example of the model, involving actual simplified ratio between the temperature and lighting intensity, assumed as the input variables, and output power as the response variable, is demonstrated in [12].

In the context of another system of classification criteria, models differ depending upon a type of the used input variables. Endogenic models are the ratio between past states of the process being simulated, and its future states. Data on the past states of the prognosticated time series are the input variables of the regression models. Testing results of such classical regression models as ARIMA, using Box-Jenkins approach [13], support the idea that their accuracy is comparable with similar neural network-based models. Paper [14] contains data on high efficiency of nonlinear regression algorithm Theta.

Exogenic models [14, 15] are the ratio between the predicted values of photovoltaic power and simultaneous influential factors being sky cover, temperature, humidity, wind velocity, solar insolation level etc.

Furthermore, in terms of models of a 'black box' type, mathematical structure of the ratio is formed in the process of the model learning using retrospective database.

Papers [12, 13] demonstrate interesting research results as well as comparison of different approaches to develop prognostic models of time series. Among other

things, it concerns processes of electricity generation by means of photovoltaic power plants. On the one hand, the findings show clearly that machine learning methods (i.e. neural networks, fuzzy mathematics, quantile forests etc.) cannot improve significantly prognosis accuracy to compare with classical approaches, which minimizing standard deviation (e.g. with the use of Moore–Penrose pseudoinverse matrices). On the other hand, they consider insufficient accuracy of meteorological forecasts within the geolocation points of photovoltaic power plants as the main reasons for the errors. In this context, meteorological errors within the retrospective database, used for model learning, results in the fact that future forecasting involves the following: even accurate projected meteorological data are interpreted by the model in the form of erroneous predictions of photovoltaic power generation. Origin of the errors is as follows: in many cases, retrospective database forms itself as a model instead of being a result of direct and rather accurate measurements within the geolocation points of photovoltaic power plants. That happens due to the lack of local certified meteorological stations.

Recently, the interest to a problem of improving accuracy of prognostic models of photovoltaic plants in terms of nonideal meteorological prognosis has been increasing. For instance, paper [16] proposes a model of a short-term prognosis of photovoltaic power based on the online sequential extreme learning machine with for-getting mechanism (FOS-ELM) which can replace constantly the old data with the new ones providing high learning accuracy and speed during any season. The paper [17] proposes a hybrid improved multi-step algorithm (HIMVO) for optimizing a support vector machine for predicting the output of photovoltaic cells.

The HIMVO algorithm introduces random sequences to initialize the set. This significantly accelerates the convergence of the algorithm.

Paper [18] proposes prognosis methods for time series of the output power a day-ahead where ideal weather type nd nonideal one are discussed separately. In the context of ideal weather conditions, a prognosis method based on the next-day meteorological data is proposed where long short-term memory (LSTM) networks are used. In the context of nonideal weather conditions, topicality of time series and specific characteristic of nonideal weather type are considered in the LSTM model by introducing adjacent day time series and weather type data.

The forecited short review of the problems, concerning simulation of photovoltaic processes (as well as approaches to solve them), prove the topicality of studies of prognostic models being robust ones relative to the errors of meteorological database used for their learning. The paper is aimed at the development of original methods to construct such a type of prognostic models. The methodology is based upon the idea of reflexive learning.

2 Method

2.1 Structural Pattern of a Prognostic Model of a Photovoltaic Power Station

Conceptual framework of the approach to model a photovoltaic power plant, presented in the paper, can be characterized with the help of following classification features: exogenic model of a 'grey box' type (see Fig. 1). The ideal level of solar insolation $S(t)$ (inclusive of both direct and diffuse components) under the 'clear sky' conditions, and proportional to it power quantity, generated under the perfect conditions, are identified by means of the determined algorithm as a function from the time moment (i.e. hour, day, month) characterizing mutual arrangement of the
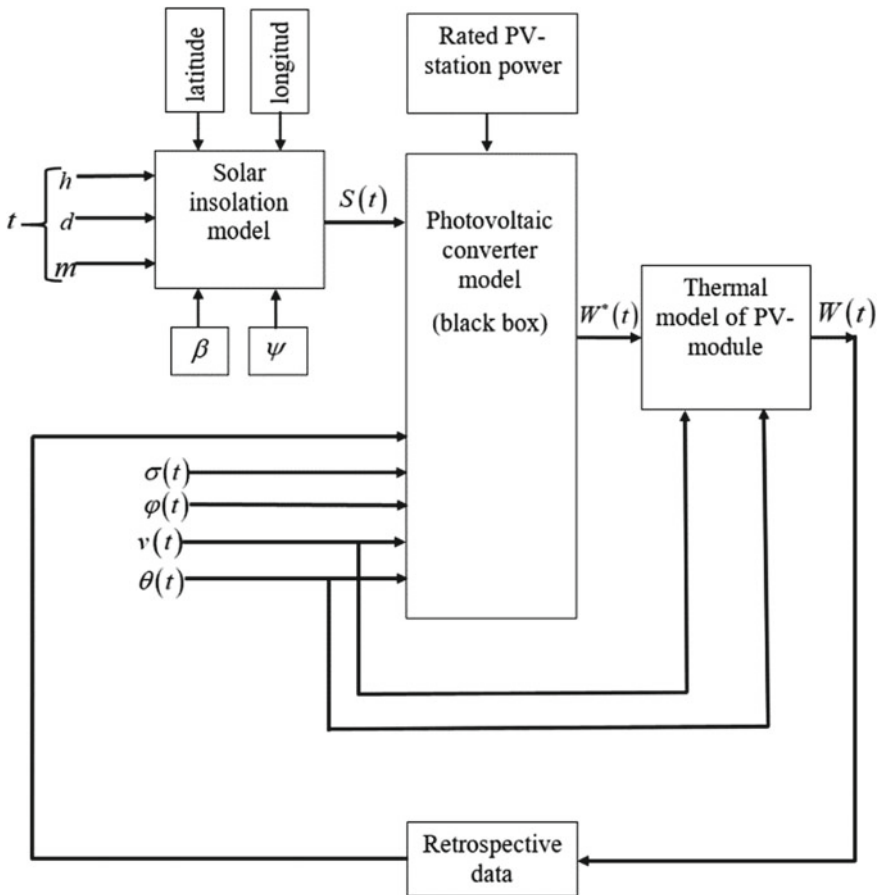
Fig. 1 Structural scheme of a PV-station model
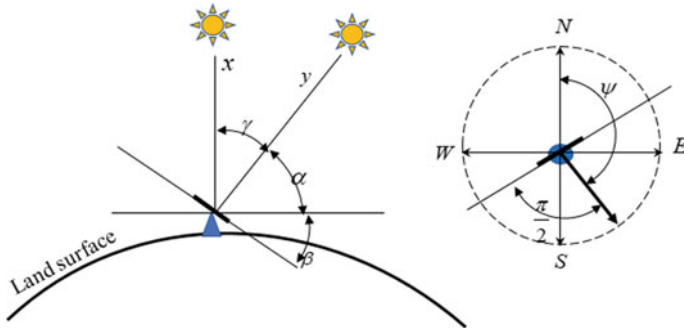
Fig. 2 Parameters of mutual orientation of the sun and a photovoltaic module

Sun, the Earth, and a conventional point at its surface indicating location of the prognosticated photovoltaic power plant as well as a function of angles at which sun rays fall onto the photovoltaic panels (see Fig. 2). Actual electric energy quantity $W(t)$, generated by a photovoltaic power plant, is always less than the ideal level depending upon numerous influencing meteorological factors. The considered models takes into consideration the following: air temperature $\theta(t)$ and PV module temperature $\Theta(t)$; air humidity $\varphi(t)$; wind velocity $v(t)$; and cloudiness $\sigma(t)$. Distinctive feature of photovoltaic power plant is as follows: all the meteorological factors influence the energy output in a multiplicative manner. If they vary, the energy output may experience its percentage-wise increase or decrease; however, it can neither exceed the ideal level nor become negative. The changes in the capacity of photovoltaic power plant, resulting from the influence of meteorological factors, take place owing to variations of a coefficient of performance rather than owing to summing of some extra energy by other sources with solar power.

In this context, meteorological factors stipulate certain variable resistivity of the model relative to the input solar power.

X—shortest length of the path passed by the light passes through the atmosphere, y—length of the path passed by the light through the atmosphere at the specified moment of time, γ—zenith angle, α—angle of sun elevation, β—angle of module inclination, ψ—polar angle by which the module is turned relative to the direction 'towards the North Pole'.

2.2 Model of a Photovoltaic 'black Box'

While ignoring real physical nature of photovoltaic converter, represent 'input–output', characterizing it, as follows

$$W^*(t) \approx S(t) \cdot F(\sigma, \varphi, v, \theta, h, d) \tag{1}$$

where $W^*(t)$ is electrical energy, generated by a power station in terms of precritical temperature limits above which abnormal decrease in the generation level takes place; $F(\sigma, \varphi, \theta, v, h, d)$ is function of photovoltaic conversion taking into consideration dependence of its efficiency upon weather conditions within the PV station location as well as upon circular time represented by hour h during the day and day d within a month.

The problem to develop a model of photovoltaic 'black box' is to express the most accurately in an explicit form dependence of $F(\sigma, \varphi, \theta, v, h, d)$ function on its arguments based upon the retrospective data concerning solar insulation value $S(t)$, power being generated $W^*(t)$ and the considered meteorological data obtained during the model learning.

Our approach for the problem solution relies on a hypothesis, in terms of which function $F(\sigma, \varphi, \theta, v, h, d)$ may be factorized and represented in the form of product of partial influence functions

$$F(\sigma, \varphi, v, \theta, h, d) = f_1(\sigma) \cdot f_2(\varphi) \cdot f_3(v) \cdot f_4(\theta) \cdot f_5(h) \cdot f_6(d) \qquad (2)$$

Further, represent each influence function as an exponential of the correspondent polynom of k degree

$$f_1(\sigma) = \exp\{a_{10} + a_{11}\sigma^1 + a_{12}\sigma^2 + \cdots + a_{1k}\sigma^k\},$$
$$f_2(\varphi) = \exp\{a_{20} + a_{21}\varphi^1 + b_{22}\varphi^2 + \cdots + b_{2k}\varphi^k\},$$
$$f_3(v) = \exp\{a_{30} + a_{31}v^1 + a_{32}v^2 + \cdots + a_{3k}v^k\},$$
$$\cdots$$
$$f_6(d) = \exp\{a_{60} + a_{61}d^1 + a_{62}d^2 + \cdots + a_{6k}d^k\}. \qquad (3)$$

In the process of the model learning, such values of a_{ij}, coefficients are selected, in terms of which mean-square imbalance of approximation equation is minimized (1). While taking logarithms of right and left members of (1) equation, represent the minimized functional polynomially

$$\sum_t \left\{\ln\left(\frac{W^*(t)}{S(t)}\right) - \ln[F(\sigma, \varphi, v, \theta, h, d)]\right\}^2 = \min \qquad (4)$$

Involving conditions of extremum achieving

$$\frac{\partial}{\partial a_{ij}}\left[\sum_t\left\{\ln\left(\frac{W^*(t)}{S(t)}\right) - \ln[F(\sigma, \varphi, v, \theta, h, d)]\right\}^2\right] = 0 \qquad (5)$$

we obtain $6 \cdot k$ system of linear equations relative to the required a_{ij}, coefficients to identify values of photovoltaic converter function $F(\sigma, \varphi, \theta, v, h, d)$.

In the light of the determined values of a_{ij} coefficients, the predicted $\hat{W}^*(t)$ value of power generation is

$$\hat{W}^*(t) = S(t) \cdot F(\sigma, \varphi, v, \theta, h, d) \tag{6}$$

2.3 Thermal Value of PV Module

In the context of convective heat exchange with the environment, the equation of thermal balance of a PV model may be represented as follows

$$cm\frac{dT_M^0}{dt} = S(t)\frac{\varepsilon}{\xi} - \left(T_M^0 - \theta\right) \cdot \alpha(V) \tag{7}$$

where T_m^0 is module temperature; $S(t)$ is solar insolation energy [W/м²]; θ is air temperature; v is wind velocity; $\alpha(V)$ is heat-exchange coefficient of the module involving air thermal conductivity in terms of the given wind velocity, its kinematic viscosity, expansion rate, and Prandtl, Rayleigh as well as Nusselt criteria; is module mass normalized to a unit of its surface area [kg/m²], ε—modulus black ratio, ξ is cooling ratio of the module ($\xi = 1$, if cooling is performed from the one side of the module surface; and $\xi = 2$, if cooling is performed from both sides).

According to the calculations, time constant to heat PV modules, used as the components of solar panels, is not more than 10 min. Thus, the module temperature, averaged in terms of an hour, may be determined from a thermal balance condition

$$T_M^0 = \frac{S(t) \cdot \varepsilon}{\alpha(v)\xi} + \theta \tag{8}$$

Electric generation, predicted involving the module temperature, is determined using the formula

$$\hat{W}(t) = \begin{cases} \hat{W}^*(t)\left[1 - \frac{k_p}{100}\left(T_M^0 - \Theta_p\right)\right], \text{if } T_M > \Theta, \\ \hat{W}^*(t), \text{if } T_M^0 \leq \Theta \end{cases} \tag{9}$$

where $\Theta(t) = 25\ °C$ is thermal threshold of power reduction of the module; and $k_p \approx 0.45\%/°C$ is temperature coefficient of power reduction of photovoltaic module.

2.4 Reflexive Learning of a Photovoltaic Converter Model

To compare influence of various factors, stipulating operation of a photovoltaic power plant, on the value of electric energy, generated by it, consider modeling results of three power stations with 10mW, 2mW, and 4mW capacities (see Fig. 3) located within different geographic areas of Ukraine. Figure 3 shows graphs of influence functions resulted from model learning using retrospective data arrays. Learning period for the first two power stations was 12 months; and 8 months for the third station. Factors σ, φ, θ, v, h, d, involved in (3) formulas, influencing generation
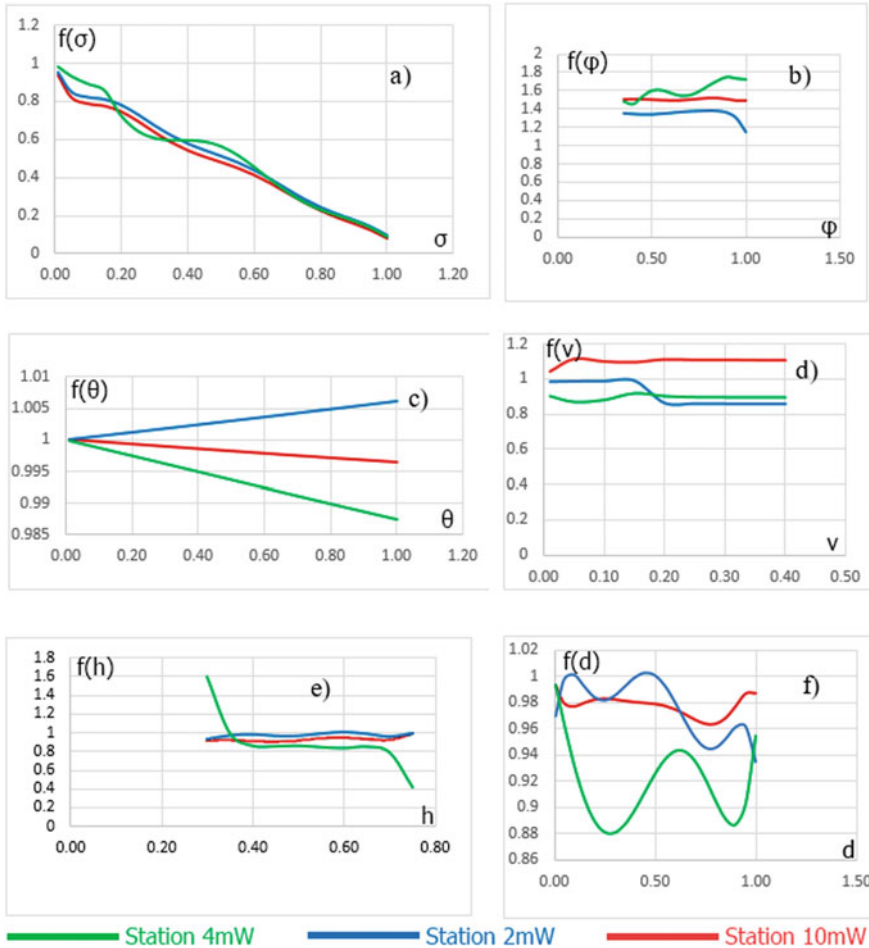
Fig. 3 Graphs of dominant functions. **a** dominant function of cloud cover $f(\sigma)$; **b** dominant function of humidity $f(\varphi)$; c) dominant function of temperature $f(\theta)$; **d** dominant function of wind $f(v)$; **e** dominant function of time (hour per day $f(h)$); **f** dominant function of time (day per month $f(d)$)

level of photovoltaic power, are in relative units. 100% cloud cover, 100% humidity, 50 °C temperature, 50 m/s wind velocity, cyclical hour—24, and cyclical day—31 correspond to value 1. −50 °C corresponds to 0 within the temperature effect graph.

The value range, which may be taken by an influence function, is indicative of a degree the corresponding factor influences a ratio between the power output and solar insolation of a 'clear sky'.

The demonstrated graphs of the influence function explain that cloud cover is the most influential factor (as one could expect). Unfortunately, in this regard actual cloud cover above a photovoltaic power plant often differs greatly from the predicted one. That may result from erroneous prediction as well as from unpredictable time-dependent changes in the correct prediction of average density of cloud cover right above the power station. Taking into consideration the fact that generally the weather archives, used to learn PV models, are climatic models for the given geolocation points, we obtain that the errors of climate modeling are brought to the models of photovoltaic energy conversion.

To achieve robustness relative to errors of climatic models, we proposed reflexive method to learn PV models. According to the method, retrospective data of integral cloud cover level right above a photovoltaic panel are determined approximately in the form of a ratio between the measured specific capacity of the generated electric energy and the predicted level of insolation of a 'clear sky' taking into consideration corrections for technological factors of fullness of solar panel use, and ignoring slightly varying dominant functions of other factors. The data array of retrospective cloud cover, obtained in such a way, is further applied to learn a PV model together with other meteorological data. Hence, in consequence of the reflexive learning, a factor of the enhanced sensitivity of the PV model to the errors of climatic models (i.e. cloud cover dominance as an influential factor) is applied to minimize the sensitivity.

3   Discussion of the Results

Figure 4 demonstrates examples of classical learning of PV models (in terms of a criteria minimizing mean-square error) as well as reflexive one.

Classical learning errors (see Fig. 4a) depend upon the classical model sensitivity for errors of retrospective data according to which learning is performed. In Fig. 4a, significant errors within 10:00–14:00 time interval result from incorrectly learned function of cloud cover influence rather than erroneous forecasts for the period. One can see that due to the incorrect learning, the predictable output is even more than solar insolation which is impossible from physical viewpoint. In this context, for the same period, cloud cover percentage, determined reflexively on the electric energy generation level as well as upon the predicted solar insolation (17%), differs insignificantly from the data of meteorological forecast (20%).

Diagram in Fig. 5 explains the results of comparative estimation of reflexive learning efficiency on the criterion of relative mean absolute error rMAE for 13 month learning period. For this purpose, rMAE is determined as follows
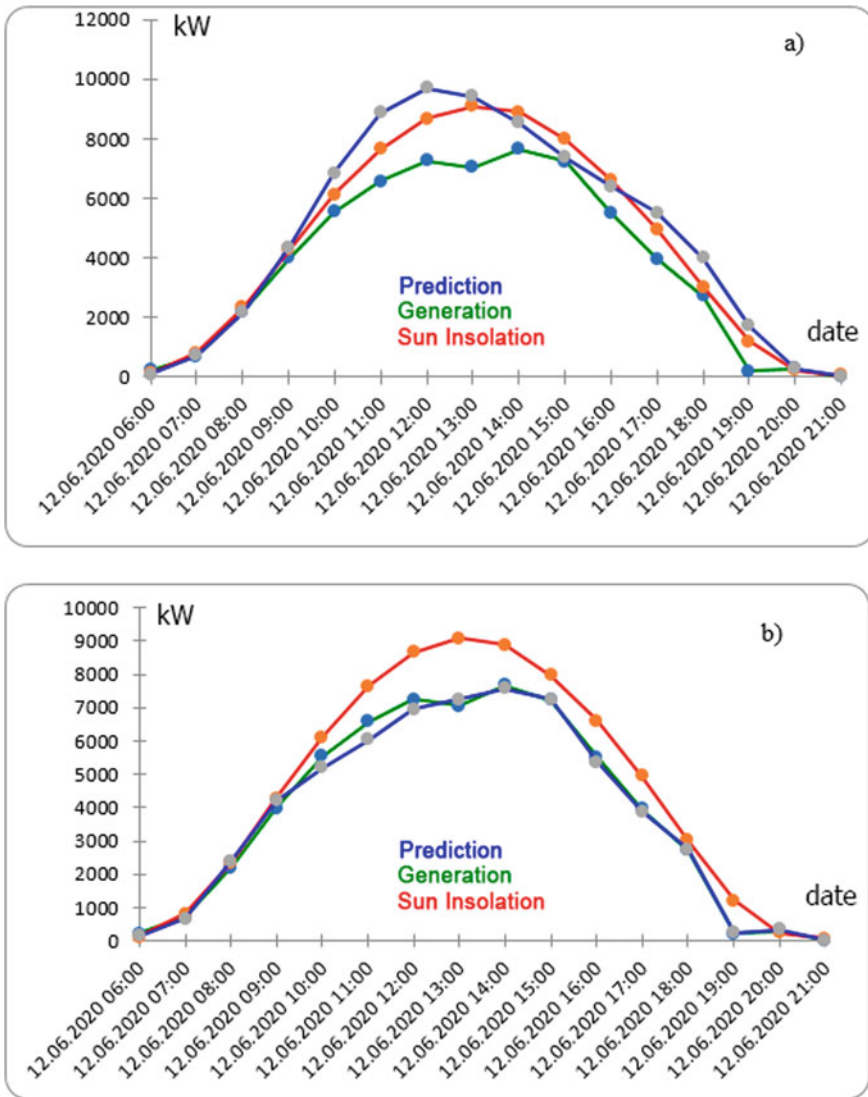
Fig. 4 Results of classical (**a**) and reflexive (**b**) learning of PV-models

$$rMAE = \frac{\sum_t |P(t) - G(t)|}{\sum_t G(t)} \cdot 100 \qquad (10)$$

where $P(t)$ is prognosticated output per current hour; and $G(t)$ is generated power measured during the current hour.s

The diagram in Fig. 5 supports clearly significant dependence of the classical method error upon the accuracy of climatic models applied for learning. Accuracy
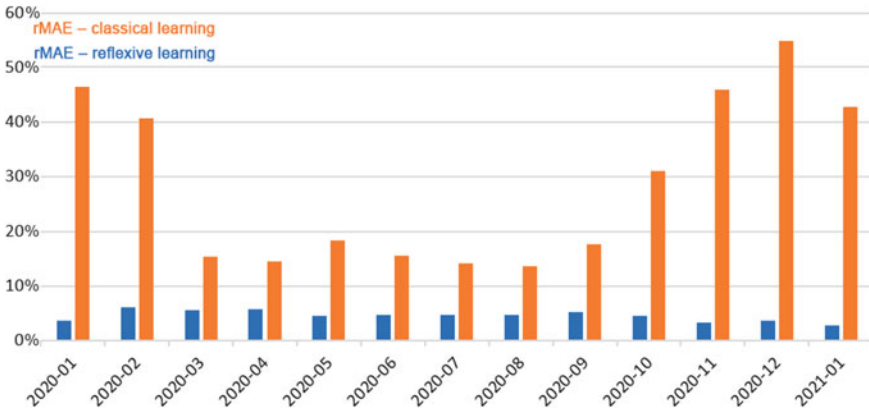
Fig. 5 Results of the comparative assessment of the reflexive learning efficiency according to the criterion of mean absolute error

of the models increases during summer decreasing in winter and autumn. As it is understood from the diagram, behavior of the classical model is similar. At the same time, reflexive model accuracy does not correlate noticeably depending upon a season.

4   Conclusions

The following can be considered as the specific features of the proposed model of photovoltaic conversion of solar energy:

1.   factorization of the photovoltaic conversion function in the form of a product of partial dominant functions;
2.   analytical calculation of the functions on the criterion of minimum mean-square prediction error;
3.   reflexive learning of photovoltaic models.

Factorization of photovoltaic transformation function simplifies drastically the model while reducing learning period since numerous weakly influencing combinations of simultaneously acting input factors are excluded from the consideration.

Analytical calculation of the influence functions while solving systems of algebraic equations saves time required to learn PV models to compare, for instance, with the methods of successive approximations based upon the use of neural systems. Digital systematization and averaging of influence functions, derived with the help of direct calculation, creates conditions for PV model start in terms of minimum depth of array of retrospective data applied for the model learning. That is extremely important to predict operations of new power stations being connected to electric energy system.

Reflexive learning of PV models reduces drastically their sensitivity for errors of climatic models used for the learning.

The developed model is used in the specialized software for a day-ahead advance prognosis of 50 photovoltaic plants located in different regions of Ukraine. Nominal power of the plants is 1–15 mW. Depth of the archive database, in terms of which learning is provided, is the basic limiting condition to use the proposed model. It should be more than the length of any sequence, being a part of it, of similar (unaltering) data. Otherwise, singularity of system matrix results in noncomputable functions of influence determining the forecast. Further development of the proposed approach is seen as the improvement of calculation of algorithm solar radiation taking into consideration diffusion and reflected components as well as taking into account the trend of errors of meteorological weather forecasts.

References

1. Bizzarri F, Bongiorno M, Brambilla A, Gruosso G, Gajani GS (2013) Model of photovoltaic power plants for performance analysis and production forecast. IEEE Trans Sustain Energy 4(2):278–285
2. Raza MQ, Nadarajah M, Ekanayake C (2016) On recent advances in PV output power forecast. Sol Energy 136:125–144
3. Antonanzas J, Osorio N, Escobar R, Urraca R, Martinez-de-Pison FJ, Antonanzas-Torres F (2016) Review of photovoltaic power forecasting. Sol Energy 136:78–111
4. Gensler A, Henze J, Sick B, Raabe N (2016) Deep learning for solar power forecasting—an approach using autoencoder and LSTM neural networks. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, pp 002858–002865
5. Jang HS, Bae KY, Park H-S, Sung DK (2016) Solar power prediction based on satellite images and support vector machine. IEEE Trans Sustain Energy 7(3):1255–1263
6. Biau G (2010) Analysis of a random forests model. Technical report, University Paris 6. 2010. pp 1–31. URL: http://hal.archivesouvertes.fr/docs/00/47/65/45/PDF/article2.pdf
7. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29(5):1189–1232. http://www.jstor.org/stable/2699986
8. Yona A, Senjyu T, Funabashi T, Kim C-H (2013) Determination method of insolation prediction with fuzzy and applying neural network for long-term ahead PV power output correction. IEEE Trans Sustain Energy 4(2):527–533
9. Ren Y, Zhang L, Suganthan PN (2016) Ensemble classification and regression-recent developments, applications and future directions. IEEE Comp Int Mag 11(1):41–53
10. Yang H-T, Huang C-M, Huang Y-C, Pai Y-S (2014) A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. IEEE Trans Sustain Energy 5(3):917–926
11. Liu J, Fang W, Zhang X, Yang C (2015) An improved photovoltaic power forecasting model with the assistance of aerosol index data. IEEE Trans Sustain Energy 6(2):434–442
12. Gigoni L, Betti A, Crisostomi E, Franco A, Tucci M, Bizzarri F, Mucci D (2019) Day-ahead hourly forecasting of power generation from photovoltaic plants. arXiv:1903.06800v1 [cs.LG] 26 Feb 2019
13. Makridakis S, Spiliotis E, Assimakopoulos V (2018) Statistical and machine learning forecasting methods: concerns and ways forward. PloS One 13(3):e0194889
14. AlKandari M, Ahmad I (2020) Solar power generation forecasting using ensemble approach based on deep learning and statistical methods, Appl Comput Inform. https://doi.org/10.1016/j.aci.2019.11.002

15. Larson DP, Nonnenmacher L, Coimbra CFM (2016) Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest. Renew Energy 91:11–20
16. Wang J, Ran R, Zhou Y (2017) A short-term photovoltaic power prediction model based on an FOS-ELM Algorithm. Appl Sci 7(4):423. https://doi.org/10.3390/app7040423
17. Li L-L, Wen S-Y, Tseng M-L, Wang C-S (2019) Renewable energy prediction: a novel short-term prediction model of photovoltaic output power. J Cleaner Prod 228:359–375. https://www.sciencedirect.com/science/article/abs/pii/S0360544219315105
18. Gao M, Li J, Hong F, Long D (2019) Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. Energy 187:115838. https://doi.org/10.1016/j.energy.2019.07.168

Investigation of Stationary Processes in Vortex Energy Separator Through Its Computational Fluid Dynamics Model

Anatoliy S. Kulik ⓘ**, Sergey N. Pasichnik** ⓘ**, and Dmytro V. Sokol** ⓘ

Abstract The investigation of vortex effect arisen in a vortex energy separator, also known as a Ranque-Hilsch tube is considered in the paper. The simple device with no moving parts is widely used in propulsion, medicine, acoustic systems, industry, thermoregulation, aviation, etc. In order to understand and investigate the physics of stationary processes a computational fluid dynamics (CFD) model of gas-dynamic flow in a vortex energy separator is proposed to use. Relationship of such parameters of the working air flow as temperature, pressure and flow rate at each section of the vortex energy separator was considered. CFD simulation results were used to obtain nonlinear static characteristics of a vortex energy separator. The interval linearization of nonlinear static characteristics was carried out and linear mathematical models of the stationary process in the vortex energy separator as an automatic stabilization object were formed.

Keywords Ranque-Hilsch tube · Vortex energy separator · Temperature separation · CFD model · Static characteristic · Interval linearization

1 Introduction

For the first time, the vortex effect was recorded by the French engineer Joseph Ranque in 1928. In order to study the thermal characteristics and increase the efficiency of temperature separation, the German physicist Robert Hilsch in 1945 conducted experiments at different inlet pressures and various geometric parameters of the experimental setup, later called the Ranque-Hilsch tube. In the series of experiments to improve the performance of the Ranque-Hilsch tube, its design had been changed, so the device was renamed as "vortex energy separator". The simplicity of obtaining the vortex effect in a vortex energy separator has led to its widespread application in various technical devices in engine building, medicine, acoustic systems, industry, thermoregulation systems, aviation, etc. [1].

A. S. Kulik · S. N. Pasichnik · D. V. Sokol (✉)
National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine
e-mail: d.sokol@khai.edu

The instability of the gas output parameters in the known devices has led to the need to study the process of converting potential energy into kinetic energy. A number of studies made it possible to clarify partially the temperature separation process essence and identify the vortex effect instability causes [2]. These studies in major are aimed at the performance coefficient increasing by changing the vortex energy separator design and using the known heat and gas dynamics equations. Accepting the fact that the optimal design searching did not give conventional result, this problem cannot be solved using such an approach. Automatic stabilization systems have the fundamental ability to avoid instability [3]. For the synthesis of high-precision temperature stabilizing system, knowledge is required in the form of mathematical models of a vortex energy separator. In the scientific literature, you can find materials devoted to a detailed step-by-step description of the processes occurring in a vortex energy separator, from the standpoint of heat transfer and hydrodynamics [4–6]: starting with turbulence models and ending with a mathematical description and methods for calculating swirling flows. However, no articles were found which consider the efficiency increase of a vortex energy separator using such a mathematical representation in form of transfer function or in the state space which could be utilized for the synthesis of the temperature stabilization system. Therefore, it is relevant to study the vortex effect process to form its linear mathematical models as an automatic stabilization object.

The known results of scientific research on the vortex effect physics do not allow to form analytical models of the input air flow potential energy conversion into the output flow thermal energy [7]. Therefore, the most constructive approach is the parameter identification to obtain mathematical model of vortex energy separators as the automatic stabilization object of the outlet gas flows temperature.

The purpose of this research is to describe the physics of the processes occurring in the prototype vortex energy separator using the computational fluid dynamics (CFD) model, to obtain nonlinear static characteristics as dependences of the cold and hot air flows temperature on the adjusting valve position at various values of the inlet air flow pressure, and to form the linearized mathematical models of the considered vortex energy separator.

2 Experimental Setup

The view of the experimental setup of vortex energy separator [8] is shown in Fig. 1. The vortex energy separator is the reverse-flow Ranque-Hilsch tube with one tangentially swirl nozzle inlet $F_n = 1$ mm^2, energy separation chamber with diameter $D = 5.8$ mm and length $L = 116$ mm, adjusting valve that is located on the hot flow outlet side with travel range $\Delta\zeta = 2$ mm, diaphragm diameter $d = 2.5$ mm. The inlet compressed air pressure is $P = 5-7$ atm., environment temperature $T_{env} = 18.85°C$, relative humidity $\overline{\varphi} = 85\%$.

The sensors include three temperature sensors based on platinum thermistors for both inlets (hot and cold flows) and pressure sensor. Each sensor is connected to the

Fig. 1 The view of the experimental setup of vortex energy separator

matching unit where DC signals are generated that proportionally vary in the range 0 to + 5 V according to the physical parameters change. The executive device is implemented on the basis of stepper motor with four-phase control and angular step $\varphi = 1.8°C$. The stepper motor shaft moves the adjusting valve linearly through the gearbox.

All experiments were carried out at fixed pressure generated by the compressor: 5, 6, and 7 atm. The temperature was recorded for a certain period of time after each change of the adjusting valve position throughout its operating range.

The functional diagram that reflects the air flow through the composed functional elements of the vortex energy separator is shown in Fig. 2.

The settings of experimental setup are presented in [8].

3　CFD Model

CFD model is a complex of physical, mathematical and numerical methods for computing the characteristics of flow processes. The drawings according to which the model was developed, and the results of experiments carried out on the experimental setup (see Fig. 1) were used to create the CFD model.
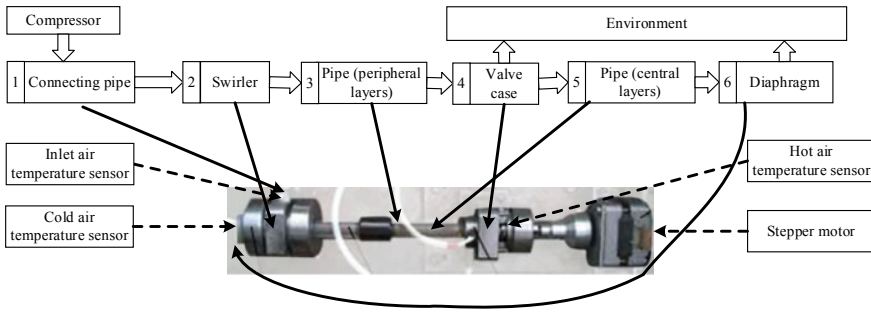
Fig. 2 Functional diagram of the experimental setup

The design was made in the SolidWorks modeling environment, numerical modeling of the fluid flow in the vortex energy separator was performed using the built-in tool for flow processes modeling Flow Simulation Library. Figure 3 shows the results of modeling the vortex effect for compressed air temperature—18.8 °C, pressure—6 atm. and the adjusting valve position—0.5 mm.

The parameters of the fluid region determine the steady-state conditions of the vortex energy separator. Temperature, pressure and air flow velocity were selected as the air flow parameters for CFD modeling analysis.
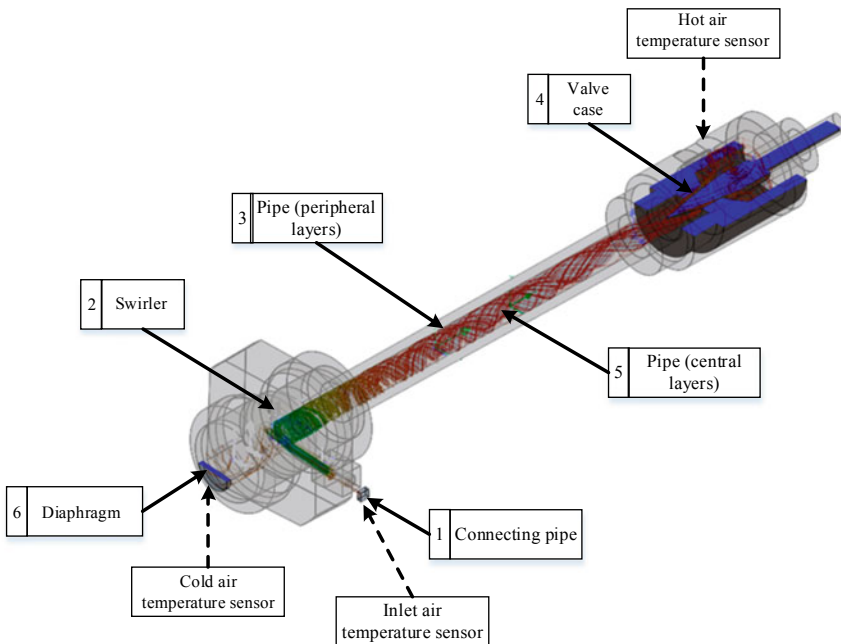
Fig. 3 Flow structure in the vortex energy separator

The computational experiments were performed with constant air flow parameters and various positions of the adjusting valve: from the complete overlap of the corresponding outlet of the vortex energy separator to its extreme position. As a result, the flow paths and its surface parameters were obtained.

The analysis of the obtained processes confirms the following hypothesis [9]: in the center of the vortex is that part of the inlet air flow, which initially had an insignificant supply of kinetic energy, and the mechanism that ensures that this particular part of the flow enters the center of the vortex is the separation in the field of centrifugal forces flow elements with different tangential velocity.

At the moment compressed air enters through the nozzle 1 into the vortex energy separator, the uniform flow has higher potential energy, expressed in air pressure, and lower kinetic energy, characterized by the air flow velocity. The compressed air flows through the vortex energy separator according to Fig. 2.

The velocity of air flow increases abruptly and the temperature and pressure decrease when the flow enters the tangential swirl inlet 2 due to the narrowing of the working area. In the other words, energy is converted from potential to kinetic.

After the swirling flow enters the peripheral layers of the energy separation chamber 3, the kinetic energy gradually transforms into thermal energy, the flow velocity decreases due to the fact that the friction force significantly increases between the air particles and the inner wall of the vortex energy separator, therefore the flow temperature increases.

Before passing through the boundary between the energy separation chamber 3 and the inner content of the valve case 4, the air flow rate increases due to the fact that the valve closes a significant part of the working area of the vortex energy separator. At this moment, the temperature along the entire perimeter of the section decreases, while the pressure stays constant. Consequently, the reverse transformation occurs: the heat energy is converted into kinetic energy.

Figure 4 shows that when air enters the internal volume of the valve case 4, the flow pressure decreases to atmospheric pressure, and the temperature rises again.

Here the flow becomes chaotic. As such, the part of the flow that exits through the valve case 4 is uneven in temperature and velocity, but uniform in pressure. However, the average temperature value is greater than the temperature from which the compressed air entered the tangential swirl inlet 2.

The residual air flow through the axial layers of the energy separation chamber 5 is directed to the diaphragm 6 located at the opposite end of the vortex energy separator. The flow has the pressure that is slightly higher than the pressure of the counter flow at this section, low temperature and speed. The latter is due to the fact that the volume of the flow moving to the opposite direction is less than the volume that filled the primary flow.

When this happens, on the surfaces of adjacent air flows a friction force arises, due to the adhesion of air particles to each other. It leads to a decrease in velocity of part of the flow that moves along the axial layers of the energy separation chamber 5, in relation to the air flow that moves along the peripheral layers 3.

At the outlet of the diaphragm 6, the temperature of the residual flow is several times lower than the temperature of that part of the flow that was previously separated
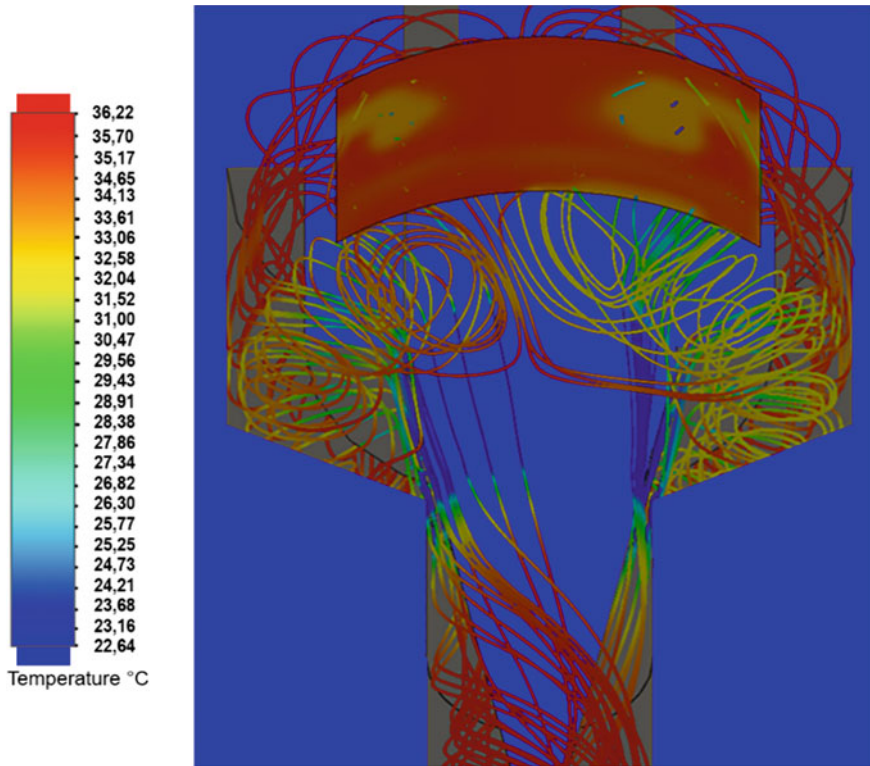
Fig. 4 Temperature distribution of the air flow when passing through the valve case

after passing through the tangential swirl inlet 2, but was not directed to the far outlet of the vortex energy separator. Due to the fact that the peripheral layers of the air flow leaving the diaphragm are swirling, the center relative to which the cold residual flow is slightly displaced relative to the axis of the energy separation chamber 5. The temperature distribution of the cold air flow at the diaphragm outlet is shown in Fig. 5.

Generally, it is impossible to form an idea of the point parameters of temperature separation in the vortex energy separator based on experimental data, since the internal parameters cannot be measured without distorting the vortex effect. Adequate computer model of the vortex energy separator allows measuring and evaluating parameters at any point of its model.

Thus, the CFD model of the considered vortex energy separator made it possible to display the trajectories of the air flow, thereby confirming the hypothesis [9], as well as to carry out a number of experiments to obtain quantitative characteristics of the output air flow. The CFD model confirmed the adequacy of the data obtained in [8] for the prototype setup (see Fig. 1).
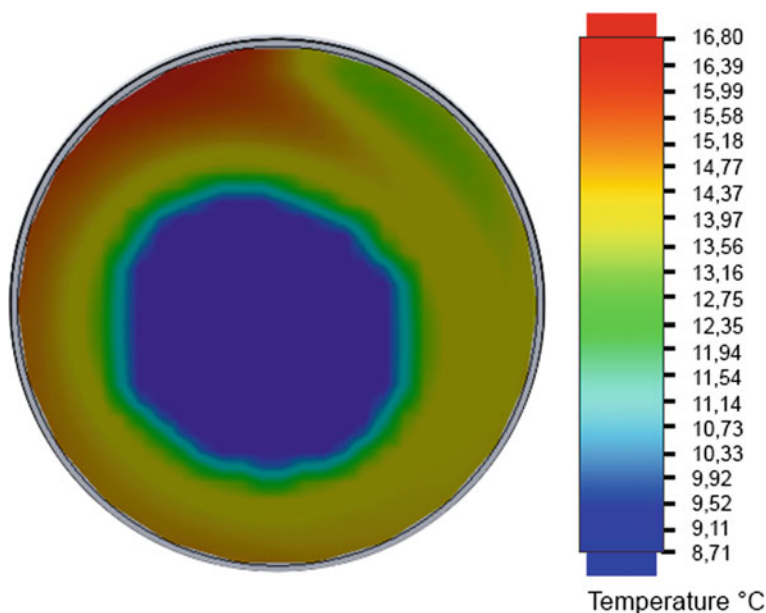
Fig. 5 Temperature distribution of the cold air flow at the outlet of the diaphragm

4 Results and Discussion

Based on the geometry of the vortex energy separator, experimental data were obtained using CFD modeling and represented in the form of temperature dependencies between cold and hot air flows and the adjusting valve position at values of the input air flow pressure of 5, 6, and 7 atm. in Figs. 6 and 7 respectively. Dependences are shown in Fig. 6. They have the nonlinear ambiguous character with evident extrema. This is also observed in investigations [10]. Figure 7 shows that the characteristics are also nonlinear.

The presented dependencies have the following explanation. When the adjusting valve moves, the area through which the air flow enters the internal volume of the valve case 4 increases, that leads to in hot air flow parameters decrease in Fig. 7. Another dependency can be observed in Fig. 7: the pressure increase leads to the temperature increase, and the adjusting valve position does not significantly affect this increment.

The volume of the air flow that leaves the vortex energy separator through the valve case 4 directly determines the residual volume of the air flow, which subsequently exits through the diaphragm 6 (the flow rate of the input compressed air flow is unchanged). The volume of the air flow moving to the diaphragm 6 decreases as the valve moves, while the area through which this flow returns to the energy separation chamber 5 increases. This leads to a decrease in velocity of this part of the flow and friction between two oppositely moving air flows so the cold air flow temperature
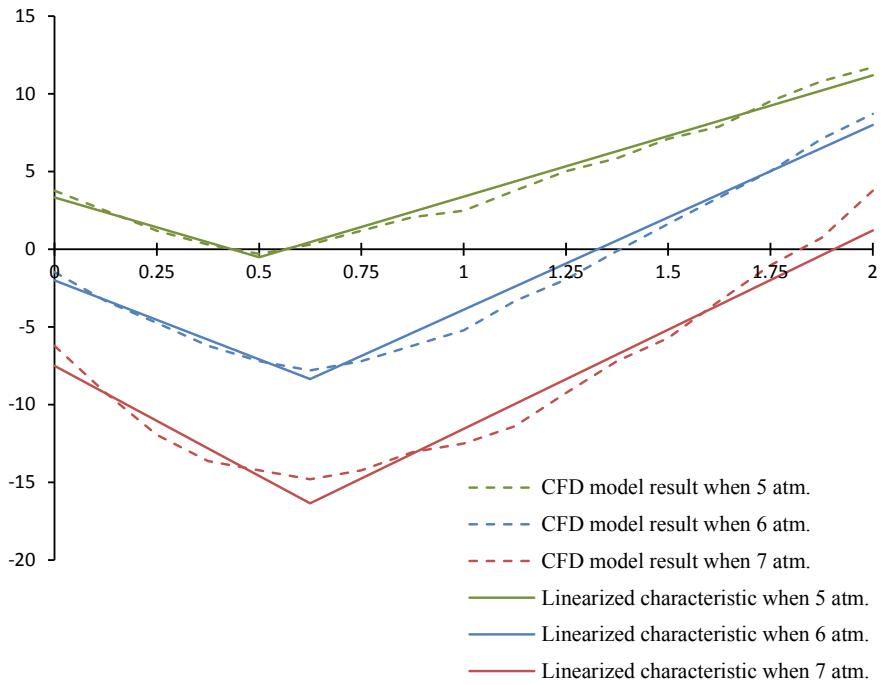
Fig. 6 Dependence of the cold air flow temperature (°C) on the adjusting valve position (mm)
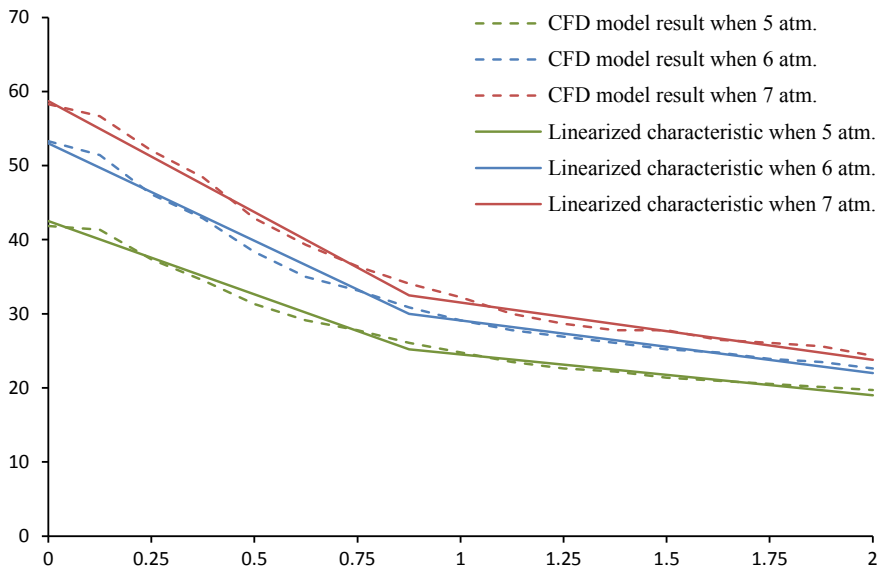
Fig. 7 Dependence of the hot air flow temperature (°C) on the adjusting valve position (mm)

decreases. However, there is a critical adjusting valve position ($\xi \approx 0.5-0.6$ mm), in which the volume of hot air outlet is large enough for the kinetic energy of the cold air to decrease significantly so the temperature of the cold air in the outlet starts to rise again. Figure 6 shows a decrease in temperature with increasing inlet air flow pressure, and the adjusting valve position does not affect this increment as well as the hot air temperature described above.

The simulation of the vortex energy separator was carried out with various parameters of the vortex flow, starting with the condition when the valve case 4 is completely closed and ending with the condition when the valve case 4 is completely opened in order to calculate the gain as the major parameter of mathematical model. The obtained nonlinear characteristics give the possibility to construct the simplest mathematical models of static transformation by dividing them into two pronounced intervals and performing the linearization of the presented dependencies.

Figures 6 and 7 show the results of the interval linearization of the dependences of the air flow temperature on the adjusting valve position at various pressure values over the entire range of input signal. The maximum linearization error is 2.56 °C, and the integral error is 26% for cold air flow temperature at 7 atm.

The following mathematical model attributes of the vortex energy separator have been obtained for the cold air flow at a pressure of 5 atm.:

- errors 0.44 and 0.9 °C;
- operation points {0.25; 1.4} and {0.75; 5.33};
- input signal ranges [0 0.5] and [0.5 2] mm;
- output signal ranges [3.32 −0.52] and [−0.52 11.18] °C;
- gains k_C −7.68 and 7.8 °C/mm;
- linearized characteristic equation $\Delta T_C = k_C \cdot \Delta \xi$,
- where $\Delta \xi$—increment of adjusting valve displacement relative to its initial
- position, mm, ΔT_C—increment of cold air flow temperature, °C.

The following mathematical model attributes of the vortex energy separator have been obtained for the hot air flow at a pressure of 5 atm.:

- errors 1.31 and 0.86 °C;
- operation points {0.4375; 33.85} and {1.4375; 22.1};
- input signal ranges [0 0.875] and [0.875 2] mm;
- output signal ranges [42.5 25.2] and [25.2 19] °C;
- gains k_H −19.77 and 5.51 °C/mm;
- linearized characteristic equation $\Delta T_H = k_H \cdot \Delta \xi$,
- where ΔT_H—increment of hot air flow temperature, °C.

In a similar way, the attributes of the linearization were obtained at pressures of 6 and 7 atm. for cold and hot air flows.

Thus, the presented approach to the static characteristics linearization made it possible to obtain the parameters of the mathematical model of the vortex energy separator for entire possible range of the adjusting valve position and all considered values of the input air flow pressure.

5 Conclusions

The complex air flow inside the vortex energy separator was investigated using the CFD model. The use of CFD gives the opportunity to understand the vortex effect physics and obtain adequate stationary processes which arise in vortex energy separator due to the temperature separation hypothesis. The simulation results were used to construct nonlinear static characteristics as dependencies between the input parameters (valve displacement and inlet air flow pressure) and output parameters (cold and hot flows temperature). Interval linearization of each characteristic over the entire range makes it possible to designate the operating modes of the considered vortex energy separator for certain values of the input flow pressure and the adjusting valve position without significant loss of accuracy.

Maximum linearization error in the dependence of the hot air flow temperature on the valve position is 1.7°C at 7 atm. Maximum linearization error in the dependence of the cold air flow temperature on the valve position is 2.56°C at 7 atm.

Interval linear mathematical models of stationary processes in the prototype of vortex energy separator are obtained as mathematical models of the automatic stabilization object, which allow to synthesize algorithms for outlet air flows temperature stabilization in steady-state operation mode.

References

1. Gupta US, Chaturvedi A, Patel N, Pandey NK, Patel N (2017) A review on vortex tube refrigeration and applications. Int J Adv Res Sci Eng 6(9):167–175
2. Xue Y, Arjomandi M, Kelso RM (2013) The working principle of a vortex tube. Int J Refrig 36(6):1730–1740
3. Zhmud VA, Dimitrov L (2017) Designing of the precision automatic control systems. LAP LAMBERT Academic Publishing, Saarbrücken
4. Khalatov AA, Avramenko AA, Shevchuk IV (2000) Heat transfer and hydrodynamics in the fields of centrifugal mass forces, vol 3. Swirling threads. Inst. tekhn. teplofiziki NAN Ukrainy
5. Alekseenko VP, Stukalov AS, Yakimov PYu (2009) Calculation of the characteristics of the vortex temperature control systems. Izvestiya Samarskogo nauchnogo tsentra Rossiiskoi akademii nauk 11(5):170–176
6. Salem TK, Khosroshahi FS, Aric M, Hamdan MO, Budakli M (2017) Numerical and experimental analysis of a heat-pipe-embedded printed circuit board for solid state lighting applications. Experim Heat Transfer 32(1):1–13
7. Xue Y, Arjomandi M, Kelso RM (2014) Energy analysis within a vortex tube. Experim Thermal Fluid Sci (EXP THERM FLUID SCI) 52:139–145
8. Kulik AS, Dzhulgakov VG, Pasichnik SN (2010) Hardware and software complex for investigating the vortex effect. Visnyk KhNTUSH 102:85–87
9. Gutsol AF (1997) The Ranque effect. Phys Usp 167(6):639–658
10. Dhillon AK, Bandyopadhyay SS (2015) CFD analysis of straight and flared vortex tube. IOP Conf Ser Mater Sci Eng 101(1)

Numerical Analysis of the Gas Turbine Rotor Blades Thermal State Using a Refined Mathematical Model

Serhii Morhun

Abstract In this paper the problem of gas turbine rotor thermal state has been studied. The physical model of the gas flow and heat transfer in the turbine rotor can be formulated as the flow of a viscous heat-conducting gas through a row of impellers with non-uniformity of flow parameters at the rotor inlet and outlet. Such processes can be correctly described by using the system of Navier–Stokes equations for a compressible heat-conducting continuum with the correct determination of turbulent components. For this problem solution, the new, more correct mathematical model on the base of the hexagonal finite elements has been designed. Taking this mathematical model into consideration, the flow temperature in the rotor root and peripheral sections have been calculated. The temperature fields on the inlet and outlet impeller blades have also been determined. The obtained results can be applied in the further study of the gas turbine rotor stress–strain state.

Keywords Gas turbine engine · Turbine rotor blades · Hexagonal finite elements · Heat exchange · Blades surface temperature field · Navier–Stokes equations system

1 Relevance of the Research

1.1 Factors, Influencing the Heat Transfer in the Gas Turbine Rotor

One of the most challenging tasks in the design of gas turbines is the task of determining the heat flows from the gas to the blades. The processes of heat and mass transfer in the near-wall regions of rotor blades are characterized by the complex effect of a number of factors: turbulence, separation of the boundary layer, transient flow regimes from laminar to turbulent flow, flow gradient and compressibility, surface roughness; the influence of inertial forces etc.

S. Morhun (✉)
Admiral Makarov National University of Shipbuilding, Mykolaiv, Ukraine
e-mail: serhii.morhun@nuos.edu.ua

1.2 Methods of Rotor Blades Surface Temperature Determining

There are various methods for the heat transfer calculating in gas turbine rotors. But it is rather difficult to find out an exact method that would make it possible to determine the conditions for external heat transfer between the gas and the blade, taking into consideration all of the above factors. Therefore experimentally substantiated generalizations (criterion dependences), approximate integral calculation methods, as well as numerical methods for solving the boundary layer equations, and averaged Navier–Stokes equations (Reynolds equations) are used [1, 2].

It is practically impossible to take into account all of the above factors during calculating external heat transfer. At the same time, the determination of the heat transfer coefficient from the gas flow to the blade wall should be carried out with the highest possible accuracy, since an error in setting the heat transfer coefficient as a boundary condition for external heat transfer leads to an error in determining the temperature of the blade surface, which, in turn, can lead to a significant error in determining the rotor stress–strain state.

The full-scale experimental methods of the external heat transfer used for the fairly accurate determination of the blades temperature are given in literature [3−6]. However, the process of a full-scale experiment is very expensive, therefore, it is necessary to solve the problem of determining the conditions of external heat transfer at the design stage. Progress in the development of computer technology, together with the latest advances in numerical methods makes it possible to determine the temperature on the turbine blades surface [1, 5−9].

Based on the analysis of the above literature sources, it can be concluded that at the moment there are still a number of unsolved problems regarding the improvement of the determination of the temperature field of the gas turbine rotors blades. Thus, the development of a refined mathematical model for determining the temperature field on the surface of the turbine rotor blades is actual.

2 Research Object and Aim

The research object is the gas flow through the turbine rotor. Turbine rotor can be considered as a system of blades of the same shape spaced on the surface of impellers. The impellers form an assembly by means of a special shaft. Flowing through the impellers, the gas flow changes the speed and direction of its movement. In stationary nozzle blades, the potential energy of the flow is converted into kinetic energy to ensure the required speed and direction of the flow. In the rotating impellers, the kinetic energy of the flow is converted into mechanical energy of rotation.

On the basis of the above-described real process, the physical model of the flow and heat transfer in the turbine rotor can be formulated as the flow of a viscous heat-conducting gas through a row of impellers with non-uniformity of flow parameters at

the inlet, separation phenomena, compressibility effects, turbulence, and changes in the parameters of the boundary layer on the blade surface. For a correct description of the parameters of heat and mass transfer, taking into account the above factors, it is necessary to develop a refined mathematical model that describes complex physical, thermo- and gas-dynamic processes in the turbine rotor.

Thus the aim of the research is to develop a refined mathematical model for calculating the temperature field on the gas turbine blades surface by solving a system of averaged Navier–Stokes equations and a model of turbulent gas flow.

To achieve this aim, the following tasks must be solved:

- To develop a refined mathematical model for determining the external heat transfer on the turbine blades based on the solution of the system of averaged Navier–Stokes equations and its verification;
- Determine the temperature fields on the surface of the turbine rotor blades.

3 Formulation of the Problem

As the stationary coordinate system, the Cartesian right-handed *xyz* coordinate system with the center at point *O* located on the gas turbine engine axis is taken. The *z* axis is perpendicular to the turbine axis, and the *x* axis coincides with this axis. The rotating coordinate system rotates together with the turbine rotor at a constant angular velocity Ω equal to the angular velocity of rotor.

3.1 The Navier–Stokes Equations System for a Compressible Heat-Conducting System

The system of Navier–Stokes equations for a compressible heat-conducting continuum can be written in the form:

$$
\begin{aligned}
&\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho \tilde{U}_j\right) = 0; \\
&\frac{\partial}{\partial t}\left(\rho \tilde{U}_i\right) + \frac{\partial}{\partial x_j}\left(\rho \tilde{U}_i \tilde{U}_j\right) = -\frac{\partial p}{\partial x_i} - \frac{\partial}{\partial x_j}\left(\tau_{ij} + \rho u'_i u'_j\right); \\
&\frac{\partial}{\partial t}\left(\rho \tilde{H}\right) + \frac{\partial}{\partial x_j}\left(\rho \tilde{U}_j \tilde{H}\right) = \frac{\partial p}{\partial t} - \frac{\partial}{\partial x_j}\left(U_i \tau_{ij} + u'_i \tau'_{ij} + Q_j + \rho u'_j H'\right),
\end{aligned}
\tag{1}
$$

where t is the time, s; ρ is the density, kg/m^3; p is the pressure, Pa; $\tilde{U}_i$ and u'_i are the components of the average and pulsation velocity; τ'_{ij} is the Reynolds stress tensor; $\rho \cdot u'_j \cdot h'$ is the heat flux due to turbulent transfer.

Symbol " ~ " refers to time and density averaged parameters; symbol "/" refers to ripple components.

$\tau'_{ij} = -\frac{2}{3}\mu\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right)$—the average stress tensor; $Q_j = -\lambda\frac{\partial T}{\partial x_j}$—the average heat flux.

The term $\rho u'_i u'_j$ characterizes the effect of the flow turbulence. It has the mathematical form of a second-order tensor and acts like a stress. Therefore, this tensor is called the Reynolds stress tensor. The term $\rho u'_j h'$ characterizes the heat transfer due to turbulent pulsations.

To close the system of Eq. (1), it is necessary to express the unknown turbulent pulsation components. For these purpose we use the equation of state in form:

$$p = \rho \cdot R \cdot T. \tag{2}$$

3.2 The Turbulent Components Determination

The Reynolds stress tensor is expressed by the following relationship:

$$-\rho u'_i u'_j = \mu_i\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right) - \frac{2}{3}\left(\mu_i\frac{\partial U}{\partial x} + \rho k\right), \tag{3}$$

$$k = \rho\frac{u'_i u'_j}{2},$$

where k is the turbulent kinetic energy.

Turbulent heat flow h is modeled in accordance with the hypothesis of eddy gradient diffusion [1, 7, 10−15]:

$$-\rho u'_i h = \frac{\mu_i}{Pr}\frac{\partial h}{\partial x}, \tag{4}$$

where Pr, is the turbulent Prandtl number.

The proportionality coefficient is selected on the dimensions theory base. The gradient diffusion hypothesis implies a relationship between the turbulent transfer of mass or heat and the gradient of the transferred substance.

The term responsible for the work of viscous forces due to turbulent pulsations is approximated as follows:

$$\frac{\partial\left(u'_i \tau'_{ij}\right)}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\mu\frac{\partial k}{\partial x_j}\right). \tag{5}$$

Relations (3)−(5) express the turbulent pulsation terms in the Reynolds equations only as functions of time-averaged quantities and they are could be determined only if the turbulent viscosity μ and the turbulent kinetic energy k are known. So, the

Reynolds equation system (1) can be rewritten into form:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho \cdot U_j\right) = 0;$$

$$\frac{\partial}{\partial t}\left(\rho \cdot U_i\right) + \frac{\partial}{\partial x_j}\left(\rho \cdot U_i \cdot U_j\right) = -\frac{\partial p^*}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\mu\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right) - \frac{2}{3}\mu\frac{\partial U_i}{\partial x_i}\right);$$

$$\frac{\partial}{\partial t}\left(\rho \cdot H\right) + \frac{\partial}{\partial x_j}\left(\rho \cdot U_j \cdot H\right) - \frac{\partial p}{\partial t} = \frac{\partial}{\partial x_j}\left(\lambda\frac{\partial T}{\partial x_j} + \frac{\mu}{\Pr}\frac{\partial h}{\partial x_j}\right)$$

$$+ \frac{\partial}{\partial x_j}\left\{U_i\left[\mu\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} - \frac{2}{3}\mu\frac{\partial U_i}{\partial x_i}\right)\right] + \frac{\partial k}{\partial x_j}\right\}, \tag{6}$$

where $p^* = p + \frac{2}{3}\rho k$; $H = h + \frac{1}{2}U_i U_j$.

By analogy with the kinetic theory of gases, turbulent viscosity can be represented in the form proposed by Prandtl and Kolmogorov [5, 12, 16−20]:

$$\mu = \rho \cdot c_\mu \cdot L \cdot v, \tag{7}$$

where $c = 0.09$ is an empirical constant, L and $v = \sqrt{k}$ are the characteristic scales of the length and velocity of turbulence.

From (7) it follows that to determine the turbulent viscosity it is sufficient to know the values of two parameters characterizing the scales of the length and velocity of turbulence. Also it is more useful to take into consideration the turbulence kinetic energy dissipation rate (ε) or turbulence kinetic energy pulsation frequency (ω):

$$\varepsilon = \frac{\sqrt[3]{k^2}}{L}; \quad \omega = \frac{\varepsilon}{k}, \tag{8}$$

To simulate turbulence and correctly describe the dynamic and thermal boundary layer on the outer surface of a turbine blade, the semiempirical *SST* (Shear Stress Transport) Menter model is used [7, 21, 22]. This model combines the advantages of the *k-ε* and *k-ω* turbulence models and is a combination of them.

4 Solution to the Problem

Analytical solutions of equations system (6) does not exist, therefore, to obtain solutions for the case of a real gas flow in a turbine, it is necessary to use numerical methods, namely, the finite volume method.

Consequently, the spatial computational region of the blades channel of the turbine rotor is covered with a discrete mesh, which is divided into finite elements of the hexagonal type [23].

Using the FEM dependencies, the system of Eq. (6) is transformed into a system of matrix equations in the considered finite element. The shape functions for a hexagonal finite element are given here:

$$
\begin{aligned}
&N_1 = (1 - \eta) \cdot (1 - \xi) \cdot (1 - \zeta); \ N_2 = \eta \cdot (1 - \xi) \cdot (1 - \zeta); \\
&N_3 = \eta \cdot \xi \cdot (1 - \zeta); \qquad\qquad\quad N_4 = (1 - \eta) \cdot \xi \cdot (1 - \zeta); \\
&N_5 = (1 - \eta) \cdot (1 - \xi) \cdot \zeta; \qquad\ N_6 = \eta \cdot (1 - \xi) \cdot \zeta; \\
&N_7 = \mu \cdot \xi \cdot \zeta; \qquad\qquad\qquad\quad\ N_8 = (1 - \eta) \cdot \xi \cdot \zeta,
\end{aligned}
\tag{9}
$$

where η, ξ, ζ are the finite element local coordinates.

There are the dependencies between the finite element local coordinate system (9) and the turbine rotor global coordinate Cartesian coordinate system:

$$
\begin{bmatrix} \frac{\partial N}{\partial x} \\ \frac{\partial N}{\partial y} \\ \frac{\partial N}{\partial z} \end{bmatrix}
=
\begin{bmatrix}
\frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \zeta} \\
\frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \zeta} \\
\frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \zeta}
\end{bmatrix}^{-1}
\begin{bmatrix} \frac{\partial N}{\partial \eta} \\ \frac{\partial N}{\partial \xi} \\ \frac{\partial N}{\partial \zeta} \end{bmatrix}.
\tag{10}
$$

Taking into account the shape functions, the temperature matrix of any finite element modeling the surface of the turbine rotor blades is determined as follows:

$$
T_e = [N_i]^T \cdot \{T_i\}^T \quad (i = 1...n),
\tag{11}
$$

where T_i are temperatures at the nodes of the considered finite element.

The global matrix describing the temperature field of the blade surface is obtained using the superposition method from the temperature matrices of individual finite elements T_e.

5 Main Results and Their Analysis

The power of the investigated gas turbine is 25 MW. The gas temperature at the entrance to the turbine rotor is 1583 K; total pressure at the stage inlet 0.7578 MPa; rotor speed 9390 rev/min.

5.1 The Flow Temperature Structure in the Rotor Axisymmetric Section

The spatial structure of the flow temperature is represented by axisymmetric sections in the root (Fig. 1a) and peripheral (Fig. 1b) sections of the turbine rotor.

Analyzing the data, presented on the Fig. 1 we can find a sharp temperature gradient between the root and peripheral blade sections. The matter is that the gas flow after passing the inlet firstly penetrates the peripheral part of the rotor. Only after heat exchange between the flow and periphery the whole blade is heated up to the root.
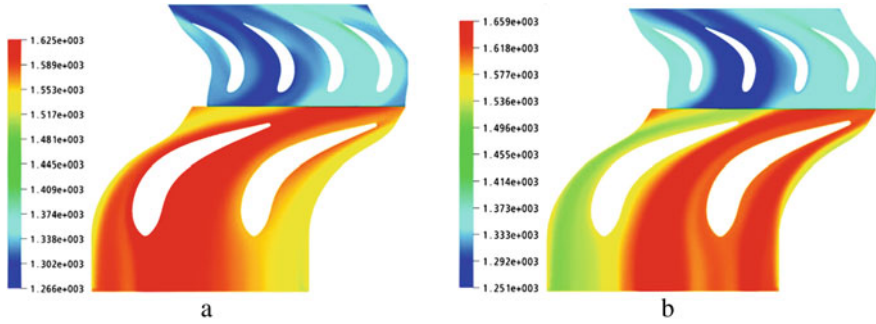
Fig. 1 Temperature field in the turbine rotor sections: **a** root section; **b** peripheral section
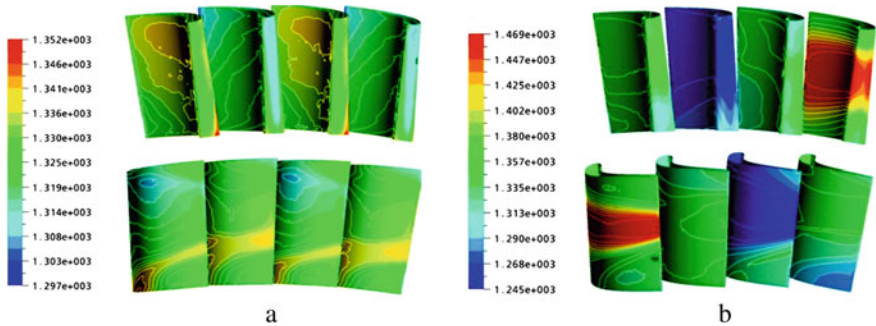
Fig. 2 Temperature field on the turbine rotor blades surface: **a** outlet impeller; **b** inlet impeller

5.2 *Temperature Fields on the Turbine Rotor Blades Surface*

But due to the blade construction and twisting of the rotor, the cross-sectional area decreases from the blades root to their periphery. Thus the temperature field has a gradient in the radial section too. The temperature field on the inlet and outlet impellers blades surface is presented on the Fig. 2.

Analyzing data, presented on the Fig. 2, we need to highlight that the temperature field for both impellers has sharp gradients. The main temperature gradients are located from the blades back side in the peripheral area. Such an inconsistency of temperatures produces an additional temperature stresses and demands to take them into consideration while determining the turbine rotor stress–strain state.

6 Conclusion

The analysis performed shows the relevance of the gas turbine rotor blades temperature fields determining. Using the Navier–Stokes equations for the case of turbulent

viscous flow, a refined mathematical model on the base of hexagonal finite elements has been developed. Using this mathematical model, the fields of flow temperatures in the rotor axisymmetric section and the turbine blades surface temperature on the rotor inlet and outlet impellers have been calculated. The obtained results can be applied in the further study of the gas turbine rotor stress–strain state.

References

1. Vassiliev V, Kostege V, Schnedler P, Trishkin A (2008) Thermal state analysis of industrial gas turbine blades. In ASME Turbo expo: power for land sea and air proceedings, pp 813−824. https://doi.org/10.1115/GT2005-68951
2. Wang S, Shouzuo L, Lei L, Zhao Zh, Wei D, Bengt S (2021) A high temperature turbine blade heat transfer multilevel design platform. Numer Heat Transfer Part A Appl 79(2):122−145. https://doi.org/10.1080/10407782.2020.1835104
3. Luo L, Wang CL, Wang L, Sunden B, Wang ST (2016) Heat transfer and friction factor performance in a pin fin wedge duct with different dimple arrangements. Numer Heat Transfer Part A Appl 69(2):209−226. https://doi.org/10.1080/10407782.2015.1052301
4. Luo L, Qiu D, Du W, Sunden B, Wang Zh, Zhang X (2018) Surface temperature reduction by using dimples/protrusions in a realistic turbine blade trailing edge. Numer Heat Transfer Part A Appl 74(5):1265−1283. https://doi.org/10.1080/10407782.2018.151533
5. Das R, Kundu B (2017) Prediction of heat generation in a porous fin from surface temperature. J Thermophys Heat Transfer 31(4):781–790. https://doi.org/10.2514/1.T5098
6. Reihani MR, Alizadeh M, Fathi A, Khaledi H (2013) Turbine blade temperature calculation and life estimation—a sensitivity analysis. Propulsion Power Res 2(2):148–161. https://doi.org/10.1016/j.jppr.2013.04.004
7. Shevchenko I, Rogalev N, Rogalev A, Vegera A, Bychkov N (2018) Verification of thermal models of internally cooled gas turbine blades. Adv Fluid Dynamics Turbomachinery 2018. https://doi.org/10.1155/2018/6780137
8. Sadowski T, Pietras D (2016) Heat transfer process in jet turbine blade with functionally graded thermal barrier coating. Solid State Phenomena 254:170−175. https://doi.org/10.428/www.scientific.net/SSP.254.170
9. Mukherji BPVS, Raja K, Naresh G, Rao VNB, Kumar INN (2015) An investigation of transient thermal analysis of 1st stage gas turbine blade manufactured by directional solidification and mechanically alloyed nickel-based superalloys. Int J Adv Sci Technol 85:17−28. https://doi.org/10.14257/ijast.2015.85.03
10. Singh P, Shukla OP (2015) Heat transfer analysis of gas turbine rotor blade through staggered holes using CFD. Int J Eng Res General Sci 4(2):263−271
11. Luo L, Wang C, Wang L, Sunden B, Wang ST (2016) Endwall heat transfer and aerodynamic performance of bowed outlet guide vans with on- and off-design conditions. Numer Heat Transfer Part A Appl 69(4):352–368. https://doi.org/10.1080/10407782.2015108102
12. Luo L, Sunden B, Wang ST (2015) Optimization of the blade profile and cooling structure in a gas turbine stage considering both the aerodynamics and heat transfer. Heat Transfer Res 46(7):599–629. https://doi.org/10.1615/HeatTransRes2015012370
13. Taslim ME, Nongsaeng A (2011) Experimental and numerical cross-over jet impingement in an airfoil trailing edge cooling channel. J Turbomach 133(4):1–10. https://doi.org/10.1115/1.4002984
14. Su WT, Binama M, Li Y, Zhao Y (2020) Study on the method of reducing the pressure fluctuation of hydraulic turbine by optimizing the draft tube pressure distribution. Renew Energy 162:550–560. https://doi.org/10.1016/j.renene.202008057

15. Brahmaiah KH, Kumar ML (2014) Heat transfer analysis of gas turbine blade through cooling holes. Int J Comput Eng Res 04(7):2250–3005
16. Singh K, Das R, Kundu B (2016) Approximate analytical method for porous stepped fins with temperature-dependent heat transfer parameters. J Thermophys Heat Transfer 30(3):661–672. https://doi.org/10.2514/1.T4831
17. Wang J, Liu J, Wang L, Sunden B, Wang S (2018) Conjugated heat transfer investigation with racetrack-shaped jet hole and double swirling chamber in rotating jet impingement. Numer Heat Transfer Part A Appl 73(11):768–787. https://doi.org/10.1080/10407782.2018.1454753
18. Yang L, Ren J, Hongde J, Luan Y, Ligrani P (2015) Assessment of six turbulence models for modeling and predicting narrow passage flows. Part 1: impingement jets. Numer Heat Transfer Part A: Appl 69(5):109−127. https://doi.org/10.1080/10407782.2018.1069665
19. Yang L, Ren J, Jiang H, Luan Y, Ligrani P (2014) Experimental and numerical investigation of unsteady impingement cooling within a blade leading edge passage. Int J Heat Mass Transf 71:57–68. https://doi.org/10.1046/ijheatmasstransfer.2013.12.006
20. Furlani L, Armellini A, Casarsa L (2016) Rotational effects on the flow field inside a leading edge impingement cooling passage. Exp Thermal Fluid Sci 76:57–66. https://doi.org/10.1016/j.expthermflusci.2016.03.004
21. Yershov SV, Yakovlev VA (2021) The influence of mesh resolution on 3D RANS flow simulations in turbomachinery flow parts. J Mech Eng 24(1):13−27. https://doi.org/10.15407/pmach2021.01.013
22. Morhun S (2017) Improving the mathematical models applied for the solution of solid assembly constructions thermoelasticity problem. J Mech Eng 20(2):42−46. https://doi.org/10.15407/pmach2017.02.042
23. Morhun S (2019) Numerical analysis of working processes in the blade channels of the highly loaded turbine of a marine gas turbine engine, using a refined finite element model. J Mech Eng 22(3):14−20. https://doi.org/10.15407/pmach2019.03.014

Elliptical Methods for Surface Meshing

Larysa Khalanchuk and Serhii Choporov

Abstract A variety of engineering processed could be modelled by differential equation with partial derivatives. Typically, differential equations for complex domains are solved by numerical methods. The initial step of such methods often requires a discrete representation of the domain by a grid or a mesh. In many practical cases domains could be combinations of cylindrical or conic shells. Such domains could be meshed using structured or block-structured grids. The general strategy is an adaptation of the grid to the domain's shape. Such adaptation should process properties of the problem, e.g. forces. In this case the grid should be non-uniform with the increased density of elements in some regions. The elliptical method is a general method for a mesh adaptation. This method uses elliptical differential equations with partial derivatives to handle density of elements over the domain. In this article, we developed the elliptical method for cylindrical and conic shells meshing. Controlling functions are used to manage density of elements in the mesh increasing it near specified coordinate lines.

Keywords Discrete Model · Grid · Poisson equation · Control functions · Thickening of grid · Surface of the body rotation

1 Introduction

Taking into account the intensive development of computer technologies, mathematical modeling of various physical processes acquires special importance. Many such processes are described by equations with partial derivatives, the analytical solution of which sometimes cannot be obtained through the complexity of the equations obtained, complex geometry of the domain or other properties of the problem. Therefore, numerical methods for solving equations with partial derivatives are widely used

L. Khalanchuk (✉)
Dmytro Motornyi Tavria State Agrotechnological University, Melitopol, Ukraine

S. Choporov
Zaporizhzhya National University, Zaporizhzhya, Ukraine

in various fields of science and technology. Many numerical methods build the solution using a discrete representation of a domain. The discrete representation is also known as a mesh or a grid and related to subdivision of the domain into the finite number of simple shapes (elements). A grid is a set of elements that might only have intersections in common nodes and edges. The problem of gridding includes adaptation of nodes to the domain shape and often to the problem properties. For example, density of elements in the grid could increase in regions loaded by high forces. The gridding problem arises, for example, in space engineering problems where domains are typically cylindrical or conic shells. Such shells could be modelled by structured or block structured grids. Elliptical methods allow to use partial differential equations to manage the density of elements in the structured grid. Such methods require to define controlling functions and solution of an elliptical equation.

2 Literature Analysis and Problem Statement

To solve differential equations with partial derivatives, the adaptive strategy is often desirable, one of the important approaches of which is the method of deformation of the grid [1]. It applies to various physical and engineering problems, such as combustion, shock waves, diffusion reactions and two-phase streams [2–4].

When the main problems depend on time, deformed grids vary in the process of changing time, and thus the method is called the method of moving mesh. In the method of finite elements of a moving grid, the grid changes in the computing region is used to control the redistribution of the grid in the physical region at each time step. Then the solution is updated by a certain formula. This method was used in many applications [5].

Elliptical equations with discontinuous coefficients and separate sources are of great importance in the dynamics of liquids, material science and biological systems. To solve such equations, methods based on finite volume [6], lump-polynomial sampling [7], the method of communication interface [8] can be used [7]. The method of agreed interface (MIB—The Matched Interface and Boundary) was originally introduced by Zhao and Wei to solve Maxwell equation with material interfaces [9], and later generalized to solving elliptical equations with distant coefficients and special sources [10–13], as well as the Helmholtz equation with material interfaces [14]. Most approaches are usually used simple regular Cartesian grids. Such grids are definitely not optimal for problems with localized dramatically different solutions. One of the optimal strategies of localized cardinal changes is to clarify the local grid. Methods of interface based on finite elements and final volumes can embed locally adaptive grid generation algorithms [15]. However, it is very difficult to construct convergent methods of convergent finite elements or final volumes based on complex interfaces, in particular interfaces with geometric features [10, 16].

Analysis of approaches to the deformation of the grid is considered in [17]. Most studies relating to the deformation method of linear elasticity, the method of finite elements was used to sample the equations of linear elasticity, and then solved the

resulting linear system a generalized minimum residual method (GMRES) [18]. The deformation of the grid by interpolation analogies can be used to arbitrary types of grids that contain common multifaceted elements or hanging nodes [19]. Interpolation schemes achieve higher computing efficiency and smaller memory requirements compared to physical schemes [20–22]. However, any interpolation process is associated with a certain error field.

In the study [23], the 3d AGARD enclosure body has been verified as a problem with a complete fluid and structure, relatively the deformation results of the weighing method at the reverse distance (IDW) and the interpolation of the radial baseline function (RBF). The method of interpolation of the radial baseline function can be used as an interpolation function for transferring movements known to the structure of the structural grid, in a grid of a liquid that produces high-quality grids with reasonable orthogonality preservation near deforming boundaries [24]. In the article [25], it is proposed in the method of interpolation of the radial baseline function of the technique of maintenance, which limits the deformation of the grid to the surrounding area of a moving surface.

In [26, 27] methods of deformation of a structured grid using control functions of the Differential Equation of Poisson are considered. Investigation of condensation of mesh nodes through control functions specifying the region of diagonal lines, circle, sinusoids with the equation of Euler used in gas dynamics are shown in the article [28]. Construction of a structured grid surface that simulates the density of the probability of finding an electron in a given area of the quantum point of different species due to the solution of the Schrödinger wave equation is executed in [29].

Taking into account the widespread use of structured nets in applied problems, it can be argued that the development of methods of generating discrete models, finite elements of which are thickened in places of high forces concentration and in places where the design has a special form, is an urgent task.

3 Materials and Methods of Research

3.1 The Purpose and Tasks of Research

The objective of this article is to develop elliptical method for structured grid adaptation in cylindrical and conic domains.

3.2 Materials of Research

In structured grids, a neighborhood between nodes allows to define neighbor elements. The design of such grids makes it easy to increase the number of nodes for the assessment of convergence, errors and to increase the accuracy of numerical

methods of solving boundary problems. Differential methods based on elliptical and parabolic equations give smooth internal coordinate lines, so it is possible to build orthogonal lines and thickening lines.

Consider the elliptical method on the example of an elliptic equation. For the curvilinear estimated area, the conversion of coordinates, which allows the curvilinear physical region in the coordinate system (x, y), is used to convert the grid, which allows the curvilinear physical region in the coordinate system (x, y) to transfer to a rectangular estimated area in the system (ξ, η). The connection between physical and estimated regions is determined by dependencies:

$$x = x(\xi, \eta), \quad y = y(\xi, \eta). \tag{1}$$

In the simplest generalized form, taking into account the transformation from the physical to the estimated area of coordinates (1), we have the Laplace equation for generation a structured grid. To obtain a thickening of the grid in the desired areas, control functions P (ξ, η), Q (ξ, η) and the Poisson equation:

$$\nabla^2 \xi = P(\xi, \eta), \quad \nabla^2 \eta = Q(\xi, \eta). \tag{2}$$

Since grid lines are set in space (ξ, η), then it is necessary to obtain dependencies x = x (ξ, η), y = y (ξ, η), so dependent and independent variables in Eq. (2) need to be changed.

Solution (2) in the estimated area of the coordinate system (ξ, η) is given the form:

$$\begin{cases} g_{22}\frac{\partial^2 x}{\partial \xi^2} - 2g_{12}\frac{\partial^2 x}{\partial \xi \partial \eta} + g_{11}\frac{\partial^2 x}{\partial \eta^2} + g\left(P(\xi, \eta)\frac{\partial x}{\partial \xi} + Q(\xi, \eta)\frac{\partial x}{\partial \eta}\right) = 0, \\ g_{22}\frac{\partial^2 y}{\partial \xi^2} - 2g_{12}\frac{\partial^2 y}{\partial \xi \partial \eta} + g_{11}\frac{\partial^2 y}{\partial \eta^2} + g\left(P(\xi, \eta)\frac{\partial y}{\partial \xi} + Q(\xi, \eta)\frac{\partial y}{\partial \eta}\right) = 0, \end{cases} \tag{3}$$

where

$$g_{11} = \left(\frac{\partial x}{\partial \xi}\right)^2 + \left(\frac{\partial y}{\partial \xi}\right)^2, g_{12} = \frac{\partial x}{\partial \xi}\frac{\partial x}{\partial \eta} + \frac{\partial y}{\partial \xi}\frac{\partial y}{\partial \eta}, \\ g_{22} = \left(\frac{\partial x}{\partial \eta}\right)^2 + \left(\frac{\partial y}{\partial \eta}\right)^2, g = \left(\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial y}{\partial \xi}\frac{\partial x}{\partial \eta}\right)^2. \tag{4}$$

In general, Eq. (3) are nonlinear, so they are solved numerically using, for example, difference schemes. Consider the condensation of nodes to a certain coordinate line, you can use control functions that can be given formulas:

$$P(\xi, \eta) = -\sum_{n=1}^{N} a_n \frac{(\xi - \xi_n)}{|\xi - \xi_n|} e^{-c_n|\xi - \xi_n|}, \tag{5}$$

$$Q(\xi, \eta) = -\sum_{n=1}^{N} a_n \frac{(\eta - \eta_n)}{|\eta - \eta_n|} e^{-c_n|\eta - \eta_n|}, \tag{6}$$

where N is the number of lines (coordinate lines $\xi = \xi_n$ and $\eta = \eta_n$), near which the grid should be thickened, and the coefficients a_n, c_n- positive parameters.

The function (5) leads to the combination of lines $\xi = const$ to line $\xi = \xi_n$, and the function (6) leads to the combination of lines $\eta = const$ to line $\eta = \eta_n$. Influence of parameters of control functions (5) and (6) on the quality of the grid during 2D modeling was considered in articles [26, 27].

Consider the influence of parameters of control functions (5) and (6) to a grid that is a surface model. We perform a survey of a surface modeling a cylinder and a cone. Since cylindrical and conical forms are widespread in engineering structures, then the study of thickening of the grid in the places of high forces concentration is an up-to-date problem. From a mathematical point of view, the scanning of the cylinder is a rectangle, so the thickening of the grid on the surface of the cylinder is performed initially on the scan, and then a spatial figure is formed. We perform a grid generation in a Scilab program package using formulas (3)–(4), and condensation by vertical lines by formula (6). As a result of calculations we build condensation on the surface of the cylinders (Fig. 1) followed by the conditions of the formula (6):

$$N = 1, \quad a_n = 5, \quad c_n = 1, \quad \eta_n = 0.3. \tag{7}$$

$$N = 1, \quad a_n = 5, \quad c_n = 1, \quad \eta_n = 0.8. \tag{8}$$

For a conical surface, we carry out a similar to a cylinder generation of a structured grid using Eqs. (3)–(4). As a result of calculations, we build condensation on the surface of the cones (Fig. 2), followed by the conditions of formula (6):

$$n = 20, \quad N = 1, \quad a_n = 5, \quad c_n = 1, \quad \eta_n = 0.2. \tag{9}$$

$$n = 20, \quad N = 1, \quad a_n = 5, \quad c_n = 1, \quad \eta_n = 0.7. \tag{10}$$
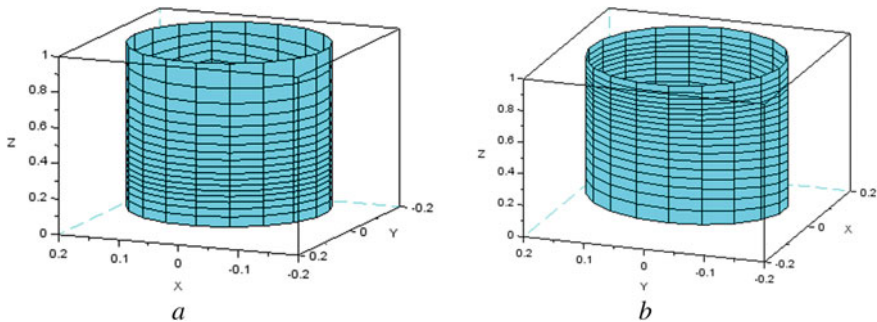
Fig. 1 Creation of the cylinder surface grid: **a** thickening condition (7), **b** thickening condition (8)
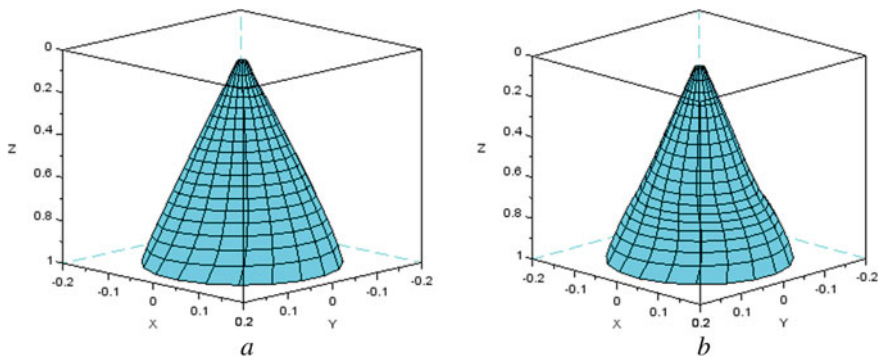
Fig. 2 Thickening of the grid of the cone surface: **a** thickening condition (9), **b** thickening condition (10)

In order to study the influence of parameters of control functions on the quality of the grid, compare the results of the generation for different values of parameters a_n, c_n with $\eta_n = 0.5$, the latter value will ensure better visibility. Also, for the clarity of thickening of the grid, we arrange a tapered surface to the top of the bottom. Consequently, for Eq. (6) we take the following general conditions:

$$n = 20, \quad N = 1, \quad \eta_n = 0.5. \tag{11}$$

Then we can observe the grid changes due to the change of parameters for cylindrical (Fig. 3) and conical surfaces (Fig. 4).

According to the results of our research (Figs. 3 and Fig. 4) we can conclude that an increase in the second parameter makes a grid close to orthogonal.

In practice, the most frequent thickening of the grid occurs in the places of high forces concentration, which may be a place of connection of structures of various forms. During the mathematical modeling of a given process, there is a problem of combining the nodes of the grid in places of connection of the design, the thickening of the grid adds a complication to the process of combining nodes. In aircraft construction and rocket buildings most often use surfaces with the best properties of
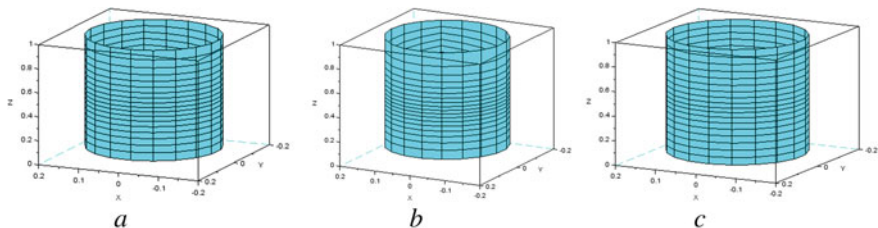
Fig. 3 Creation of the cylinder surface grid: **a** thickening condition $a_n = 3$, $c_n = 1$; **b** thickening condition $a_n = 5$, $c_n = 1$; **c** thickening condition $a_n = 5$, $c_n = 5$
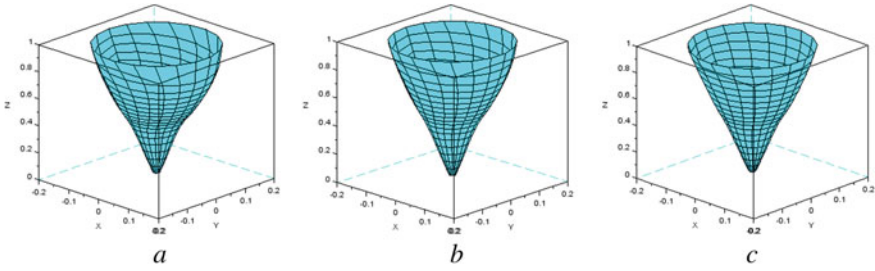
Fig. 4 Thickening of the grid of the cone surface: **a** thickening condition $a_n = 8, c_n = 1$; **b** thickening condition $a_n = 6, c_n = 1$; **c** thickening condition $a_n = 6, c_n = 3$
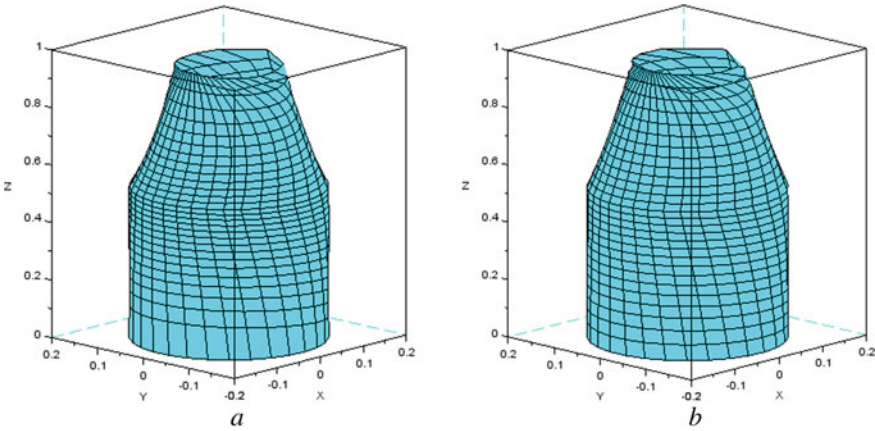
Fig. 5 Thickening of the grid during a combination of a cylinder surface and a cone surface: a) $a_n = 6, c_n = 1$; b) $a_n = 3, c_n = 1$

wrap, which do not have clearly pronounced angles, for example, to reduce air resistance. An example of such domains is the cylindrical or conic shells. The study was considered a combination of surfaces of the following bodies of rotation: cylinder and cone (Fig. 5), cylinder and spheres (Fig. 6) and two conical surfaces (Fig. 7). To construct the specified surfaces, the parameters of control functions were determined, which gave the best option of combining nodes of the grid: to combine the cylinder with a cone and the sphere and to combine the two cones.

3.3 Research Results

The grid generation shells and combinations of shells could be employed in different engineering fields, for example, in the air and rocket industry. The thickening of the grid in different places of a homogeneous surface and the connection of different
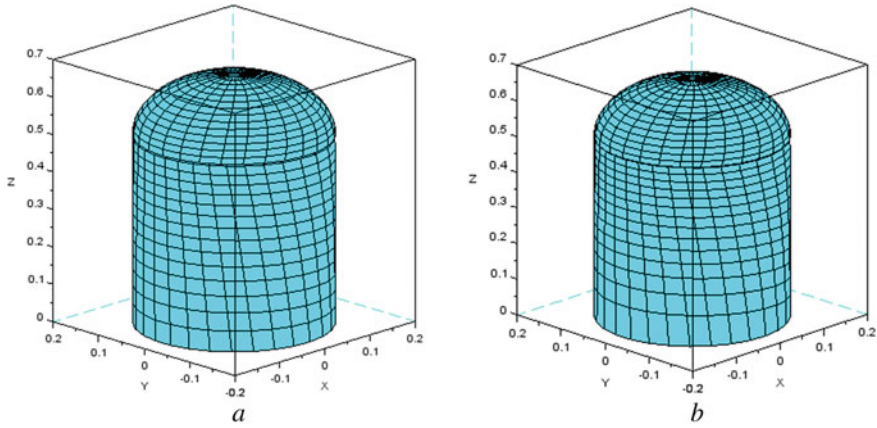
Fig. 6 Thickening of the grid during a combination of a cylinder surface and a sphere surface: **a** $a_n = 3, c_n = 1$; **b** $a_n = 5, c_n = 1$
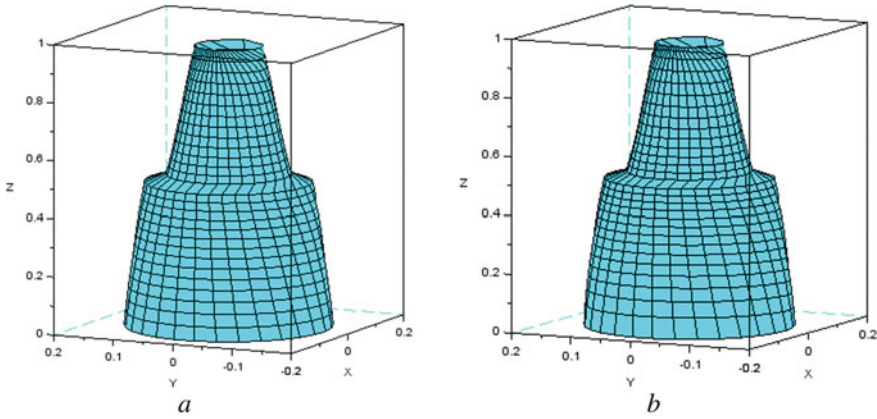
Fig. 7 Thickening of the grid during a combination of two cone surfaces: **a** $a_n = 3, c_n = 1$; **b** $a_n = 5, c_n = 1$

surfaces is investigated. The thickening of the grid was performed using the control functions of the Poisson equation. The influence of parameters of control functions to the values of each mesh cell is investigated. It has been found that the quality of the grid (its orthogonality) improves with an increase in the value of the second parameter c_n of the formulas of control functions (5), (6).

The observed patterns of influence parameters of control functions on the orthogonality of the structured grid are confirmed by similar results spent on flat regions [26, 27].

4 Conclusions

Elliptical methods allow to use partial differential equations to manage the density of elements in the structured grid. Such methods require to define controlling functions and solution of an elliptical equation. Particular attention was paid to the investigation of the influence of control functions on the management of the grid's density. For the elliptical method, control functions are considered by which can be condensed to coordinate lines. The influence of parameters of control functions of the elliptical method on the mesh quality was investigated, namely its orthogonality.

References

1. Xia K, Zhan M, Wan D, Wei GW (2012) Adaptively deformed mesh based interface method for elliptic equations with discontinuous coefficients. J Comput Phys 231(4):1440–1461
2. Huang W, Ma JT, Russell RD (2008) A study of moving mesh PDE methods for numerical simulation of blowup in reaction diffusion equations. J Comput Phys 227:6532–6552
3. Wan DC, Turek S (2007) An efficient multigrid-FEM method for the simulation of solid-liquid two phase flows. J Comput Appl Math 203:561–580
4. Wan DC, Turek S (2007) Fictitious boundary and moving mesh methods for the numerical simulation of rigid particulate flows. J Comput Phys 222:28–56
5. Tang T (2005) Moving mesh methods for computational fluid dynamics. Contemp Math 383
6. Oevermann M, Klein R (2006) A cartesian grid finite volume method for elliptic equations with variable coefficients and embedded interfaces. J Comput Phys 219:749–769
7. Chen T, Strain J (2008) Piecewise-polynomial discretization and Krylov-accelerated multigrid for elliptic interface problems. J Comput Phys 227:7503–7542
8. Chern I, Shu YC (2007) A coupling interface method for elliptic interface problems. J Comput Phys 225:2138–2174
9. Zhao S, Wei GW, Xiang Y (2005) Dsc analysis of free-edged beams by an iteratively matched boundary method. J Sound Vib 284(1–2):487–493
10. Yu SN, Wei GW (2007) Three-dimensional matched interface and boundary (MIB) method for treating geometric singularities. J Comput Phys 227:602–632
11. Yu SN, Zhou YC, Wei GW (2007) Matched interface and boundary (MIB) method for elliptic problems with sharp-edged interfaces. J Comput Phys 224(2):729–756
12. Zhou YC, Wei GW (2006) On the fictitious-domain and interpolation formulations of the matched interface and boundary (MIB) method. J Comput Phys 219(1):228–246
13. Zhou YC, Zhao S, Feig M, Wei GW (2006) High order matched interface and boundary method for elliptic equations with discontinuous coefficients and singular sources. J Comput Phys 213(1):1–30
14. Zhao S (2010) High order matched interface and boundary methods for the helmholtz equation in media with arbitrarily curved interfaces. J Comput Phys 229:3155–3170
15. Wu CL, Li ZL, Lai MC (2011) Adaptive mesh refinement for elliptic interface problems using the nonconforming immerse finite element method. Int J Numer Anal Model 8:466–483
16. Yu SN., Geng WH., Wei GW (2007) Treatment of geometric singularities in implicit solvent models. J Chem Phys 126:244108
17. Selim MM, Koomullil RP (2016) Mesh deformation approaches—a survey. J Phys Math 7:181
18. Dwight RP (2009) Robust mesh deformation using the linear elasticity equations. J Comput Fluid Dynam 12:401–406
19. Luke E, Collins E, Blades E (2012) A fast mesh deformation method using explicit interpolation. J Comput Phys 37:586–601

20. Maruyama D, Bailly D, Carrier G (2012) High-quality mesh deformation using quaternions for orthogonality preservation. In: 50th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition, USA (2012)
21. Zhou X, Li S (2015) A novel three-dimensional mesh deformation method based on sphere relaxation. J Comput Phys 298:320–336
22. Sun S, Lv S, Yuan Y, Yuan M (2016) Mesh Deformation Method Based on Mean Value Coordinates Interpolation. Acta Mech Solida Sin 29:1–12
23. Witteveen JAS (2010) Explicit and robust inverse distance weighting mesh deformation for CFD. In: 48th AIAA Aerospace sciences meeting including the new horizons forum and aerospace exposition, USA
24. Boer AD, van der Schoot MS, Bijl H (2007) Mesh deformation based on radial basis function interpolation. J Computers Struct 45:784–795
25. Michler AK (2011) Aircraft control surface deflection using RBF-based mesh deformation. Int J Numer Meth Eng 88:986–10079
26. Khalanchuk LV, Choporov SV (2020) Research of non-uniform structured discrete models generation for two-dimensional domains. Visnyk of Zaporizhzhya National University. Physical and Mathematical Sciences 1:106–112
27. Khalanchuk LV, Choporov SV (2020) Development of a method for constructing irregular meshes based on the differential poisson equation. Appl Quest Math Model Kherson 3(№ 2.2):274–282
28. Valger SA, Fedorova NN (2012) Primenenie algoritma k adaptatsii raschetnoi setki k resheniu Eilera. Vychislitelnye tehnologii 17(№ 3):24–33
29. Sosnytska N, Morozov M, Khalanchuk (2020) Modeling of electron state in quantum dot structures. In: 2020 IEEE problems of automated electrodrive. Theory and Practice (PAEP), Kremenchuk, Ukraine, pp 1–5 (2020)

Research and Mathematical Modeling of Filter Elements Production Processes

Svitlana Popereshnyak and **Anastasiya Vecherkovskaya**

Abstract The article reviews the industrial production of polypropylene filter elements. One of the main problems of polypropylene fibrous filter elements production management is technological process control. The polypropylene fibrous filter elements technological production process control system modeling is reviewed. Considering the large number of factors and parameters influencing the production process, 4 main ones were identified. And the filter element fiber thickness mathematical models that depend on the technological process parameters were created. The research of the created models was conducted. This made it possible to construct an algorithm for designing the polypropylene filtering elements fiber optimum thickness.

1 Introduction

Increasing the efficiency of a modern manufacturing enterprise is possible through the development and implementation of complex information systems and complexes, which will include technological control objects and information systems for managing technological processes.

Despite significant achievements in the development of computer technology and information technology, the creation of the information technology at chemical plants associated with the production of polypropylene fiber filter elements (PFFE) does not keep up with the growth of labor productivity in the industry. The variety of machine types and the corresponding production methods and procedures did not

S. Popereshnyak (✉) · A. Vecherkovskaya
Taras Shevchenko National University of Kyiv, 24, Bohdana Havrylyshyna Str., Kyiv 04116, Ukraine
e-mail: spopereshnyak@gmail.com

A. Vecherkovskaya
e-mail: vecherkovskaia90@gmail.com

allow to develop and improve the universal information management system for such an enterprise. In view of this, the question of finding and developing technologies on the basis of which systems invariant to production problems can be developed is important and relevant.

PFFE production is a complex production process, and is performed with the use of automation and information technologies, as a reflection of the general trend of transition from manual production operations to machine ones. If the increase in labor productivity is ensured by the use of more advanced tools, then maintaining it at a necessary high level is ensured by appropriate forms of organization of the production technological process and the use of purposeful and adaptive management of subjects that are involved in the production process.

2   Analysis of the Structure of PFFE, Depending on the Field of Application

Let us consider fibrous-porous filtering elements based on fiber-forming polypropylene for filters for purifying natural gas, air, water and aqueous solutions, hydraulic and turbine oils, petroleum products and other organic liquids. The basis for the production of PFFE is a bulky filtering nonwoven material obtained by aerodynamic spraying of a polymer melt. This polymer has high chemical resistance, low tendency to swell in bases, acids and oil products. This property allows this material to be used for manufacturing filter elements for a wide range of applications. Namely, in various filters and separator filters designed to clean natural and associated gas from solid particles, droplet moisture, oil aerosol and gas condensate, as well as in filters used to filter various liquids. And such a feature of the PFFE structure, as the autohesion of fibers at their intersection points, allows them to be effectively used as coalescers [1].

Let us consider the process of manufacturing elements from porous polypropylene by pneumoextrusion. The method consists in the formation of molten polypropylene through forming heads, followed by stretching the uncured extrudate with a stream of hot air and applying it to a rotating helical cylindrical rod. The design of the receiving device allows for a continuous process of forming frameless elements (see Fig. 1).

The resulting structure of the cartridges is rigid and stable, since it is fixed by thermal stitching of microfibers and slightly pressed with a special roller. With the seeming simplicity of the technology, the operator must control a sufficiently large number of process parameters, taking into account the incoming quality control of the feedstock. By changing the key parameters, you can get cartridges of different micron ratings from 0.3 to 100 microns, different lengths and diameters [2].

The area of application of elements made of porous polypropylene is quite diverse [3, 4] (Table 1).
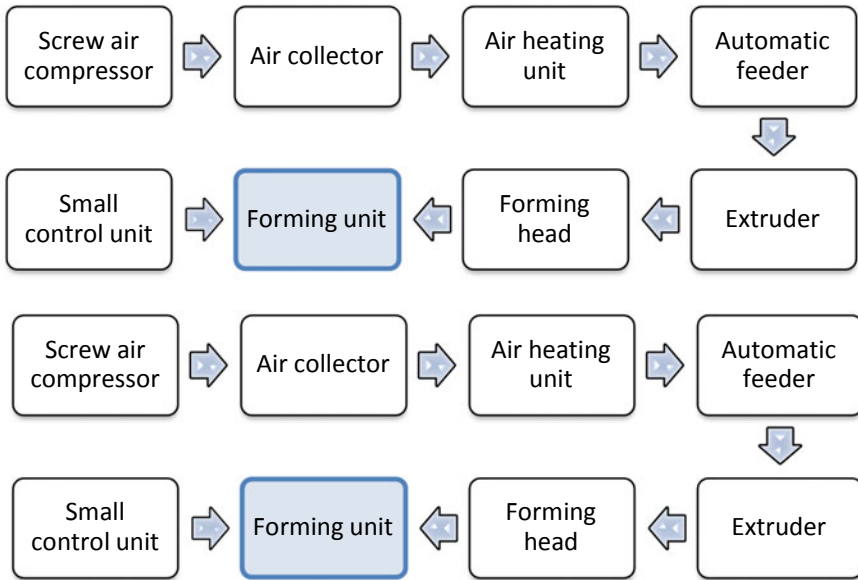
Fig. 1 Diagram of the production of elements from porous polypropylene by pneumoextrusion

Table 1 Application examples of filter elements in various industries

Branch	Application
Water preparation	Industrial filters Household filters Industrial wastewater treatment
Food industry	Purification of drinking water Filtration of wine, beer, juices Purification of vegetable oil
Chemical industry	Purification of water and chemical solutions Purification of galvanic solutions Cleaning of varnishes, emulsions, paints
Biotechnology	Purification of culture fluids Micro filtration Filtration as sterilization
Mechanical engineering industry	Cleaning of transformer and turbine oils Cleaning of lubricating and cooling liquids Fuel cleaning
Medical industry	Purification of aqueous solutions Purification of non-aqueous solutions

This includes the arrangement of water wells, filtration of fuel and fuels and lubricants, the use of filters in industry, filtration of gases and air, the use of drainage systems in agriculture and construction, etc. [5, 6].

The structure of filter elements is formed on the basis of calculations made on the basis of technical requirements for a specific filter-separator and liquid composition. The task of the calculations is to determine the design parameters of the PFFE (the design of filter elements and their number in a particular filter), which, in turn, must provide the main operational parameters: density, filtration efficiency, permissible initial pressure drop [7].

The feasibility of using a stochastic form of Pennes bioheat model within a 3-D finite element based Kalman filter (KF) algorithm is critically evaluated for the ability to provide temperature field estimates in the event of magnetic resonance temperature imaging (MRTI) data loss during laser induced thermal therapy [8].

3 The Set of Parameters of the Technological Process of Manufacturing Elements of Porous Polypropylene by Pneumoextrusion

During the production of filter elements and drainage elements by pneumo-extrusion method, a significant number of parameters of the technological process can be changed to ensure a certain structure of each layer [9, 10].

These are parameters such as:

- Melting temperature in the heating zones of the extruder and the forming head;
- The speed of the dosing pump and, accordingly, the mass of the polymer melt;
- Pressure and temperature of the air supplied to the forming head;
- The rotational speed of the receiving mechanism and its distance from the forming head;
- Parameters of the calendering device (if used).

It should also be taken into account that all of these options will change dramatically when used in the manufacture of raw materials of various grades.

If we analyze all the input parameters, then the main task of automating the production of filter elements from porous polypropylene by extrusion becomes to reduce the number of control parameters of said process.

The number of layers in this case can be increased, starting from the filtering layer to the outer, accumulative layer, while the density of the material will decrease in the direction of the outer layer. This process in production can be either discrete or continuous.

As indicated above, we note that we will consider the production of elements from porous polypropylene by extrusion from one grade of polymer.

It has been experimentally substantiated that the most influential are two main input parameters that affect the diameter of the fiber, which is formed from molten
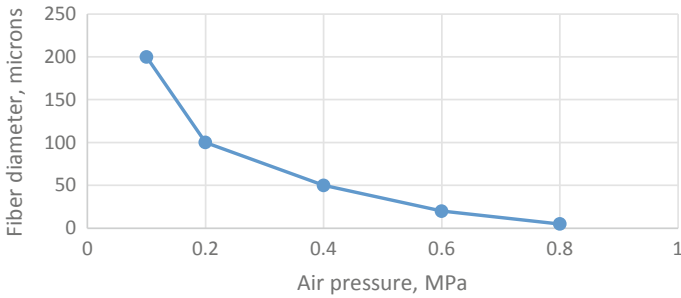
Fig. 2 Diagram of the dependence of the fiber diameter on the air pressure supplied to the forming head

polypropylene, and, accordingly, the density of the layer on the product is the air pressure and the material feed rate (molten polypropylene) [11].

The fiber diameter for the production of the layers of the filter element, which interchange the fine-fiber and thick-fiber layers, depends on the air pressure when the fibers are applied to the product (see Fig. 2).

Large surfaces such as rods with many holes give low resistance and the melt pressure in clean nets can be as low as 6–10 MPa at the worm tip.

It has been experimentally established that with an increase in the polypropylene feed rate to the mold from 1000 to 5000 m/min, a significant change in the mechanical properties of the resulting fibers is observed: a decrease in tensile elongation, an increase in the modulus and fiber strength (see Fig. 3).

For the production of a thinner layer, where a fiber thickness of about 20 microns, usually 10–25 MPa is supplied, and a small wire insulation can give up to 50 MPa or more.

A lower pressure forming head may allow faster production, but if the thickness is not uniform, there will be excess material consumption or the risk of scrap.
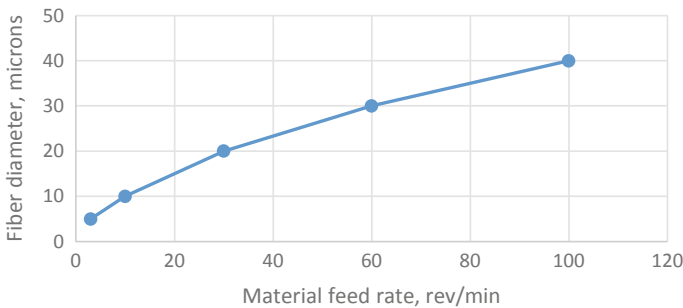
Fig. 3 Graph of the dependence of the fiber diameter on the feed rate of the material

3.1 The Method of Selecting System Parameters Depending on the Industry

To select the required parameters for the selected application, follow these steps:

1. Determine the technical conditions (characteristics) of the industry for which the filter element will be produced.
2. If a drainage element is made, which consists of a single layer of fibers, then it is necessary to set control parameters for it, the main of which are air pressure and material supply rate.
3. If the production of a filter element is going to take place, then for it it is first necessary to determine how many layers and in what order will be applied. Then, in the same way as for the drainage element, control parameters are set before starting each layer.
4. Monitor parameters at the beginning and during production, avoid equipment malfunctions.

3.2 Research and Mathematical Modeling of PFFE Fiber Thickness

The melting temperature in the last heating zone of the extruder, the revolutions of the dosing pump and, accordingly, the mass of the polymer melt, the air pressure supplied to the forming head and the air temperature supplied to the forming head are the main characteristics for the process of polypropylene fiber filter elements (PFFE) forming. To obtain a thin fiber with a diameter of up to 150 microns. Technological characteristics must have the following ranges [3]:

- Melting temperature in the last heating zone of the extruder 290–300 °C
- The dosing pump revolutions and, accordingly, of the polymer melt mass 25–33 rev/min
- The air pressure supplied to the forming head 0.6–1 MPa
- The air temperature supplied to the forming head 15–150 °C.

Modeling and processing of experimental data were performed using the mathematical package MATLAB.

We introduce the following notation:

y—fiber thickness;
x_1—melting temperature in the extruder last heating zone °C;
x_2—dosing pump revolutions and, accordingly, the polymer melting mass revolutions rev/min;
x_3—the air pressure supplied to the forming head MPa;
x_4—the air temperature supplied to the forming head °C.

We receive a mathematical model in the form of a multifactor regression equation:

$$y = 2643.28x_1 - 6.12x_2 + 7.7x_3 - 186.56x_4 - 7.32$$
$$- 0.008x_1x_2 + 0.199x_1x_3 + 0.008x_1x_4 - 0.029x_2x_3$$
$$- 0.001x_2x_4 + 0.022x_3x_4$$

The point estimates coefficients calculation of the regression equation was performed using the least squares method, which combined statistical and mathematical calculations. In General, the equation of fiber thickness is described by the following polynomial:

$$y = \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} x_i x_j + b_0$$

where
$\{x_i, x_j\} \in R^n$—production parameters vector;
$\{b_i, b_{ij}, b_0\}$—weight coefficients of the regression equation.
The smallest and largest values of separate technological parameters are presented to define their degree of influence on the basic physical characteristics of PFFE fabric (Table 2).

Using the mathematical package MATLAB, graphic 3D models were built that describe the dynamics of fiber thickness with pairwise changes of two independent factors (melting temperature in the last heating zone/speed of the metering pump, air pressure/air temperature supplied to the spray head) and are presented in Figs. 4, 5, 6, 7, 8 and 9.

In the selected range of technological parameters changes, fiber with a thickness of 15–150 microns was produced, which significantly affects the filtering properties of the final product.

Table 2 Limit values of separate technological parameters

Fiber diameter μm	Melting temperature in the last heating zone of the extruder t_1, °C	The dosing pump revolutions and, accordingly, of the polymer melt mass N, rev/min	The air pressure supplied to the forming head P, MPa	The air temperature supplied to the forming head t_2, °C
	Max	Min	Max	Max
150	300	33	0.6	150
110	300	25	0.6	150
60	300	25	1	150
15	300	25	1	160

$$y = -583{,}26x_1^2 + 6{,}16x_2^2 + 2{,}63x_1 - 0{,}01x_2 - 0{,}013$$
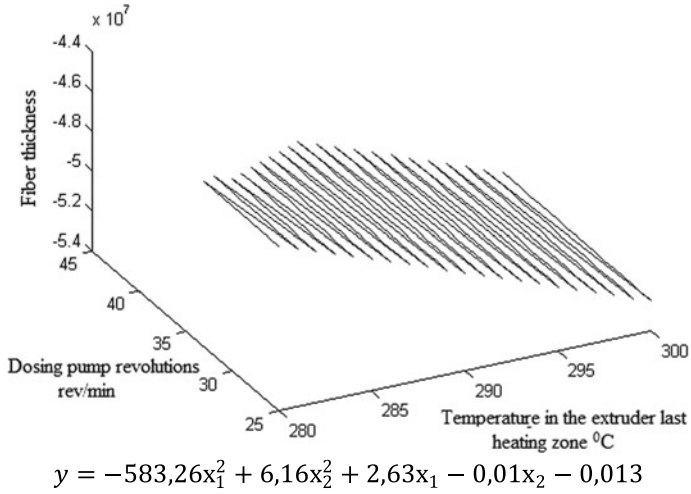
Fig. 4 Graphic 3D model of fiber thickness (x_1—melting temperature, x_2—dosing pump speed)

$$y = -1270{,}42x_1^2 + 11{,}74x_3^2 + -37{,}71x_1 - 0{,}02x_3 - 7{,}43$$
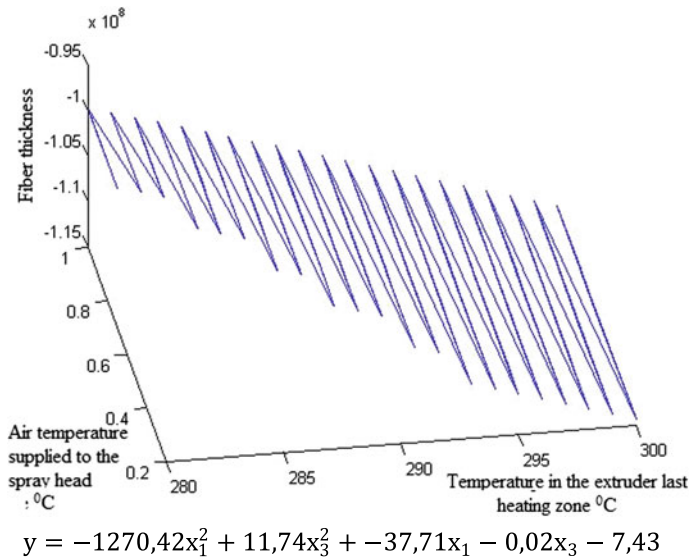
Fig. 5 Graphic 3D model of fiber thickness (x_1—melting temperature, x_3—air temperature supplied to the spray head)

$$y = -1090{,}76x_1^2 + 9{,}74x_4^2 + 2{,}96x_1 - 0{,}02x_4 - 0{,}02$$
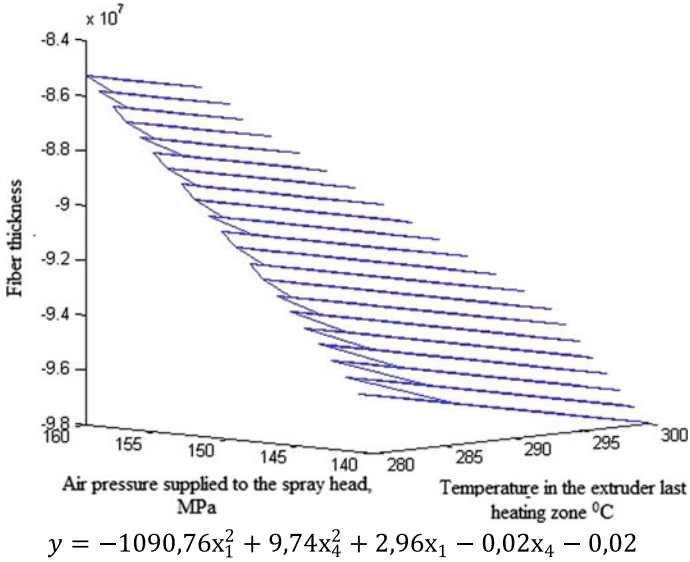
Fig. 6 Graphic 3D model of fiber thickness (x_1—melting temperature, x_4—air pressure supplied to the spray head)

$$y = 82{,}43x_2^2 + 2{,}94x_3^2 - 39{,}71x_2 - 0{,}018x_3 - 3{,}19$$
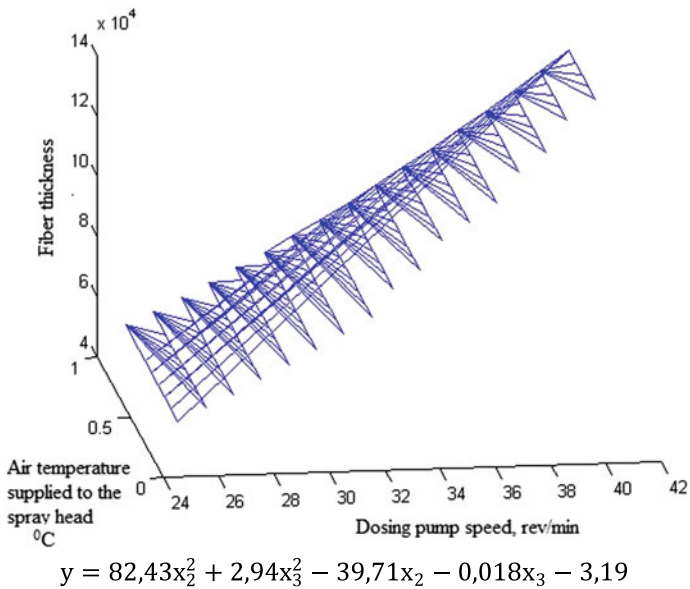
Fig. 7 Graphic 3D model of fiber thickness (x_2—dosing pump speed, x_3—air temperature supplied to the spray head)
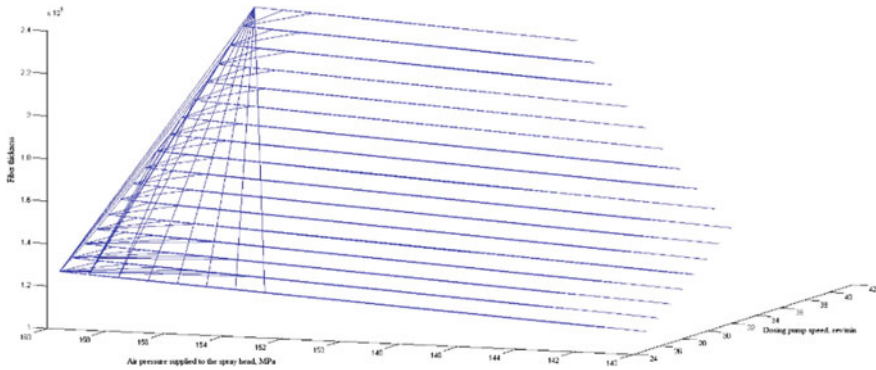
$$y = 90{,}37x_2^2 + 2{,}67x_4^2 + 1{,}34x_2 - 0{,}014x_4 - 0{,}01$$

Fig. 8 Graphic 3D model of fiber thickness (x_2—dosing pump speed, x_4—air pressure supplied to the spray head)
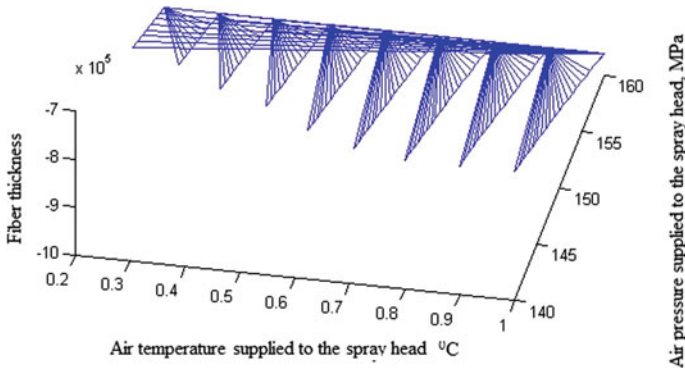
$$y = -6{,}32x_3^2 - 37{,}26x_4^2 + 4{,}14x_3 - 7{,}57x_4 - 0{,}02$$

Fig. 9 Graphic 3D model of fiber thickness (x_3—air pressure supplied to the spray head, x_4—air temperature supplied to the spray head)

3.3 Algorithm for Developing the Allowable Thickness of PFFE Fiber

The carried out experimental and theoretical studies made it possible to form a universal algorithm for obtaining the permissible fiber thickness, depending on the physical characteristics of the technological process.

The algorithm of the optimum polypropylene fibrous filter elements fibers thickness as the main task requiring automation is presented as a block diagram in Fig. 3.

The presented algorithm allows solving the following problems:

- supports the creation of polypropylene fibrous filtering elements formulations with specified consumer characteristics, depending on the application field and optimal cost;
- forms recommendations on the filtering characteristics of the received polypropylene fibrous filter elements according to current standards (DSTU, TU, etc.);
- provides search of optimum physical characteristics at all stages of technological process;
- analyzes the final product quality, identifies technological problems and suggests solutions.

The presented algorithm can be used to select the fiber structure of a polypropylene filter element depending on the application field (Fig. 10).

4 Conducting Experimental Studies and Modeling of Fiber Thickness

Experimental studies were carried out on the production of filtering polypropylene fiber elements. TPP D30S grade of polypropylene was used.

As a result of the analysis of the influence of the factors of the technological process for the formation of the PFFE fiber, the following were chosen: the melting temperature in the last heating zone of the extruder, the speed of the metering pump, the air pressure supplied to the forming head, the temperature of the air supplied to the forming head.

To obtain fiber with a thickness of up to 25–50 microns technological characteristics should have the following ranges: melting temperature in the last heating zone of the extruder 290–300°C, speed of the metering pump and, accordingly, the mass of polymer melt 25–32 rev/min, air pressure supplied to the forming head 0.8–1 MPa, temperature of the air supplied to the forming head 153–160 °C.

To control and obtain temperature values in the last heating zone of the extruder, a thermocouple ТХК 2488 (1A … 8A) with an operating temperature range of −40 … 400 °C was used, and two-channel POD temperature regulators "MICRA-604", which provide control and indication of the temperature of the melting mass in temperature zones 1–4 located in the region of the extruder barrel in which the screw rotates, in the region of the "knee" and in the region of the beam on which the metering pumps are located, respectively; control the power of thermocouples.

To determine the air pressure supplied to the forming head, a proportional pressure regulator "VEAA-L-D2-Q4-V1-1-R1" (10A) was used.

The control and indication of the consumption of the polymer melt mass is carried out using the microprocessor controller "MIK-50", which sends commands to the frequency converter "GD100 2r2g-4" (11A), which in turn controls the AIR 90L4Y2 (M2) engine, the engine sets in motion mechanical pumps mounted on the extruder frame.
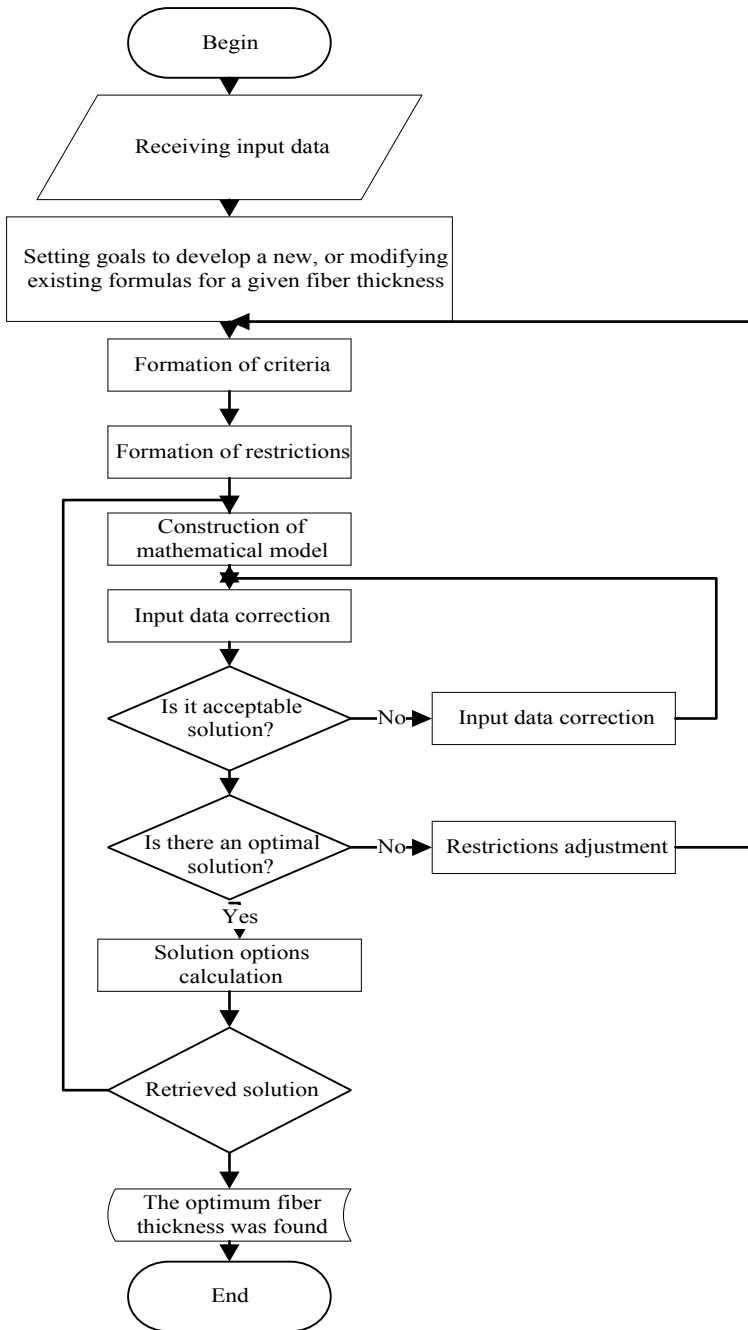
Fig. 10 The polypropylene fibrous filter elements optimum thickness modeling algorithm

When considering possible variants of the experimental design, a full factorial experiment (FFE) was chosen.

The investigated parameter is the FE fiber thickness.

The investigated factors, the nature of their change and coding schemes are shown in Table 3.

To avoid repetition of experiments, we will use the principle presented in Table 4.

The factor designated "x_0"—the so-called dummy variable—is introduced to calculate the free term of the desired model—the coefficient "b_0". After drawing up the table, we check the correctness of its filling. The symmetry condition of the replica requires that each column contains an equal number of minuses and pluses. The condition is met.

Table 3 Investigated factors

Characteristic	Factors			
	t_1, °C	N, rev/min	P, MPa	t_2, °C
Code	x_1	x_2	x_3	x_4
Base level	295	30	0.9	157
Interval of variation	1	1	0.1	1
Lower level (coded value −1, −)	293	25	0.8	153
Top level (coded value + 1, +)	300	32	1	160

Table 4 Drawing up an experimental plan

Experiment number	x_1	x_2	x_3	x_4	x_0
1	+	+	+	+	+
2	−	+	+	+	+
3	+	−	+	+	+
4	−	−	+	+	+
5	+	+	−	+	+
6	−	+	−	+	+
7	+	−	−	+	+
8	−	−	−	+	+
9	+	+	+	−	+
10	−	+	+	−	+
11	+	−	+	−	+
12	−	−	+	−	+
13	+	+	−	−	+
14	−	+	−	−	+
15	+	−	−	−	+
16	−	−	−	−	+

Experiment plan implementation

To determine the error of the experiment, the experiments will be duplicated (at the main level). To determine the variance of the experience, additional experiments 17–20 were implemented at the main level.

After the implementation of all the experiments of the planning matrix, based on their results, we build a mathematical model of the process. To do this, when using FFE, we calculate the regression coefficients of the equation using the formula:

$$b_j = \sum_{j=1}^{n} \frac{x_{jn} \cdot y_n}{n},$$

b_j—value of the jth regression coefficient ($j = 0, 1, 2, ..., n$);
x_{jn}—the value of the jth factor in the nth experiment in coded form;
y_n—the value of the optimization parameter in the nth experiment;
n—the number of experiments in the planning matrix.

As a result, we obtain a model that looks like this:

$$y = 2315 - 5x_1 + 5x_2 - 125x_3 - 5x_4,$$

where

y—fiber thickness;
x_1—melting temperature in the extruder last heating zone °C;
x_2—dosing pump revolutions and, accordingly, the polymer melting mass revolutions rev/min;
x_3—the air pressure supplied to the forming head MPa;
x_4—the air temperature supplied to the forming head °C.

For it we find paired correlation coefficients R in Table 5.

In our case, all pairwise correlation coefficients $|r| < 0.7$, which indicates the absence of multicollinearity of factors.

For our model, we can determine the direct influence of the factor signal x_1 on the result of Y in the regression equation, β_j is measured and is -1.135; indirect (non direct) influence of this factor signal on the result is defined as:

Table 5 Paired correlation coefficients

–	y	x_1	x_2	x_3	x_4
y	1	−0.2468	0.2468	−0.1518	−0.2468
x_1	−0.2468	1	0.2697	−0.2633	−0.2697
x_2	0.2468	0.2697	1	0.2633	0.2697
x_3	−0.1518	−0.2633	0.2633	1	−0.2633
x_4	−0.2468	−0.2697	0.2697	-0.2633	1

$$r_{x_1 x_2} \beta_2 = 0.3063$$

Testing hypotheses regarding the coefficients of the regression equation (checking the multiple regression equation parameters significance).

Number $v = n - m - 1$ is called the number of degrees of freedom. It is believed that when estimating multiple linear regression, to ensure statistical reliability, it is required that the number of observations is at least 3 times higher than the number of estimated parameters.

t-statistics

$$T_{tabl}(n - m - 1; \alpha/2) = 2.263$$

$$t_i = \frac{b_i}{S_{b_i}}$$

Let's calculate t_i, for our model. We receive: $t_i > 2.263$. Which suggests that the statistical significance for all regression coefficients $b_0 \ldots b_4$ confirmed.

Checking the overall quality of the multiple regression equation.

Because the actual value $F > F_{kp}$, then the coefficient of determination is statistically significant and the regression equation is statistically reliable (ie, the coefficients b_i are jointly significant).

5   Conclusions

Modeling is one of the main research methods in all fields of knowledge and a scientifically justified method for evaluating the complex systems characteristics used in various fields of engineering decision making. We have reviewed the polypropylene fibrous filter elements technological production process control system modeling.

With the help of mathematical modeling, it was determined that the degree of influence on the fiber thickness and filtering properties of technological parameters, namely, the melting temperature in the last heating zone of the extruder and the temperature of the air supplied to the forming head is extremely high, therefore the PFFE manufacturing process must be controlled.

Based on the results of experimental statistical modeling, it can be concluded that the filtering properties of the finished PFFE directly depend on its manufacturing technology and technological indicators of this process, as well as the grade of polypropylene (which we cannot control) from which the product is made.

This made it possible to construct an algorithm for designing the polypropylene filtering elements fiber optimum thickness.

References

1. Myk YA, Horev VN (2018) Simulation of the hybrid filtering material. Polzunovskyy vestnyk 1:135–142
2. Sambaer W, Zatloukala M, Kimmerb D (2011) 3D modeling of filtration process via polyurethane nanofiber based nonwoven filters prepared by electrospinning process. Chem Eng Sci 66(4):613–623
3. Martínez GS, Sierla SA, Karhela TA, Lappalainen J, Vyatkin V (2018) Automatic generation of a high-fidelity dynamic thermal-hydraulic process simulation model from a 3D plant model. IEEE Access 6:45217–45232
4. Huang Y, Peng S, Yao W (2010) A study on modeling and analysis of production order and process deadlock free control for an automated production system using Petri net technique. In: Proceedings of SICE annual conference 2010, pp 2236–2241
5. Majdana R, Tkaca Z, Abrahama R, Kollarovab K, Vitazeka I, Halenar M (2017) Filtration systems design for universal oils in agricultural tractors. Tribol Ind 39(4):547–558
6. Babanin O, Butskiy O, Kovalenko O, Maksimov M (2018) Application of synthetic filters from polypropylene in diesel locomotive oil systems to improve the efficiency of cleaning engine oil. Int J Eng Technol 7(4.3):162–166
7. Yakovlev VV, Kalugin YI, Kalugin AY (2019) Mathematical modelling of enhanced gas production. J Petrol Explor Prod Technol 9:561–571. https://doi.org/10.1007/s13202-018-0495-7
8. Fuentes D, Yung J, Hazle JD, Weinberg JS, Stafford RJ (2012) Kalman filtered MR temperature imaging for laser induced thermal therapies. IEEE Trans Med Imaging 31(4):984–994. https://doi.org/10.1109/TMI.2011.2181185
9. Ittiphalin M, Chearanai T (2019) The application of mathematical model for production planning in a polystyrene factory. In: 2019 IEEE 6th international conference on industrial engineering and applications (ICIEA), 2019, pp 169–172
10. Vecherkovskaya A, Popereshnyak S (2017) Comparative analysis of mathematical models forming filter elements. In: XIII international conference on perspective technologies and methods in MEMS design (MEMSTECH) proceedings, Polyna, 20–23 April 2017, Lviv, p 113
11. Rigatos G, Siano P, Wira P, Busawon K, Jovanovic IM (2018) Nonlinear H-infinity control for optimizing cement production. In: 2018 UKACC 12th international conference on control (CONTROL) 2018, pp 248–253

Modeling of Technological Machines Spatial Rigidity by Subsystems

Volodymyr Chupryna [ORCID]

Abstract The article proposes to develop a tensor approach for obtaining a spatial rigidity of an elastic system of a technological machine based on spatial rigidity of its basic subsystems (nodes). In the technological machine there are at least two main nodes that hold the cutting tool and the preparation, and corresponding to two ellipsoids—an ellipsoid rigidity of the tool and an ellipsoid rigidity of the preparation. The main task is to develop a method and algorithm for constructing a space of spatial rigidity in a working zone of a machine for further analysis of the accuracy of processing. The direct connection of spatial rigidity of the technological machine with the precision of processing is proved. Mathematical modeling of surfaces of spatial rigidity of the machine elastic system is carried out. The method of finding rigidity (pliability) of the technological machine is determined in an arbitrary direction of three-dimensional space. On the basis of a constructed 3D surface, you can find the values of the machine deformation in the cutting zone from the active power load, in particular, the tightening of the tool from the part that directly determines the processing accuracy—the size and shape of the part. This allows you to predict the achievable accuracy of machines at the design stage.

Keywords Technological machine · Spatial rigidity · Subsystem · Node · Rigidity ellipsoid · Pliability ellipsoid · Tensor model · Precision processing

1 Introduction

Intensification of mechanical processing, in particular cutting, in modern technological processes requires the high rigidity of technological equipment and machines.

The rigidity of machines manifests itself in the form of an uneven rigidity of the machine in the working space, which affects the accuracy of processing and the emergence of deviations from the ideal geometric shape of the treated parts.

V. Chupryna (✉)

State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Striletska str. 1, Chernihiv 14033, Ukraine

e-mail: voldchu@ukr.net

As a result, macro- and micro-errors in the form of distortion of the form of parts (ellipsism, ovality, border, tongue, and other) increases in detail, and also deteriorates the roughness of their surfaces. The rigidity of machines is directly related to the precision of processing.

In fundamental works [1–3] theoretically, the relationship between the accuracy of the processing of parts on the machine with the characteristics of the rigidity of the machine and its nodes is installed. The influence of the rigidity of the machine on indicators of its accuracy, in particular the rigidity of individual nodes to the balance of geometric accuracy of the machine.

The monograph [4] shows the role of static and dynamic rigidity of machine tools and their nodes, for the accuracy of processing. The concept of rigidity ellipses (supporter groups, spindle nodes and others) are introduced. It is also determined by the causes of loss of stability, the occurrence of self-oscillations and vibrations when cutting and caused by them by means of processing errors in the form of a violation of the geometry of parts (ellipsism, border, wavy) and deterioration of the roughness of processed surfaces.

In works [5, 6] a tensor approach to analysis of static and dynamic systems of machines by mathematical modeling has been developed.

In the work of the author [7] on the basis of a tensor-geometric idea of the stiffness of the technological machine, the concept of tensor of rigidity of the spindle node is introduced. This tensor is depicted as an ellipsoid and reflects a change in the rigidity of the node not in a separate plane, and in a three-dimensional space (3D-space). Similar tensors can be found for other subsystems and nodes of the machine (support, table).

The article deals with an attempt to develop a tensor approach to obtain a rigidity of the entire technological machine based on its main subsystems. To do this, it is necessary to build a 3D model of spatial stiffness in the working zone of the machine for further analysis of the accuracy of processing.

The purpose of the article is to develop a tensor-geometric model of spatial rigidity of the technological machine, taking into account the rigidity of its main subsystems.

2 Main Part

The processing machine is a prefabricated mechanical design consisting of many individual elements (nodes) and built on an aggregate-modular principle [1, 8, 9]. In this case, we always have two main subsystems that are intended to maintain the end elements—cutting tools and details.

Each of these subsystems has its own spatial rigidity (tensor and ellipsoid of rigidity). This rigidity can be brought to the end element (to the point (zone) of cutting) as an ellipsoid of rigidity. Such subsystems in the technological machine there are at least two of which one is associated with an instrument, and another—with a workpiece. According there are two tensors (ellipsoids) of rigidity—an ellipsoid rigidity of the instrument (ERI) and an ellipsoid rigidity of the part (ERP).

Figure 1 shows examples of industrial robots and machine tools of various technological purposes with ellipsoids of rigidity of the main subsystems.

Each ellipsoid of rigidity corresponds to the tensor of the second order and the corresponding diagonal matrix of rigidity [7, 10]
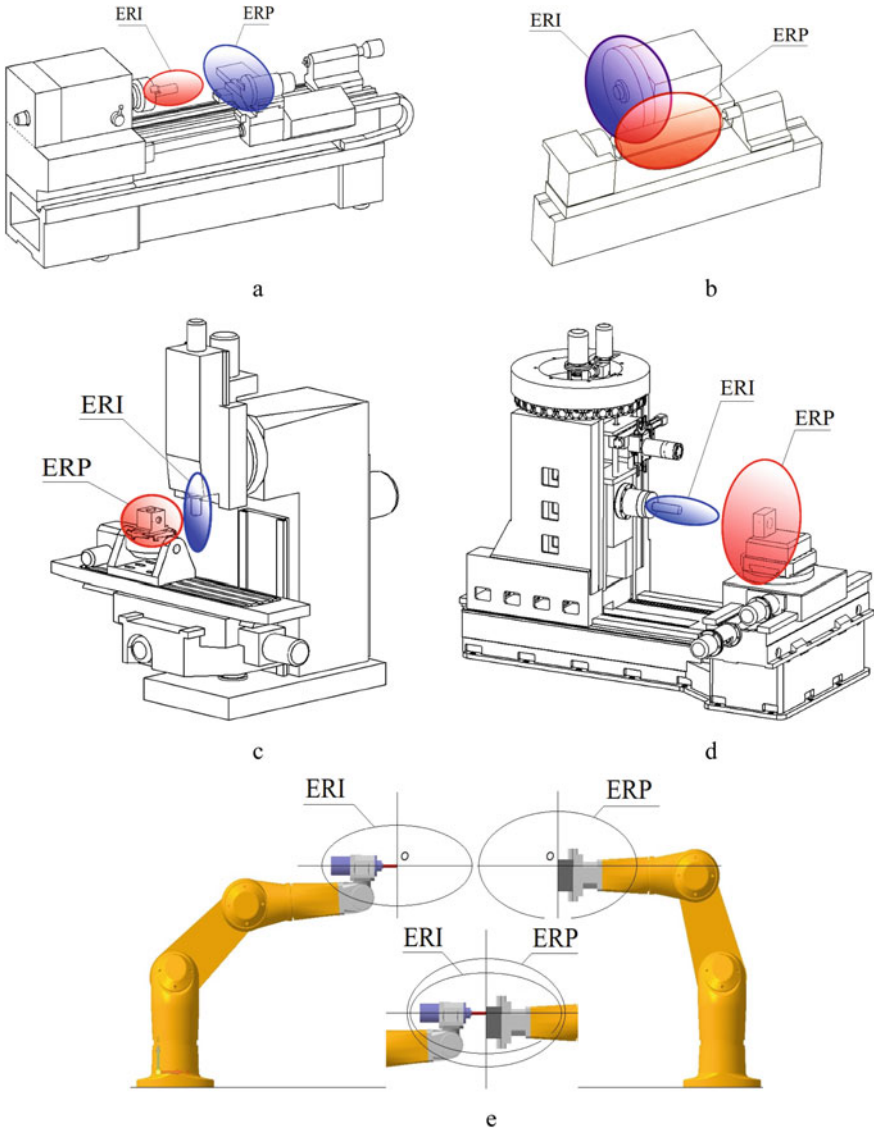
Fig. 1 Ellipsoid of the rigidity of the instrument (ERI) and parts (ERP) for machine tools of various technological groups: **a** turning; **b** round-grinding; **c** milling; **d** multioperative; e industrial robots

$$C = [c_{ij}] = \begin{bmatrix} c_{x0} & 0 & 0 \\ 0 & c_{y0} & 0 \\ 0 & 0 & c_{z0} \end{bmatrix} \tag{1}$$

where c_{x0}, c_{y0}, c_{z0}—the main rigidity of the subsystem of the node.

Or in this form [11]

$$\begin{bmatrix} x & y & z \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{c_{x0}^2} & 0 & 0 \\ 0 & \frac{1}{c_{y0}^2} & 0 \\ 0 & 0 & \frac{1}{c_{z0}^2} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 1. \tag{2}$$

We will show the interaction of ellipsoids of rigidity on an example of a turning machine.

As can be seen from Fig. 2a in the non-working state of the machine tools centers of rigidity ellipsoids are located at the points of bringing rigidity (endpoints) of the tools and parts (n and m).

When cutting (Fig. 2b) points n and m converge (practically coincide) and the load is loaded with a force of cutting P. This force acts both on the tool at the point n and to the detail at the point m (the reaction of force).

Deformation of the elements of the machine under the influence of cutting forces cause the appearance in the system of generalized displacements q, both absolute displacements of points q_n and q_m, and relative $\Delta q = q_m - q_n$. These movements form a deviation from the ideal geometry of the part, that is, the processing errors (preferably in the direction of normal to the processing surface). In order for the operation of the machine, these deviations did not exceed the installed perceptions, it is necessary to provide a sufficient rigidity of the machine.

When solving spatial-deformation tasks, it is more convenient to use the values of pliability than rigidity.
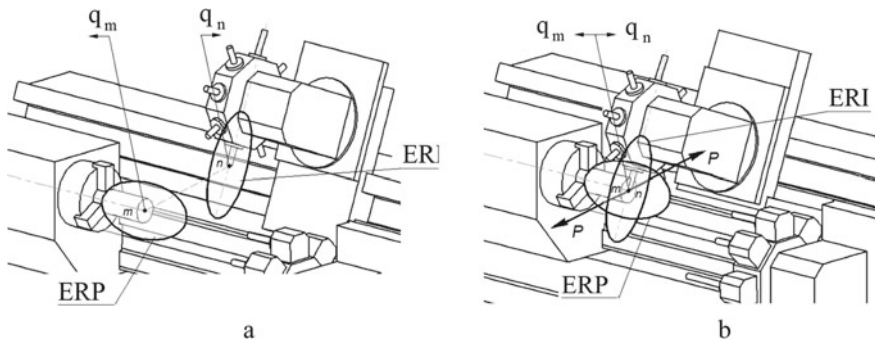
Fig. 2 Constructive schemes of machine tool with rigidity ellipsoids: **a** in a non-working state; **b** in working order (when cutting)
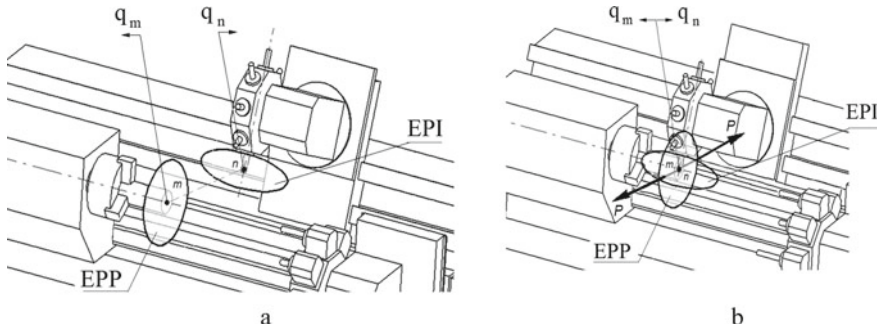
Fig. 3 Constructive schemes of machine tool with pliability ellipsoids: **a** in a non-working state; **b** in working order (when cutting)

For the three main rigidity of the node c_{x0}, c_{y0}, c_{z0} there are three reciprocal values—the main pliability δ_{x0}, δ_{y0}, δ_{z0} and, of course, another tensor—the ellipsoid of pliability.

The ellipsoid of the pliability of the node reflects a change in the magnitude of the system in different directions of the working space from the center of ellipsoid. The magnitude of the pliability is determined by the length of the radius-vector ellipsoid in a given direction.

The ellipsoid of pliability is mathematically described by the formulas similar to the rigidity ellipsoid, only instead of c is used $1/\delta$.

Figure 3 shows the structural schemes of a turning machine with ellipsoids of the support of two main subsystems—tools and details. These ellipsoids are conjugated to the relevant rigidity ellipsoids.

When cutting ellipsoid centers (points n and m) also coincide as shown in Fig. 3b.

In order to find the static pliability of the machine in an arbitrary direction, we will determine the total consumption of the system through the pliability of the components of the tools and parts, as well as describe the surface of spatial pliability.

The force of cutting P and its reaction act on points n and m in opposite directions. Therefore, the relative movement $\Delta q = q_m - q_n$ will be found in the form of an algebraic amount of absolute displacements of points n and m. We find it through the pliability of elastic subsystems at these points

$$\Delta q = P \cdot \delta_m - (-P) \cdot \delta_n = P \cdot (\delta_m + \delta_n). \tag{3}$$

The value Δq is actually pressing the tool from the details in the direction that directly determines the accuracy of processing on the machine.

From the formula (3) it follows that the resulting pliability of the elastic system of the machine in a given direction is equal to the total pliability of the tool and parts in the cutting zone, that is $\delta = \delta_m + \delta_n$.

Thus, the resulting pliability surface of the machine can be constructed using two ellipsoids obtained for points n and m.

In any case, this surface will not be an ellipsoid.

In some partial cases, only if the main axis of the ellipsoid coincides and for both the condition performed, the resulting surface of the pliability of the elastic system can be approximately (with a slight error) to replace the surface of the ellipsoid of pliability, which is described by the equation

$$(\delta_{x0_n} + \delta_{x0_m})^2 \cdot x_0^2 + (\delta_{y0_n} + \delta_{y0_m})^2 \cdot y_0^2 + (\delta_{z0_n} + \delta_{z0_m})^2 \cdot z_0^2 = 1. \qquad (4)$$

In general, the directions of the main axes of drafts are not coincident. In this case, the form of the resulting surface can be obtained by graphic constructs.

The algorithm for finding the resulting surface of spatial support is as follows:

1.   We set the equation of the pliability ellipsoid for the tool subsystem (point n).
2.   Determine the equation of a straight line L that passes through the center of the ellipsoid and specifies the direction of determination of pliability.
3.   Find the coordinates of the point of intersection of direct L with the ellipsoid of pliability (point K).
4.   Determine the magnitude of pliability δ_n.
5.   Repeat p.p. 1–4 for another ellipsoid for the subsystem of the part (point m) and determine the magnitude of pliability δ_m.
6.   We find the total pliability in this direction for the whole system $\delta = \delta_m + \delta_n$.

When finding points on the surfaces of ellipsoids, the direction is given by an infinite straight line L, which is written in the form of a canonical equation [11]

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}. \qquad (5)$$

where l, m, n—guide coefficients;

x_0, y_0, z_0—coordinates of a point that is located on a straight line.

If the point on direct coincides with the onset of the coordinate system, we have

$$\frac{x}{l} = \frac{y}{m} = \frac{z}{n}. \qquad (6)$$

Also, straight line L can be set as a line of intersection of two non-complainary planes in the form of a system of equations

$$\begin{cases} A_1 x + B_1 y + C_1 z + D_1 = 0 \\ A_2 x + B_2 y + C_2 z + D_2 = 0 \end{cases}. \qquad (7)$$

If the planes pass through the beginning of the coordinate system, then $D_1 = D_2 = 0$.

From Eq. (6) you can find all coefficients for the system (7):

$$A_1 = m; \ B_1 = -l; \ C_1 = 0; \ A_2 = n; \ B_2 = 0; \ C_2 = -l.$$

Thus, the coordinates of the point K intersection of the ellipsoid of straight L (see p.3 algorithm) can be found from the linear-quadratic equations of the species

$$\begin{cases} A_1 x + B_1 y + C_1 z = 0, \\ A_2 x + B_2 y + C_2 z = 0, \\ a_{11} x^2 + a_{22} y^2 + a_{33} z^2 + 2a_{12} xy + 2a_{23} yz + 2a_{13} xz + a_{44} = 0 \end{cases} \tag{8}$$

Here, the first two equations are direct, and the third one belongs to the surface of an ellipsoid. The quadratic equations have two decisions, therefore there are two—K and K_1. Their coordinates are symmetrical relative to the point due to the central symmetry of the ellipsoid.

The pliability of the tool subsystem δ_n in this direction we find as a length (module) of the radius-vector K, carried out from the ellipsoid center to the point K

$$\delta_n = \left| \sqrt{(x_K^2 + y_K^2 + z_K^2)} \right|. \tag{9}$$

By replacing the third equation of the system (8) to the equation for the second ellipsoid, according to analogies we will receive the pliability of the subsystem of the part δ_m.

The complete pliability of the elastic system of the machine in this direction is determined as a total consolidation, that is $\delta_\Sigma = \delta_m + \delta_n$, or

$$\delta_\Sigma = \left| \sqrt{(x_{K_m}^2 + y_{K_m}^2 + z_{K_m}^2)} \right| + \left| \sqrt{(x_{K_n}^2 + y_{K_n}^2 + z_{K_n}^2)} \right|. \tag{10}$$

Deferring δ_Σ from the center O in this direction, we find a point on the guide line. The set of points determines the resulting surface of the relative pliability of the elastic system.

Thus, according to this algorithm, it is possible to construct the resulting surface of the static pliability for the elastic system of the machine as a whole.

Figure 4 shows the construction of a surface of two ellipsoids by the above-mentioned algorithm.

In this case, the ratio of the axes of the ellipsoid was taken as follows: for the first one—1:2:3, for the second one—1:3:2. The resulting surface is not like an ellipsoid. It reflects the distribution of the pliability of the system of a machine tool in space.

Figure 5 shows examples of simulation in the MathCad surfaces obtained for various variants of the relation of the axes of ellipsoids (within the limits of one order). As can be seen from the above images of forms of surfaces, the spatial pliability of the machine tools in different directions is changed substantially.

The proposed tensor-geometric model of spatial rigidity (pliability) in the form of a 3D surface makes it possible to determine the magnitude of rigidity (pliability) in any direction of the working space of the machine.

On the basis of a constructed 3D surface, you can find the values of the deformation of the machine in the cutting zone from the active power load, in particular,
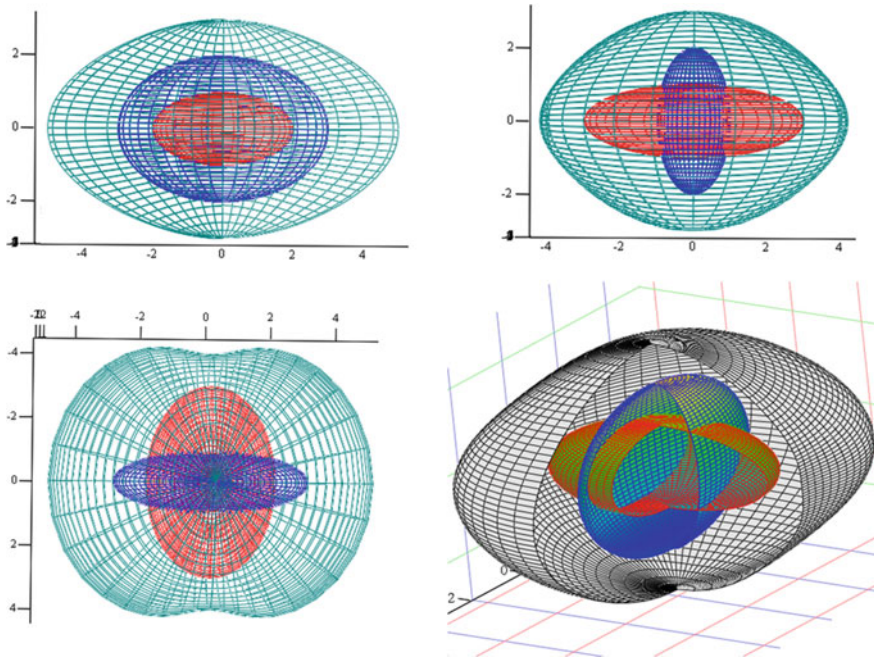
Fig. 4 Construction of the surface of the machine pliability from two ellipsoids (three orthogonal and isometric projections are shown)

the displacement of the tool from the detail that directly affects the accuracy of processing—the size and shape of the part.

Thus, the direct connection of spatial rigidity (pliability) of the machine with the precision of processing on the machine is established.

The developed method allows on the basis of the surfaces to evaluate and predict the accuracy of the construction of parts on the technological machine at known loads from the cutting strength (the limit of achievable accuracy), or to obtain the accuracy of processing was predicted to determine the technological processing modes.

3 Conclusions

1. Based on the tensor geometric approach to the description of the rigidity of the main subsystems of the machine tool (tools and parts) in the form of tensors (ellipsoids) rigidity (pliability), the estimated model of the elastic system of the technological machine to obtain its spatial rigidity (pliability) is determined. The surface of spatial rigidity (pliability) of an elastic system is expedient to build on the basis of ellipsoids of rigidity of the main subsystems (nodes) of the machine connected tool and the preparation.
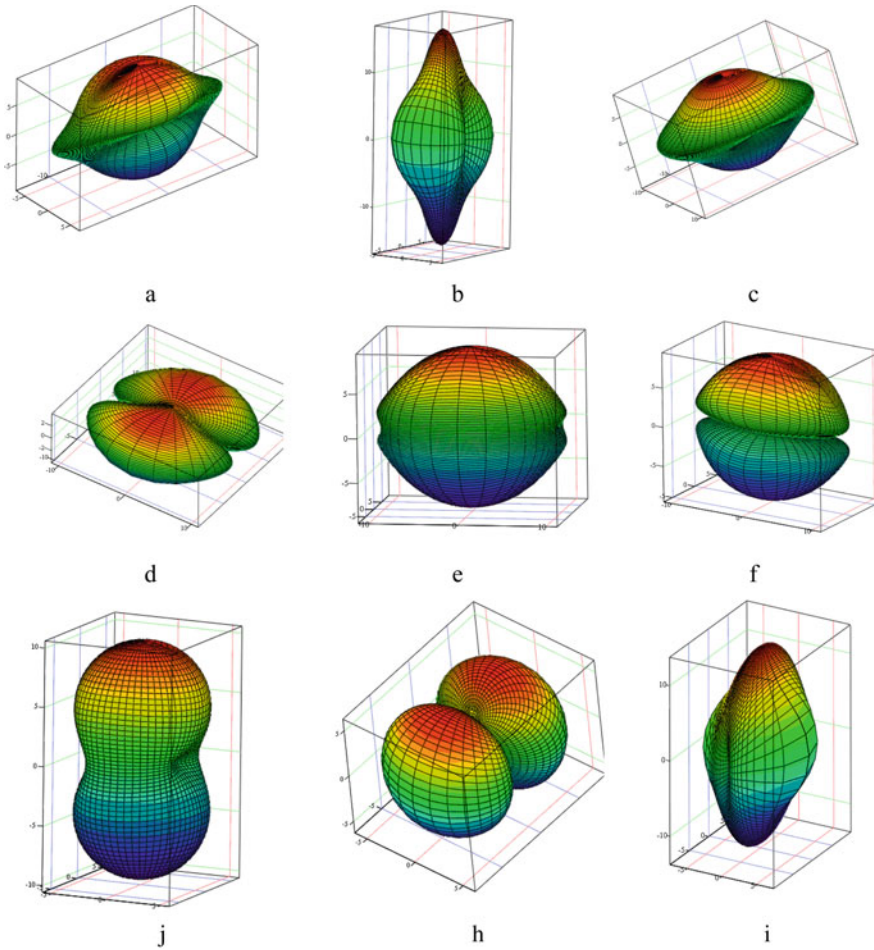
Fig. 5 Variants surfaces of machine pliability with different ratios of axles of two ellipsoids $a_1 : b_1 : c_1$ and $a_2 : b_2 : c_2$. **a** 2:8:1 and 1:7:8; **b** 2:1:8 and 1:7:8; **c** 7:8:1 and 1:7:8; **d** 7:2:1 and 1:7:2; **e** 7:2:1 and 1:7:8; **f** 7:2:1 and 1:7:8; **j** 2:1:2 and 1:4:8; **h** 2:1:2 and 1:7:1; **i** 1:7:8 and 6:1:5

2. The algorithm for finding the magnitude of static rigidity (pliability) in the cutting zone in an arbitrary direction of three-dimensional space is developed.
3. The surface of static pliability (rigidity) is entirely dependent on the parameters of the elastic system. When changing these parameters, the shape and size of the spatial pliability surfaces are also changing.
4. Anisotropy of spatial rigidity of the elastic system of the machine directly affects pliability of processing and determines the magnitude of the macro- and micro-errors of treated parts.

References

1. Aver'yanov O (1987) Modul'nyi printsip postroeniya stankov s ChPU (Modular design of numerically controlled machine tools). Mashinostroenie, Moscow (in Russian)
2. Poduraev Yu (2006) Mechatronics: foundations, methods, application: educational. Pos. for high schools. Mechanical Engineering, Moscow (in Russian)
3. Danilchenko Y, Kuznetsov Y (2003) Precision spindle nodes on rolling bearings. Economic Thought, Ternopil (in Ukrainian)
4. Kudinov V (1967) Machine dynamics. Mechanical Engineering, Moscow (in Russian)
5. Lebedev L, Cloud M, Veremejev V (2010) Tensor analysis with applications in mechanics. World Scientific Publishing Company, 363 p
6. Strutinsky V (2005) Tensor mathematical models of processes and systems: tutorial. ZhDTU, Zhytomyr (in Ukrainian)
7. Chupryna V, Rudyk A (2020) Tensor modeling of spatial rigidity of metal-cutting machine spindles. Mathematical modeling and simulation of systems (MODS'2020) selected papers of 15th international scientific-practical conference, MODS June 29–July 01, Chernihiv, Ukraine, pp 132–144
8. Ito Y (2008) Modular design for machine tool. McGraw-Hill, New York, 504 p
9. Strutinsky V, Kirichenko A (2009) Theoretical analysis of the stiffness of a six-coordinate mechanism of a parallel structure. Bulletin of the National Technical University of Ukraine "Kyiv Polytechnic Institute", Kyiv, pp 198–207 (in Ukrainian)
10. Korn G, Korn T (2000) Mathematical handbook for scientists and engineers: definitions, theorems, and formulas for reference and review. General Publishing Company, 1151 pp
11. Ilyin V, Poznyak E (2004) Analytical geometry: a textbook for high schools. Fizmatlit, Moscow (in Russian)

Mathematical Modeling of Heat Field in a Six-Strand Tundish During Filling

Kyrylo S. Krasnikov

Abstract The article is devoted to numerical research of temperature distribution in a six-strand tundish based on the Navier-Stokes equations. A long tundish is widely used at those metallurgical plants which are able to produce a large amount of iron from the convertors. However, such tundish requires specific geometry and technological parameters to work at the maximum effectiveness. The mathematical model takes into account the speed of the falling melt and the output speed. These speeds significantly affect size of stagnant zones. Also they affect heat exchange. The experiments show the importance of sufficiently high temperature of the falling melt from teeming ladle. It should be above 1870 K to avoid excessive melt freezing near the far wall of the tundish. The model allows easy modification of tundish geometry and number of strands to help metallurgists in finding rational ones.

Keywords Continuous casting tundish · Heat transfer · Navier-Stokes equations · Numerical experiment

1 Introduction

At steelworks a tundish is widely used for temporal storage of melt, which enters continuous casting machine. This work considers a long tundish with six strands. It speed ups production but requires rational technological parameters to maintain quality of steel. Temperature homogeneity in melt body is one of the important conditions for the process effectiveness. Also fluid speed should be not too small to avoid development of stagnant zones. The study of the mentioned process for optimization in industrial or laboratory conditions is associated with time and resource

K. S. Krasnikov (✉)
Dniprovskyi State Technical University, Kamianske, Ukraine
e-mail: kir_kras@ukr.net
URL: http://scitensor.com/

161

costs. Moreover, the study can include dozens of experiments. Mathematical modeling allows not only to reduce research costs, but also to avoid difficulties associated with opacity and high temperature of melt.

1.1 Literature Review

There are set of scientific works regarding to research of heat transfer in metallurgical tundish in various conditions based on numerical simulation or physical modeling.

Articles [1, 4] are devoted to studying influence of different embedded devices on melt flow in tundish. The main addressed problems are turbulent and short-circuit fluxes. In [1] author also present calculated temperature fields. Regions with the lowest temperature are near the top surface of melt.

Authors of [3] investigate three configurations of four-strands tundish to prove the importance of flow control. They use water tank similar to real tundish scaled to 1/2. There are obtained velocity profiles.

In [4] authors study hydrodynamic picture of melt using physical glass model of industrial tundish. Tracer is added to measure mixing time. They present figures with dependency of concentration on time.

In [5] authors consider five-strand tundish with baffle, turbulence inhibitor (rectangular and circular), and dams before farthest outlets. To investigate flow and heat transfer they use Navier-Stokes equations with standard k-epsilon model of turbulence. The paper presents parameters needed to simulate the process. Temperature distribution with isolines on the top surface of melt and on the front wall is presented. Also it is calculated a total temperature drop for different conditions. Conducted numerical experiments show a good potential for optimization of the process at plant.

In [6, 7] authors use mathematical model to study heat distribution taking into account heat losses from walls and free surface of melt. It is presented comparison of industrial measurements and predicted temperature [6]. In [8] author applies computational methods of FVM and SPH to various industrial vessels. It is stated that melt is non-isothermal [9], but with modifications the temperature difference can be decreased.

In [10] authors consider influence of different turbo-stoppers on melt hydrodynamics in tundish using water model. It is presented results of velocity measurements in various tundish configurations, which show ability to reduce number of unneeded swirls. The results are used to validate mathematical model.

Work [11] contains cross-sections of calculated temperature and velocity fields for single-strand tundish. Authors study a dependency of inlet cooling rate on heat distribution.

In [12, 13] authors optimize tundish geometry using Navier-Stokes equations with buoyancy term dependent on temperature difference. Authors conclude that flow control barriers positively influence tundish effectiveness and melt purification. They present temperature contours for three cases ranging from few barriers to zero.

There are metallurgical plants, which use a big eight-strand tundish. In [14] author considers various flow controllers (baffles, impact pads, dams) and numerically models dynamics of fluid based on Navier-Stokes equations. As result designs of controllers which lead to short circuits of flow were canceled. In [15] the physical and numerical modeling of molten steel dynamics in tundish with tunnel filters is presented. Author describes tundish geometry in details to clarify the position of the filter as well as baffle. In tables thermophysical properties of molten steel are listed, which can be used as initial data in numerical experiments, for example, value of heat flux at walls and slag layer. Presented results prove effectiveness of filter installation for removing of small inclusions. Comparison between numerical results and experimental one is based on RTD curves and gives sufficient correspondence.

2   Mathematical Model

The model predicts the process using conservation laws of physics—conservation of impulse and energy.

2.1   Assumptions

The following assumptions are made for the process:

1. Molten steel is an incompressible viscous Newtonian fluid represented by continuum.
2. Tundish walls are rectangular except curved one near falling jet (Fig. 1). The curviness modeled by circle. The jet has cylindrical shape and falls vertically. The barrier has shape of ellipse and occupies the whole height of melt (Fig. 2. Using mirror symmetry with respect to plane Y-Z only the half of tundish needs modeling.
3. Top surface of molten steel is flat.
4. Influence of slag on melt hydrodynamics is negligible.
5. Speed at outlets is constant.
6. Temperature gradients are small enough to influence only vertical motion of melt using Boussinesq model (buoyancy).
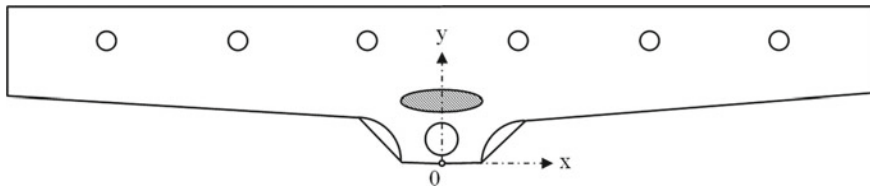
Fig. 1 Top view of a six-strand tundish

Fig. 2 Vertical cross-section
of melt body at x = 0
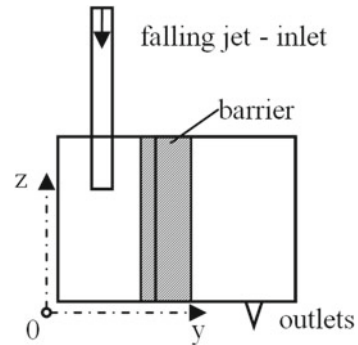
2.2 Equations

The melt velocity **v** is expressed by Navier-Stokes equations with addition of Boussinesq model for buoyancy:

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = \nu_e \Delta \mathbf{v} - \nabla P - \mathbf{g}\beta\left(T - T_0\right) \tag{1}$$

$$\nabla \cdot \mathbf{v} = 0 \tag{2}$$

where ν_e—effective kinematic viscosity, P—kinematic pressure, **g**—gravitational acceleration, β—thermal expansion coefficient of liquid steel (1.27×10^{-4} [2]), T_0—average melt temperature.

Temperature distribution is predicted by heat equation with advection term:

$$\frac{\partial T}{\partial t} + \nabla \cdot (T\mathbf{v}) = \frac{D_T}{C\rho}\nabla^2 T \tag{3}$$

where T_m—temperature of melt; C—specific heat capacity, ρ—density of melt; D_T—effective coefficient of diffusion.

2.3 Boundary Conditions

Equations are complemented by physically-based conditions. For velocity field there is no-slip condition on solid surfaces and impermeability condition—on all surfaces:

$$\mathbf{v} \times \hat{\mathbf{n}} = 0 \tag{4}$$

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \tag{5}$$

At areas occupied by jet and holes there are inlet and outlet conditions:

$$v_{jet} = -2.8 \, \text{m/s} \tag{6}$$

$$v_{hole} = -0.3 \, \text{m/s} \tag{7}$$

For temperature field it is set a heat transfer at the wall and the top surface of melt. Also there is zero heat flux at plane of symmetry x = 0:

$$T_{jet} = 1873 \, K \tag{8}$$

$$\frac{\partial T}{\partial \hat{\mathbf{n}}} |_{x=0} = 0 \tag{9}$$

$$\frac{\partial T}{\partial \hat{\mathbf{n}}} = k(T - T_{surf}) \tag{10}$$

where $\hat{\mathbf{n}}$—normal to surface; k—heat transfer coefficient; T_{surf}—temperature of wall or slag.

Curvy wall is modeled by combination of arc and line expressed by piece-wise function of x-coordinate:

$$f(x) = \left\{ \begin{array}{ll} \sqrt{R^2 - (x - 2R)^2} & x <= 2R \\ R - \frac{2R-x}{15} & x > 2R \end{array} \right\} \tag{11}$$

3   Results

The mathematical model is implemented in multithreaded software using CSharp programming language. The equations are discretized using a central difference scheme for spatial coordinates. The number of discretization steps is $174 \times 68 \times 35$. The explicit numerical scheme allows easy parallelization of calculations dividing computational domain on equal chunks assigned to threads. The software post-process and renders calculated 3D vector and scalar fields on the screen to get figures and to analyze melt state at any recorded moment of time.

Table 1 Geometric parameters

Object	Properties		
Melt body	Length—3.480 m	Width—1.32 m	Height—0.70 m
6 holes	Diameter—0.038 m	Interval—1.10 m	Y coord—0.90 m
Jet	Diameter—0.090 m		Y coord—0.22 m
Barrier	Length—0.790 m	Width—0.16 m	Y coord—0.52 m

Table 2 Thermophysical parameters of molten steel, which can be found, for example in [16]

Property	Value	Units
Initial temperature	1873	K
Density	7000	kg/m^3
Specific heat capacity	840	J/(kg K)
Thermal conductivity	41	W/(m K)
Viscosity	7	$\times 10^{-3}$ Pa s

Numerical experiment was conducted using real process parameters, which are summarized in Tables 1 and 2.

Figure 3 shows velocity distribution at the half of melt depth. As theoretically predicted the largest speed is near falling jet. Fast flows create the swirl behind barrier. As expected, the barrier slows down quick fluxes of melt. It is seen a large stream moving near curved wall. Thanks to this stream, heat is delivered to the far corners of the tundish. Also the stream prevents appearing of stagnant zones with too low speed. A region near the right wall has the lowest temperature and can reach solidus eventually.
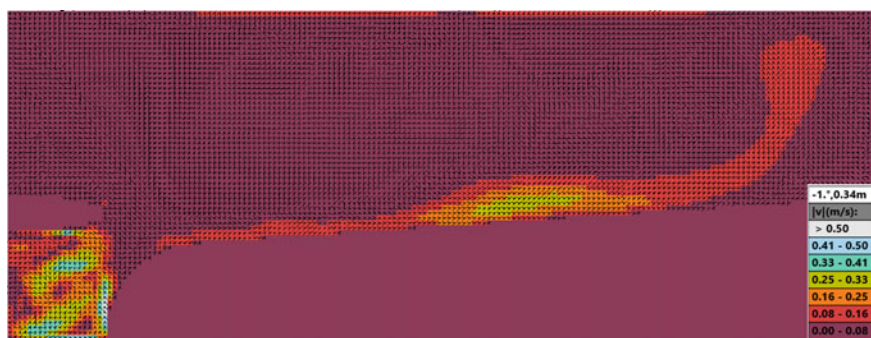
Fig. 3 Horizontal cross-section of velocity at 0.34 m of height after 60 s. Black arrows show direction of streams. Colored background of the arrow represent magnitude of velocity. Average speed of melt is 0.09 m/s
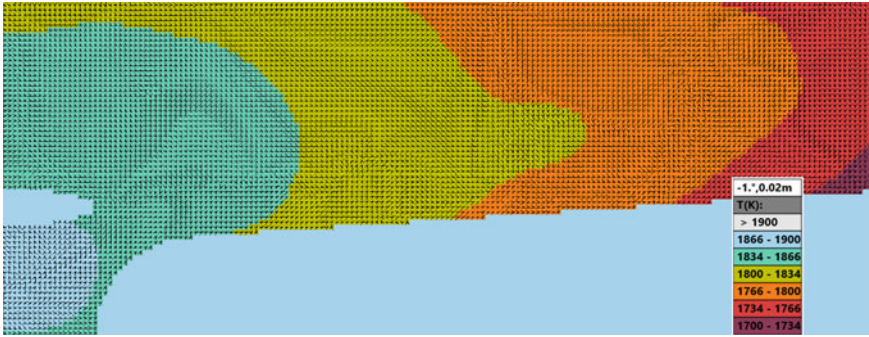
Fig. 4 Cross-section of temperature field at 0.02 m of height after 60 s. Black arrows follow fluxes. Colored background of the arrows represent amount of thermal energy
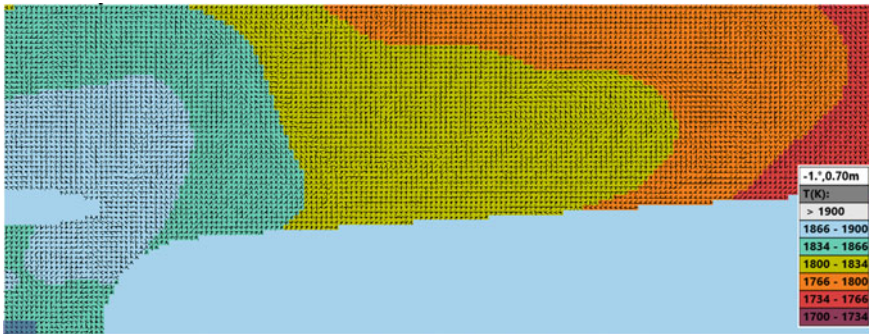
Fig. 5 Cross-section of temperature field at 0.7 m of height after 60 s

Figure 4 shows temperature gradient from the plane of symmetry x = 0 to the right. The highest temperature region is near falling jet. Barrier doesn't significantly prevent heat transfer by width. A region near the right wall has the lowest temperature and can reach solidus eventually.

On Fig. 5 the high temperature region expands farther than on Fig. 4. One can see lower temperature near walls than at center of tundish, because wall takes heat from melt.

Table 3 and Fig. 6 present evolution of temperature near outlets. At the closest outlet it slowly decreases with some fluctuations caused by fluxes from jet. Sometime, temperature at the second outlet can even be higher than at the first one (from 15 to 25 s on Fig. 6). The second and the third outlet lose heat considerably faster.

At Table 4 and Fig. 7, the calculated speed evolution near outlets is given. The time interval of 5 s was selected as sufficiently low to view overall tendency of value without excessive set of numbers.

Table 3 Calculated temperature values (in Kelvins) at the three outlets (Fig. 6)

Outlet	5s	10s	15s	20s	25s	30s	35s	40s	45s	50s	55s	60s	65s	70s
1st (0.6 m)	1849	1849	1847	1840	1840	1846	1847	1845	1843	1841	1839	1839	1838	1837
2nd (1.7 m)	1848	1841	1846	1844	1838	1827	1818	1817	1812	1805	1797	1788	1782	1778
3rd (2.8 m)	1848	1840	1831	1822	1814	1801	1802	1796	1787	1772	1761	1754	1746	1741
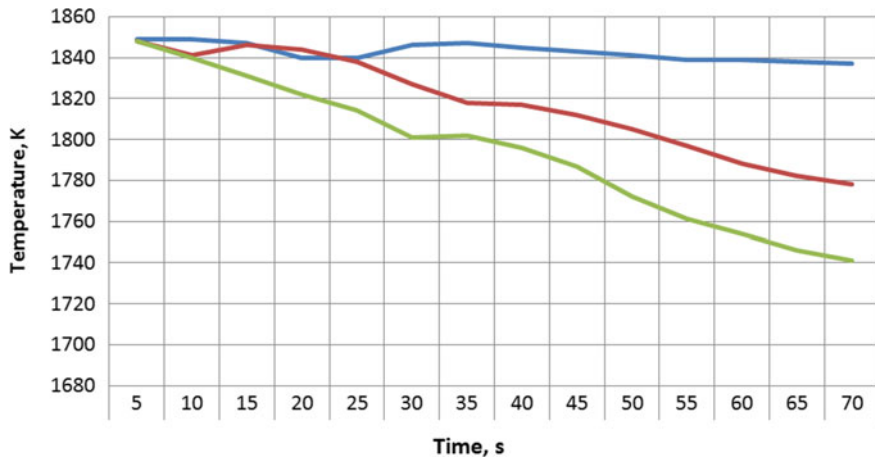
Fig. 6 Time-dependency of the temperature near each of three outlets. The blue line corresponds to the first outlet (0.6 m), the red line—to the second outlet (1.7 m), and the green line—to the third one (2.8 m)

Table 4 Calculated values of speed (in m/s) at the three outlets (see also Fig. 7)

No.	5s	10s	15s	20s	25s	30s	35s	40s	45s	50s	55s	60s	65s	70s
1st	0.203	0.204	0.203	0.202	0.213	0.216	0.212	0.21	0.208	0.207	0.207	0.208	0.213	0.214
2nd	0.202	0.206	0.216	0.217	0.208	0.204	0.209	0.218	0.218	0.219	0.22	0.223	0.219	0.218
3rd	0.203	0.202	0.202	0.203	0.2	0.198	0.226	0.208	0.206	0.203	0.215	0.218	0.214	0.214

To validate calculation the model is slightly modified using parameters and results presented in work [5], where authors use a similar tundish with five strands. Differences in results are small enough (3–9%) to recommend further usage of mathematical model (Fig. 8).

4 Conclusions

Heat distributions in molten steel are calculated using presented 3-dimensional mathematical model of melt hydrodynamics and heat transfer in tundish during filling. The model takes into account nontrivial geometry of melt body, buoyancy caused by temperature gradients and heat losses on the top surface and on the tundish walls.
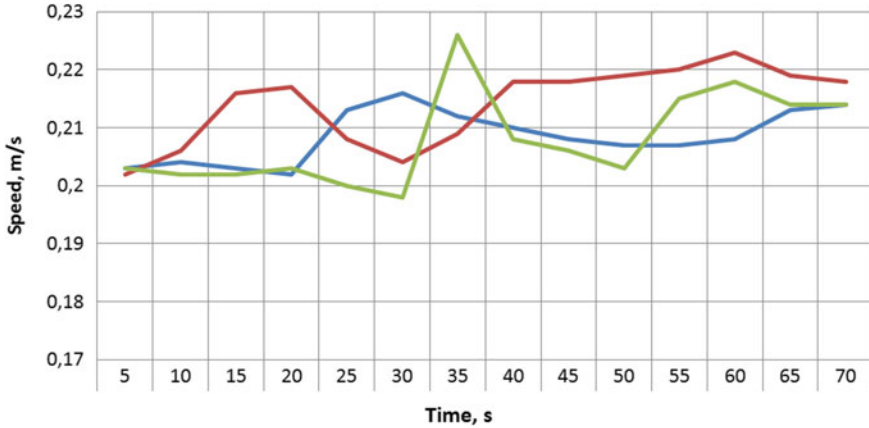
Fig. 7 Time-dependency of the melt speed near three outlets. The blue line represents speed at the first outlet (0.6 m), the red line—at the second one (1.7 m), the green line—at the third outlet (2.8 m)

Fig. 8 Vertical cross-section of predicted heat field (in Kelvins) at the line of outlets (y = 1.3 m) for comparison with data in [5]

Results of experiments show that temperature of melt in tundish can decrease significantly, especially, near right wall. So a jet temperature should be above 1870 K to have enough reserve of heat.

Speed of falling jet greatly affects flow activeness in the melt body. Potential stagnant zone is at the farthest bottom corner at the half of width, where speed is lower than 0.01 m/s (Fig. 4). As seen on Fig. 7, the speed near outlets just slightly fluctuates, because the vertical component of it is the main contributor. As expected, the mean speed is increased with time.

The flexibility of the model allows changing tundish geometry as well as count of outlets and other parameters according to industrial needs.

References

1. Chang S, Zhong L, Zou Z (2015) Simulation of flow and heat fields in a seven-strand tundish with gas curtain for molten steel continuous-casting. ISIJ Int 55(4):837–844
2. Chattopadhyay K, Isac M, Guthrie R (2012) Modelling of non-isothermal melt flows in a four strand delta shaped billet caster tundish validated by water model experiments. ISIJ Int 52:2026–2035

3. Cloete JH, Akdogan G, Bradshaw SM, Chibwe DK (2015) Physical and numerical modelling of a four-strand steelmaking tundish using flow analysis of different configurations. J S Afr Inst Min Metall 115(5):355–362. www.scopus.com

4. Cwudzinski A, Jowsa J, Gajda B (2020) Physical simulations of macro-mixing conditions in one strand tundish during unsteady period of continuous slab casting sequence. Steel Res Int 91

5. Fang Q, Zhang H, Luo R, Liu C, Wang Y, Ni H (2020) Optimization of flow, heat transfer and inclusion removal behaviors in an odd multistrand bloom casting tundish. J Mater Res Technol 9(1):347–363. https://www.sciencedirect.com/science/article/pii/S2238785419310713

6. Gaston A, Sanchez Sarmiento G, Sylvestre J (2008) Thermal analysis of a continuous casting tundish by an integrated fem code. Lat Am Appl Res 38:259–266

7. Glaser B (2012) A study on the thermal state of steelmaking ladles. Doctoral thesis, Royal Institute of Technology, Stockholm. https://www.diva-portal.org/smash/get/diva2:559686/FULLTEXT01.pdf

8. Hosain ML (2018) Fluid flow and heat transfer simulations for complex industrial applications: from Reynolds averaged Navier-Stokes towards smoothed particle hydrodynamics. Ph.D. thesis, Malardalen University, Future Energy Center

9. Jiang GZ, Kong GY, Li GF, Liu YL (2007) Simulation research on temperature distribution in tundish, vol 42, pp 27–29

10. Odenthal HJ, Bolling R, Pfeifer H, Holzhauser JF, Wahler FJ (2001) Mechanism of fluid flow in a continuous casting tundish with different turbostoppers. Steel Res Int 72:466–476

11. Qu TP, Liu CJ, Jiang MF (2012) Numerical simulation for effect of inlet cooling rate on fluid flow and temperature distribution in tundish. J Iron Steel Res Int 19(7):12–19. https://www.sciencedirect.com/science/article/pii/S1006706X12601071

12. Rocha J (2017) Modeling and numerical simulation of fluid flow and heat transfer of a steel continuous casting tundish. Thesis. http://bdtd.ibict.br/vufind/Record/UFC-7_2b01dfd1ebce69c4ee094b668da5a113

13. Rocha J, Souza E, Marcondes F, Castro J (2019) Modeling and computational simulation of fluid flow, heat transfer and inclusions trajectories in a tundish of a steel continuous casting machine. J Mater Res Technol 8(5):4209–4220. https://www.sciencedirect.com/science/article/pii/S2238785419307537

14. Srivastava A, Nurni V (2018) Numerical simulation of flow behaviour of liquid steel in an eight strands tundish of billet caster at JSPL, Angul. In: Proceedings of Asia steel international conference

15. Wang XY, Zhao DT, Qiu ST, Zou ZS (2017) Effect of tunnel filters on flow characteristics in an eight-strand tundish. ISIJ Int 57(11):1990–1999

16. Wilthan B, Schutzenhofer W, Pottlacher G (2017) Thermophysical properties of five industrial steels in the solid and liquid phase. Int J Thermophys 83:1–15

Mathematical Modeling and Simulation of Systems in Information Technology and Information Security

Hash Method for Information Stream's Safety in Dynamic Cooperative Production System

Andrii Yarmilko ⓘ**, Inna Rozlomii** ⓘ**, and Hryhoriy Kosenyuk** ⓘ

Abstract The article focuses on the enhancing methods of information security of the cooperative productive systems with dynamic communication clusters. There is a disappearance of physical barriers between people and robots in the process of their production interaction in the age of development of production systems. The problem is the lack of proper trust in the actions of the robot in cooperation. One of the steps towards achieving a proper level of trust and cooperation is to ensure the credibility of the information in the common communication system. In this research we assumed the messages exchanged by the components of the situational cluster are information blocks composed of an authentication and information component. Method that combine the principles of password authentication and identification with hashing methods have been proposed to analyze information. The method of frequent error detection and a method provides a guaranteed detection of a specified number of counterfeit fragments in the message by entering control blocks of information and performing cross hashing is offered.

Keywords Cooperative production system · Matrix cryptographic transformation · Hash function

1 Introduction

The characteristic feature of the modern human–machine systems (HMS) is combination in one productive system of people and technical devices, and especially robotic modules of different level of intellectuality and autonomy that envisages their co-operation in the process of implementation of the applied tasks. On the whole, run into the phenomenon of hybrid intellect, that is characterized by possibility of forming and use of knowledge beyond hard control of man.

The functioning of such system acquires a new level of complexity to implement tasks in productive open-space that does not provide the physical isolation of such

---

A. Yarmilko (✉) · I. Rozlomii · H. Kosenyuk
Bohdan Khmelnytsky National University of Cherkasy, 81, Shevchenko Blvd., Cherkasy 18031, Ukraine
e-mail: a-ja@vu.cdu.edu.ua

co-operative formation from the strangers of productive and unproductive factors. Under such conditions, it is necessary to solve the problems of the formed system dependability of components with natural and artificial intelligence for solving tasks both of purely industrial purpose and related to its internal and external security. One of the current challenges in the context of this problem is providing defense of intrasystem communications as precondition of cooperative system dependability on the whole.

As the described type system consists of intellectual components of different nature, in the process of performing the task assigned to it, its components can change their internal structure, in accordance with the current configuration of production and non-production factors, forming situational communication clusters, in which composition and information communication system dynamically changes according to many significant parameters.

We had the aim to develop method of information security of cooperative production systems with dynamic communication clusters and researching its properties.

2   Related Works

The attempts of creation autonomous, effective and safe mobile robots for the unstructured environments have old history. In particular, the precedent of researching robot's behavior in urbanized open-space was examined in [1]. Not only robot's planning its own work to obtain the necessary visual information about the real world, but also assessing his own behavior in terms of the possibility of its implementation in real time and visual control were addressed. Another area of research is to find methods for managing the coordinated movement of mobile robot groups, in particular, by consolidated movement based on decentralized algorithms for applying to a group of autonomous quadcopters [2]. Other precedents are the warehouse robotic systems. Indeed, the similar systems are often characterized by the greater or less level of the regime regulation of the functional environment (presence of rules of behavior, access or physical space restrictions) which helps to reduce the level of anarchic state of communication tasks.

But we watch the disappearance of physical barriers between people and robots in the process of their production interaction with the development of production systems on the whole. In terms of interaction with robots, Industry 4.0 points the relevant humans to consider that are working alongside the machines, and other staff who may be nearby. It has become clear that cooperation and collaboration are different types of interactions in human-to-work collaboration (HRC) with large conceptual differences between them. Proposed classification of levels of collaboration for HRC in production, formed to provide a conceptual model, called the levels of collaboration (LoC) [3]. In previous years [4] and now [5] work is underway to obtain realistic applications for the interaction of robot outside the traditional industrial environment with capability to supporting HRC inside sociotechnical systems.

However, the consumer response to common workspace for humans and robots, as noted by Christian Tarragona, Senior Vice President at KUKA Research and Development, is quite frank: "As soon as they have to start thinking about safety fencing, most of them opt out" [6]. The problem is the lack of proper trust in the collaboration with robots in common workspace. When humans interact with each other, all aspects of the interaction occur relatively quickly and effortlessly and for the most part it occurs subconsciously through their mindreading. There is a great challenge to achieve a similar mutual and fluent interaction between humans and robots in collaboration [3]. One of the steps towards achieving this level of trust and cooperation or collaboration is ensuring trust in the reliability of information in a common communication system.

3 Problem Statement

Talking about cooperative production systems with hybrid intelligence used in open space, the following main features and properties can be noted:

(1) presence in general productive space of the cooperative system of static and dynamic objects that are not part of it, but that can be the permanent or situation sources of hindrances for this system, including in communication processes;

(2) corresponding to the aims system, to productive activity not all of its components may be involved at different time intervals; such components expedient to combine in corresponding situation productive clusters and form corresponding communication clusters for their co-operation;

(3) in system's composition, at the same time a few clusters can function with the own system of communication;

(4) the channels of information exchange that are actually accessible and appropriate for the organization of situational production clusters in systems of the mentioned type are multimodal and open.

The problem of maintaining an adequate level of security and dependability remains relevant in the functioning process of the cooperative production system and its clusters. The proposal to reduce the level of formalization of the cluster structure by the human operator seems a rational step. It means that the overall statement of the production task is formulated at the command management level of the cooperative system. Collective Artificial Intelligence forms the cluster for a specific task at the strategic and tactical levels, independently or with the help of a person, and provides this cluster with the means of secure communication in an open space. Clustering can be performed on a production or other sign (Fig. 1).

The purpose of clusters would be to prevent (or minimize) third-party interference (intentional or accidental) in the channels of information exchange of components of such systems. We may be interested not only in identifying errors when performing information analysis between cluster components, but also in identifying the differences between two information fragments or their instances. If the purpose is to
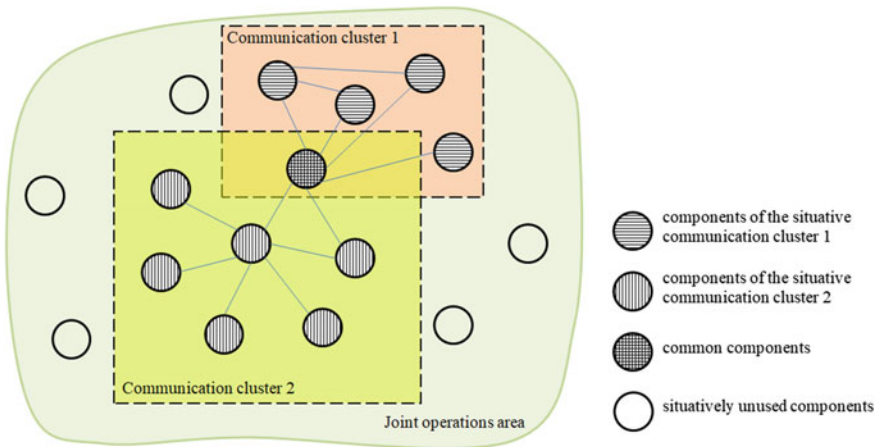
Fig. 1 An example of a cooperative production system structure with two situational clusters

control the storage of information in the communication process, then the difference will be treated as a mistake with all possible consequences of reacting to it. If we are interested in the structure or dynamics of changes in information fragments, then the differences will indicate the presence of dynamic processes, their localization and other related parameters. Therefore, considering the information security of the situational cluster of the human–machine system, we must ensure the control of the components of the system over the sources of receipt of messages and the detection of damage or change of information in them. Authentication and user identification tools are intended to provide such a cluster with secure communication in open space. Authentication provides validation of the subject [7].

One of the most common methods of authentication is password authentication [8]. First of all, this is due to the easiness of implementation and the low cost of maintaining such technology [9, 10].

In this research we assumed that the messages exchanged by the components of the situational cluster are blocks of information made up of an authentication and information component. After the authenticity of the authenticator is established, it is necessary to check the absence of falsifications in the information part of the sent message, and in case of their detection to perform the restoration of the valid information.

Analysis of the information content part of the message can also be partially or completely performed by cryptographic methods. Development of the method is necessary as they allow to detect several changes in a separate block of information and to set their coordinates for the purpose of identification and formation of the answer, as a separate case of such analysis may be establishing fact and area of localization of the change in the data block. Method that combines the principles of password authentication and identification with hashing methods to analyze shared information resources was proposed. In this research, we did not focus on the format

and size of the authentication and information components of the message. Therefore, in the future they will be considered as blocks of data, given the uniformity of the tasks of detecting falsification in each of the message components.

4 New Method for Detecting and Localization Multiple Errors in Blocks of Information and for Restoring the Integrity of the Primary Messages

4.1 Method of the Surplus Hashing of Information Control Block

The introduction of redundancy enables the detection and correction of errors in the information that is transmitted and can be modified during transmission. It is advisable to use linear code theory namely Hamming codes to identify errors in authentication data [11]. By definition, Heming codes allow you to detect a double error and correct a single error in a binary code by entering redundant information. In particular, this capability (single-bit error detection) is not enough for cryptographic purposes and for the authentication procedure. Therefore, building codes that detect more errors is appropriate.

Hamming introduced the code distance d and minimum code distance d_0 to evaluate the corrective capability of the codes and show their dependence on the code length and the redundancy entered [12]. According to Hamming, the minimum code distance corresponds to the number of positions, which distinguish two bitwise code sequences by modulo 2 addition. It is proved that the minimum code distance characterizes the corrective properties of the noise immunity code. If the two code sequences differ from each other in t ($t \geq 1$) positions (digits, symbols), but from all other code sequences of this code set will differ by more than t positions, then a minimum code distance to detect t errors must be provided

$$d_0 \geq 2 * t + 1. \tag{1}$$

If a condition (1) is executed, then an interference-resistant code is guaranteed to correct

$$t_{corr} \leq (d_0 - 1)/2 \tag{2}$$

erroneous symbols or to find $t_{det} \leq d_0 - 1$ erroneous binary symbols.

Note also that any $G_{(n, k)}$ the Hamming code in the general form can be given by a generating matrix consisting of a identity matrix $k \times k$ and the assigned matrix test elements $k \times r$:

$$G_{(n,\,k)} = \begin{vmatrix} 1\ 0\ 0\ 0\ \ldots 0 & c_{11}\ c_{12}\ c_{13}\ c_{14}\ \ldots\ c_{1r} \\ 0\ 1\ 0\ 0\ \ \ldots 0 & c_{21}\ c_{22}\ c_{23}\ c_{24}\ \ldots\ c_{2r} \\ 0\ 0\ 1\ 0\ \ \ldots 0 & c_{31}\ c_{32}\ c_{33}\ c_{34}\ \ldots\ c_{3r} \\ 0\ 0\ 0\ 1\ \ \ldots 0 & c_{41}\ c_{42}\ c_{43}\ c_{34}\ \ldots\ c_{4r} \\ .\quad.\quad.\quad.\quad.\quad. & .\quad.\quad.\quad.\quad.\quad.\quad. \\ .\quad.\quad.\quad.\quad.\quad. & .\quad.\quad.\quad.\quad.\quad.\quad. \\ 0\ 0\ 0\ 0\ \ldots 1 & c_{k1}\ c_{k2}\ c_{k3}\ c_{k4}\ \ldots\ c_{kr} \end{vmatrix} \qquad (3)$$

where $c_{11}\ldots c_{kr}$ are verification elements of the added matrix, k is the number of rows of the generating matrix (corresponds to the dimension of the identity matrix), r is the number of columns of the generating matrix. The values of k and r must satisfy two conditions: first is to obtain a generating matrix of the smallest size and, second is to provide the minimum code distance required to detect a specified number of errors. To determine the values of the check elements of the right side of the matrix, we must start from the basic properties of the correction codes [13].

Since each row of the identity matrix $k \times k$ has only one unit, then the weight of each row of the assigned matrix must not be less than $d - 1$, namely, the sum of modulo two of the two rows of the matrix must not be less than $d - 2$ for a guaranteed fix of a one-off error. In addition, the combinations of the right side of the matrix must be linearly independent [14]. Therefore, these principles of redundant coding were the basis of the method.

In addition to Hamming theory, hashing techniques are another component of the method of guaranteed fraud detection. Using hashing methods to ensure the integrity of information has several advantages, including:

- relatively low excess;
- a small number of cryptographic transformations;
- the ability to control the length of the hash code.

There are so many options for constructing hash functions for example based on the use of matrix cryptographic transformations described in writing [15].

We introduce a notation to describe the method. Let the information be presented in the form of records of arbitrary size, denote them by the set $A = \{\vec{a}_1,\ \vec{a}_2,\ ...,\ \vec{a}_n\}$ – authentication data where $\vec{a}_1,\ \vec{a}_2,\ \vec{a}_3,\ ...,\ \vec{a}_n$ in accordance set of binary vectors, blocks of information of arbitrary size. $F = \{\vec{f}_n,\ \vec{f}_{n+1},\ \ldots,\ f_{n+k}\}$– set of binary vectors, fixed size hash codes calculated by $f_i = h(a_i)$, where $i \in [1, n]$. The value of hash functions of each information block is calculated according to any algorithm, offered in the article [15]. Set of the possible hashing schemes of the blocks a_1, a_2, a_3, ..., a_n, present as a binary matrix:

$$F = \begin{pmatrix} f_{11} & f_{12} & \ldots & f_{1n} \\ f_{21} & f_{22} & \ldots & f_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ f_{m1} & f_{m2} & \ldots & f_{mn} \end{pmatrix}$$

where each row corresponds to a specific hash scheme.

The following conditions are satisfied for the rows of the matrix:

- there are no null rows of the matrix;
- all rows of the matrix are linearly independent;
- there is a minimum code distance between the rows of the matrix.

Coding theory for the generating matrix also has the following properties, which makes it possible to use linear code building rules to construct hash codes.

A hash code system is a set of hash codes that are obtained by implementing any algorithm for calculating a hash function in the order defined by a special procedure for selecting records (blocks of information) based on a mathematical apparatus of linear algebra.

4.2 An Algorithm for Constructing Hash Codes to Ensure Message Integrity

Then the protected block resulting from the hashing will look like this: $(\vec{a}_1 \, \vec{a}_2 ... \vec{a}_{n+1} \, \vec{a}_{n+l}) \rightarrow (\vec{a}_1 \, \vec{a}_2 ... \vec{a}_{n+1} \, \vec{a}_{n+l} \, \vec{f}_{n+l+1} \, \vec{f}_{n+r})$, where "$\rightarrow$" special multi-dimensional noncommutative hashing operation.

Then the protected block, received after hashing, will look like this:

$$(\vec{a}_1 \, \vec{a}_2 ... \vec{a}_{n+1} \, \vec{a}_{n+l}) \otimes \begin{vmatrix} 1\ 0\ 0\ 0 \ldots 0 & c_{11}\ c_{12}\ c_{13}\ c_{14} \ldots c_{1r} \\ 0\ 1\ 0\ 0 \ldots 0 & c_{21}\ c_{22}\ c_{23}\ c_{24} \ldots c_{2r} \\ 0\ 0\ 1\ 0 \ldots 0 & c_{31}\ c_{32}\ c_{33}\ c_{34} \ldots c_{3r} \\ 0\ 0\ 0\ 1 \ldots 0 & c_{41}\ c_{42}\ c_{43}\ c_{34} \ldots c_{4r} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ 0\ 0\ 0\ 0 \ldots 1 & c_{k1}\ c_{k2}\ c_{k3}\ c_{k4} \ldots c_{kr} \end{vmatrix}$$

$$= (\vec{a}_1 \, \vec{a}_2 ... \vec{a}_{n+1} \, \vec{a}_{n+l} \, \vec{f}_{n+l+1} \, \vec{f}_{n+r})$$

where $\vec{A}_v \otimes G_{(n,k)} = (\vec{a}_1 \, \vec{a}_2 ... \vec{a}_{n+1} \, \vec{a}_{n+l} \, \vec{f}_{n+l+1} \, \vec{f}_{n+r})$, the symbol "$\otimes$" is special multidimensional noncommutative hashing operation of the information blocks of the message. The term "syndrome" is used to control the integrity of information in the theory of linear codes and means features characteristic's set of particular phenomenon. Syndrome vector, which may have errors, allows us to recognize the most likely nature of these errors.

Mistake in the block $(\vec{a}_1 \, \vec{a}_2 ... \vec{a}_{n+1} \, \vec{a}_{n+l} \, \vec{f}_{n+l+1} \, \vec{f}_{n+r})$, which is the object of information protection, in our understanding will be the result of the discrepancy of the binary vector of the result obtained by checking the syndrome.

Checking the integrity of data in blocks of information involves the following steps:

(1) input is a data block $(\vec{a}_1^* \vec{a}_2^* ... \vec{a}_{n+1}^* \vec{a}_{n+l}^* \vec{f}_{n+l+1}^* \vec{f}_{n+r}^*)$, which is checked for integrity; a block hash operation must be performed;

(2) calculate the syndrome corresponding to the value of the predicate:

$$P(\vec{a}_n) = \begin{cases} 1, \; if \; \vec{f}_n^* = \vec{f}_n; \\ 0, \; if \; \vec{f}_n^* \neq \vec{f}_n. \end{cases}$$

(3) according to the syndrome table, errors in blocks of information must be corrected.

4.3 An Example of Fixing the Multiple Error

The obtained method has the potential to provide a guaranteed error correction of any multiplicity. Let us consider using it as an example of fixing a double error (t = 2). In this case, the minimum code distance between the rows of the generating matrix, according to (1), was $d_0 \geq 5$.

According to expression (3), one example of the generating matrix can be represented in the following form:

$$G_{(9, 4)} = \begin{vmatrix} 1\;0\;0\;0 & 0\;0\;0\;1\;0 \\ 0\;1\;0\;0 & 0\;0\;1\;0\;1 \\ 0\;0\;1\;0 & 1\;1\;0\;1\;1 \\ 0\;0\;0\;1 & 1\;1\;1\;0\;0 \end{vmatrix}$$

We calculate the minimum code distance between the rows of the generating matrix. To do this, we denote the rows of the matrix $G_{(9,4)}$, through a x_1, x_2, x_3, x_4. The calculations and values of the minimum code distances between the rows of the generating matrix are presented in Table 1.

Calculations show, the minimum code distance is stored between all code sequences (rows of the matrix) $d_0 \geq 5$, this makes it possible to assert that according to (2) this code will contribute to the guaranteed correction of double error in blocks of information. According to the matrix, hash codes are built due to the following rules:

Table 1 Calculation d_0 between the rows of the matrix $G_{(9,4)}$

x_1 100000010	x_1 100000010	x_1 100000010
$\oplus$	$\oplus$	$\oplus$
$\dfrac{x_2 010000101}{d_0=11111}$, $d_0 = 5$	$\dfrac{x_3 001011011}{d_0=11111}$, $d_0 = 5$	$\dfrac{x_4 000111100}{d_0=111111}$, $d_0 = 6$
x_2 010000101	x_2 010000101	x_3 001011011
$\oplus$	$\oplus$	$\oplus$
$\dfrac{x_3 001011011}{d_0=111111}$, $d_0 = 6$	$\dfrac{x_4 000111100}{d_0=11111}$, $d_0 = 5$	$\dfrac{x_4 000111100}{d_0=11111}$, $d_0 = 5$

$$F^k_{(9,\,4)} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_3 \oplus x_4 \\ x_3 \oplus x_4 \\ x_2 \oplus x_4 \\ x_1 \oplus x_4 \\ x_2 \oplus x_3 \end{pmatrix}$$

According to the generating matrix, $G_{(9,4)}$, we have four information blocks $\{A_1...A_4\}$ and five controls $\{F_1...F_5\}$ (Fig. 2), which values are obtained by the formulas:

$$F_1 = F(A_3) \cup F(A_4);$$
$$F_2 = F(A_3) \cup F(A_4);$$
$$F_3 = F(A_2) \cup F(A_4);$$
$$F_4 = F(A_1) \cup F(A_3);$$
$$F_5 = F(A_2) \cup F(A_3).$$
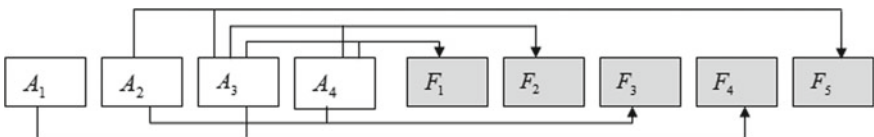
Fig. 2 Scheme for obtaining hash functions

Table 2 Error correction in data blocks

	Localization of the error									
	twice error						once error			
A_1	1*	1*	1*	0	0	0	1*	0	0	0
A_2	1*	0	0	1*	1*	0	0	1*	0	0
A_3	0	1*	0	1*	0	1*	0	0	1*	0
A_4	0	0	1*	0	1*	1*	0	0	0	1*
$F_1(A_3+A_4)$	0	1	1	1	1	0	0	0	1	1
$F_2(A_3+A_4)$	0	1	1	1	1	0	0	0	1	1
$F_3(A_2+A_4)$	1	0	1	1	0	1	0	1	0	1
$F_4(A_1+A_4)$	1	1	0	0	1	1	1	0	0	1
$F_5(A_2+A_3)$	1	1	0	0	1	1	0	1	1	0

* It is the falsified block of information

Let us consider the results of the detection and correction of errors with the given Hamming codes, depending on the multiplicity of error (Table 2).

The value of the syndrome for different combinations of errors is shown in Table 2. Matrix models of encoding and error detection devices for any Hamming code are constructed by analogy.

5 Conclusion

The methods and means of building a secure dynamic communication network in a cooperative production system in an open functional space are considered. We proposed to use cryptographic method of password authentication and identification to enhance the security of communication channels in situational communication clusters.

It is suggested to use hashing technologies to improve the message speed processing, detection of the fact and the place of damage in the sent information packet.

For the first time, the method is proposed that provides a guaranteed detection of a predetermined number of falsified fragments in a message by entering control information blocks and performing cross-hashing.

The using of hashing technology provides both rapid authentication in a multi-agent environment and the ability to quickly analyze the structure and content of sent messages.

The developed method can be one of the steps towards achieving an adequate level of information reliability in the communication system and increasing confidence in the cooperation or collaboration of people and robots.

References

1. Xu T, Zhang T, Kühnlenz K, Buss M (2008) Towards high-speed vision for attention and navigation of autonomous city explorer (ACE). In: Xiong Z (ed) Computer vision. In-teh, Rijeka, Croatia, pp 189–214
2. Virágh C, Vásárhelyi G, Tarcai N, Szörényi T, Somorjai G, Nepusz T, Vicsek T (2014) Flocking algorithm for autonomous flying robots. Bioinspiration Biomim 9(2)
3. Kolbeinsson A, Lagerstedt E, Lindblom J (2019) Foundation for a classification of collaboration levels for human-robot cooperation in manufacturing. Prod Manuf Res 7(1):448–471
4. SUCRé. Reliability and resilience for the cooperative control and management of sociotechnical systems: Human(s)-Robot(s) cooperation in hostile environment. LAMIH. https://www.uphf.fr/LAMIH/en/reliability-and-resilience-cooperative-control-and-management-sociotechnical-systems-humans-robots. Accessed 2 Feb 2021
5. Treffler R. Innovation Award 2018: and the winner is… https://www.blog.kuka.com/2018/05/10/kuka-innovation-award-2018-and-the-winner-is/?lang=en. Accessed 2 Feb 2021
6. Goetz U. As soon as they have to start thinking about safety fencing, most of them opt out. https://www.blog.kuka.com/2018/09/07/collaboration-and-coexistence/?lang=en. Accessed 2 Feb 2021
7. Velásquez I, Caro A, Rodríguez A (2018) Authentication schemes and methods: a systematic literature review. Inf Softw Technol 94:30–37
8. Song R (2010) Advanced smart card based password authentication protocol. Comput Stand Interfaces 32(5–6):321–325
9. Fazlida MR, Said J (2015) Information security: risk, governance and implementation setback. Procedia Econ Finance 28:243–248
10. Siponen M, Oinas-Kukkonen H (2007) A review of information security issues and respective research contributions. DATA BASE Adv Inf Syst 38(1):60–80
11. Cao Z, Yin Z, Hu H, Gao X, Wang L (2016) High capacity data hiding scheme based on (7, 4) Hamming code. Springerplus 5:175
12. Ma Z, Li F, Zhang X (2013) Data hiding in halftone images based on hamming code and slave pixels. J Shanghai Univ (Nat Sci) 19(2):111–115
13. Wang K, Shao Q (2011) Analysis of cloud computing and information security. In: Proceedings of the 2nd international conference on frontiers of manufacturing and design science (ICFMD'11), Taichung, Taiwan, pp 3810–3813
14. Rozlomii IA, Rudnitsky VN, Alekseeva ES (2017) Using of hash function to identify counterfit fragments of electronic document. Wschodnioeuropejskie Czasopismo Naukowe (East Eur Sci J) 3(19):68–72
15. Sahu A, Ghosh SM (2017) Review paper on secure Hash algorithm with its variants. Int J Techn Innov Mod Eng Sci (IJTIMES) 3(05)

Generation of Pseudo-random Sequences of the Maximum Period Using Elliptic Curves Transformations

Alexandr Kuznetsov, **Anastasiia Kiian**, **Yevgen Kotukh**,
Serhii Florov, and **Tetiana Kuznetsova**

Abstract A promising direction for constructing cryptographically stable pseudo-random sequence generators is an application of transformations in a group of points of elliptic and hypereliptic curves. This will allow building evidence-stable crypto algorithms, the problem of finding the private key in which is associated with solving a theoretically complex elliptic curve discrete logarithm problem. This paper proposes a method for generating pseudo-random sequences of the maximal period using transformations on elliptic curves. This method consists in the application of recurrent transformations with sequential formation of elements of points group of elliptic curves. This allows providing the maximum period of pseudo-random sequences with the reduction of the problem of finding the private key to the solution of the theoretically complex elliptic curve discrete logarithm problem. The block diagram of the device for generating pseudo-random sequences and the scheme for generating the internal states of the generator are given. We also present the results of statistical testing of some generators, which show that the generated sequences are indistinguishable in a statistical sense from truly random ones.

Keywords Pseudo-random sequences · Elliptic curves · Maximum period

A. Kuznetsov (✉) · A. Kiian · T. Kuznetsova
V. N. Karazin, Kharkiv National University, Svobody sq., 4, Kharkiv 61022, Ukraine
e-mail: kuznetsov@karazin.ua

A. Kuznetsov
JSC "Institute of Information Technologies", Bakulin St., 12, Kharkiv 61166, Ukraine

Y. Kotukh
University of Customs and Finance, Vladimir Vernadsky St., 2/4, Dnipro 49000, Ukraine

S. Florov
Dnipro University of Technology, av. Dmytra Yavornytskoho, 19, Dnipro 49005, Ukraine
e-mail: florov.s.v@nmu.one

1 Introduction

The analysis and comparative studies have shown that the most effective [1–3], in terms of indistinguishability of generated sequences with the implementation of a random process, are methods of generating pseudo-random numbers based on the use of modular transformations [4–6] or transformations in a group of elliptic curve points [1, 7, 8]. The most promising are considered to be pseudo-random sequence generators [9–11], which are constructed using transformations in a group of points of an elliptic curve [8–14].

At the same time, as studies have shown [15, 16], the main disadvantage of such methods is that they do not allow the formation of pseudo-random sequences of the maximum period. For example, when using the well-known Dual Elliptic Curve Deterministic Random Bit Generator with NIST Special Publication 800-90A [17] (in the updated version of the standard this algorithm was excluded [18]), the actual length of the sequence period is much shorter than expected [15]. As the lengths of the parameters increase, this tendency intensifies [15]. Indeed, the application of the operation of scalar multiplication of points of an elliptic curve and the display of the coordinates of the obtained point for the formation of pseudo-random numbers does not provide the maximum period of the formed sequences. In this paper the task is to develop a method of forming sequences of pseudo-random numbers which, due to additional introduction of recurrent transformation in combination with transformations in group of points of an elliptic curve will allow to form pseudo-random sequences of the maximum period.

2 Dual Elliptic Curve Deterministic Random Bit Generator

Dual Elliptic Curve Deterministic Random Bit Generator is based on the application of two scalar multiplications of the points of the elliptic curve and the mapping of the corresponding x-coordinates of the results to a non-zero integer value [17]. The block diagram of the generator is shown in Fig. 1.

The first scalar multiplication by a fixed (base) point P is performed to form an intermediate state s_i, which is cyclically updated at each iteration during the operation of the corresponding generator. Thus, the value of the state s_i depends on the value of the previous state s_{i-1} (on the previous iteration) and on the value of the base point P:

$$s_i = \varphi(x(s_{i-1}P)), \tag{1}$$

where $x(A)$ is a x-coordinate of the point A, $\varphi(x)$ is a function of mapping field elements to non-zero integers.

The initial value of the parameter s_0 is formed using the initialization procedure, which includes entering a private key (Key), which specifies the initial entropy
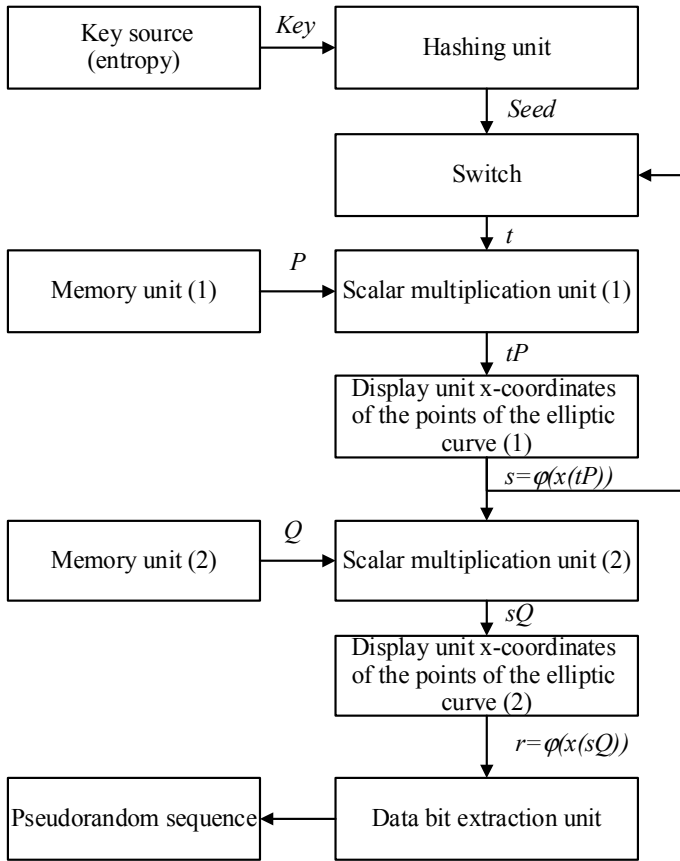
Fig. 1 Block diagram dual elliptic curve deterministic random bit generator

(uncertainty), and hashing the entered key with the formatting of the result to a certain bit length. The value *Seed* obtained in this way sows (initiates) the initial value of the parameter: $s_0 = Seed$.

The second scalar multiplication by a fixed (base) point Q is performed to generate an intermediate state r_i, which after the appropriate transformation and sets the value of the generated pseudo-random bits. The value of the parameter r_i depends on the parameter formed as a result of the first scalar multiplication s_i and on the value of the base point Q:

$$r_i = \varphi(x(s_i Q)). \tag{2}$$

The obtained value r_i is the source for the formation of pseudo-random bits, which are generated by reading a block of the least significant (right) bits of the number r_i. The pseudo-random sequence is formed by concatenation of the read bits of the

formed numbers r_i. The values of fixed (base) points are set as constants and do not change during the formation of the pseudo-random sequence.

Thus, the considered method of generation of pseudo-random sequences applies the transformation in the group of points of the elliptic curve to the formation of intermediate states s_i and r_i. Moreover, the reverse action, ie the formation of s_{i-1} by the known s_i, and/or the formation of s_i by the known r_i is associated with the solution of a theoretically complex elliptic curve discrete logarithm problem.

3   New Pseudo-random Sequence Generator on Elliptic Curves

The generation of sequences of the maximum period is solved by the additional introduction of recurrent transformations. The block diagram of the new generator is shown in Fig. 2. The basis of our proposal is the above Dual Elliptic Curve Deter-
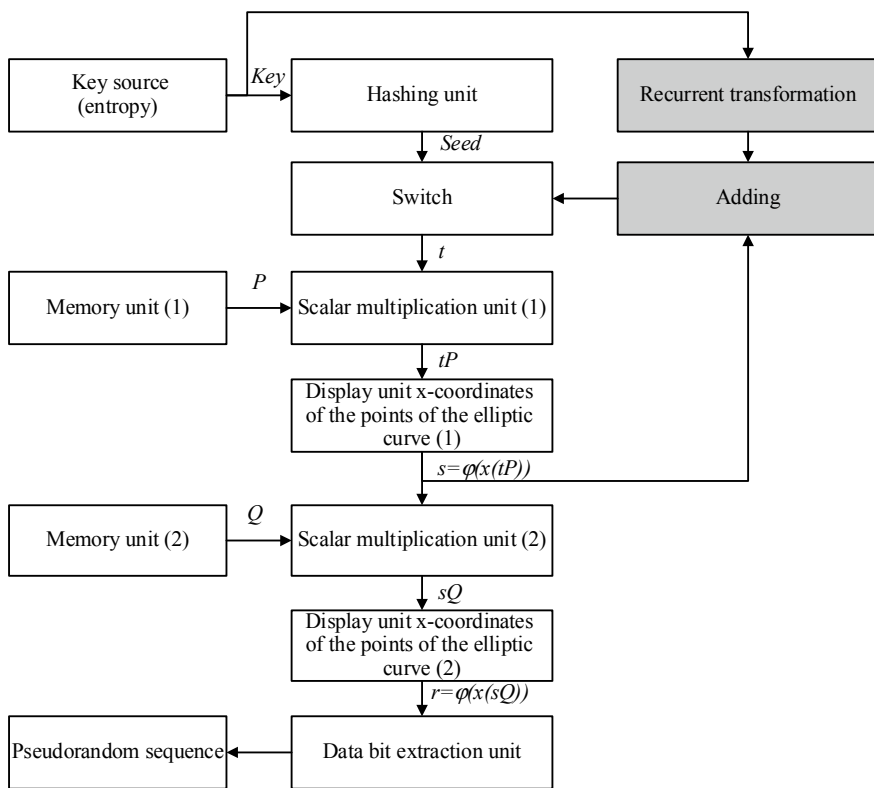
Fig. 2 Block diagram of the new generator

ministic Random Bit Generator (from the previous version of the standard NIST SP 800-90), the newly introduced elements are highlighted in Fig. 2.

The first scalar multiplication to a fixed (base) point, P as in the Dual Elliptic Curve Deterministic Random Bit Generator, is performed to generate an intermediate state s_i, which is cyclically updated at each iteration during the operation of the corresponding generator. But the fundamental difference is the process of formation of this intermediate state. To ensure the maximum period of the sequences …s_{i-1}, s_i, s_{i+1}… in our method it is proposed to use a recurrent transformation, which is initiated by the entered secret key (Key).

Thus, each subsequent value of the state depends not only on the value of the previous state s_{i-1} (on the previous iteration) and on the value of the base point P, but also on the result of the recurrent transformation (denote it by $LRR(y)$), i.e.:

$$s_i = \varphi(x((s_{i-1} + LRR(y))P)),$$

where $x(A)$ is the x-coordinate of the point A, $\varphi(x)$ is a function of mapping field elements to non-zero integers.

The recurrent transformation can be constructed in different known ways, in particular, through the simplest variant using a chain of linear recurrent registers (LRR) with feedback [19], the taps of which are given by the polynomial coefficients

$$g(x) = g_0 + g_1 x + g_2 x^2 + \cdots + b_m x^m.$$

If the polynomial $g(x)$ is primitive over a finite field $GF(2^m)$, then the sequence formed by the LRR with the corresponding feedback logic has a maximum period equal to $2^m - 1$.

The value of the private key (Key), which initiates the work $LRR(y)$, is written in the LRR as the initial value of the register.

The initial value of the parameter s_0, as in Dual Elliptic Curve Deterministic Random Bit Generator, is formed using the initialization procedure, which includes entering a private key (Key), which sets the initial entropy (uncertainty), and hashing the entered key with formatting the result to a certain bit length. The value obtained in this way sows (initiates) the initial value of the parameter: $s_0 = Seed$.

The second scalar multiplication by a fixed (base) point Q is performed to form an intermediate state r_i, which after the appropriate transformation and sets the value of the generated pseudo-random bits. Since each subsequent state value s_i depends on the result of the recurrent transformation $LRR(y)$, which provides the maximum period of the formed sequences, then the value of the parameter r_i depends on the parameter s_i and the value of the base point Q: $r_i = \varphi(x(s_i Q))$ will depend on the result of the recurrent transformation $LRR(y)$, ie the formed sequence of states …r_{i-1}, r_i, r_{i+1},… will have a maximum period.

The obtained value r_i is a source for the formation of pseudo-random bits, which are formed by reading the block from the least significant (right) bits of the number

r_i. The pseudo-random sequence is formed by concatenation of the bits of the formed numbers r_i.

The values of fixed (base) points are set as constants and they do not change during the formation of the pseudo-random sequence. Thus, the periodic properties of the states of the proposed generator are determined by the periodic properties of the additionally introduced recurrent transformation $LRR(y)$.

Let's denote the original sequence of transformation $LRR(y)$ by $\ldots y_{i-1}$, y_i, $\ldots y_{i+1}\ldots$ and schematically show the influence of the periodicity of this sequence on the periodicity of the sequences $\ldots s_{i-1}$, s_i, $\ldots s_{i+1}\ldots$ and $\ldots r_{i-1}$, r_i, r_{i+1},$\ldots$.

Because the periodic properties of the sequences $\ldots s_{i-1}$, s_i, $\ldots s_{i+1}\ldots$ and $\ldots r_{i-1}$, r_i, r_{i+1},$\ldots$ directly depend on the properties of the sequence $\ldots y_{i-1}$, y_i, $\ldots y_{i+1}\ldots$, the use of recurrent transformation $LRR(y)$ with the maximum period of the original sequences provides the maximum period of the original sequence.

4   Statistical Testing of Pseudo-random Sequences

Various statistical tests are used to experimentally test the efficiency of generators of random and pseudo-random numbers. The most common NIST STS statistical tests are used to evaluate the performance of generators in cryptographic applications [20].

4.1   NIST STS Statistical Tests

NIST Statistical Tests is a statistical test suite developed by the Information Technology Laboratory of the National Institute of Standards and Technology (NIST). This package includes 15 statistical tests:

- Frequency (Monobit) Test;
- Frequency Test within a Block;
- Runs Test;
- Test for the Longest Run of Ones in a Block;
- Binary Matrix Rank Test;
- Discrete Fourier Transform (Spectral) Test;
- Non-overlapping Template Matching Test;
- Overlapping Template Matching Test;
- Maurer's "Universal Statistical" Test;
- Linear Complexity Test;
- Serial Test;
- Approximate Entropy Test;
- Cumulative Sums (Cusum) Test;
- Random Excursions Test;
- Random Excursions Variant Test.

The purpose of each test is to determine the measure of randomness of binary sequences generated by either hardware or software random number generators.

NIST statistical tests are based on various statistical properties that are unique to random sequences.

Knowing the probabilistic properties of a truly random sequence, one can use them to test the hypothesis about how similar the generated sequence is to a random one. For this, a suitable statistic is selected for each test, and its values are calculated for the ideal and generated sequence. If the difference between these values exceeds a certain critical value set in advance, then the sequence is considered non-random.

Let us define the p-value as the probability that the ideal generator generated a sequence less random than the one under study. Then, if the p-value is greater than α, then the sequence under study is considered random and vice versa, otherwise.

In total (taking into account various parameters for 15 tests), statistical studies are performed using 188 tests. In this case, the value of the probability p calculated for each test must be at least $\alpha = 0.01$. Otherwise ($p < \alpha$), the binary sequence is not completely random.

Thus, to test the generator of random and pseudo-random numbers, a long output sequence is formed (it is recommended to use 10^8 bits). And for this long sequence, 188 p-values are calculated. These 188 values make it possible to evaluate the similarity of the generator output sequence to a random sequence.

4.2 Test Results

We have performed statistical testing on several pseudo-random number generators. Specifically, we tested:

- generator on elliptic curves;
- Micali Shnorr generator;
- ANSI X9.17 generator (3-DES);
- generator using SHA-1.

We presented the test results (188 p-values) in the form of diagrams in Figs. 3, 4, 5, and 6.

The ordinate scale shows the test numbers (from 1 to 188). The p-value corresponding to these 188 tests is plotted on the abscissa scale. Thus, each diagram shows the full spectrum of 188 values of the probabilities p, which characterize how similar the generated sequence is to a random one.

In each figure, the range of values $p < \alpha$ is highlighted in red. A hit in this area indicates that the generated binary sequence is not completely random. In other words, if at least one of 188 p-values is in the area $p < \alpha$, then the corresponding generator has drawbacks, the sequences it generates differ from truly random ones.

The test results showed very good statistical properties. All the generators under study actually form pseudo-random sequences. For all 188 statistical tests, the obtained values are $p > \alpha = 0.01$. This means that for each test, the long binary
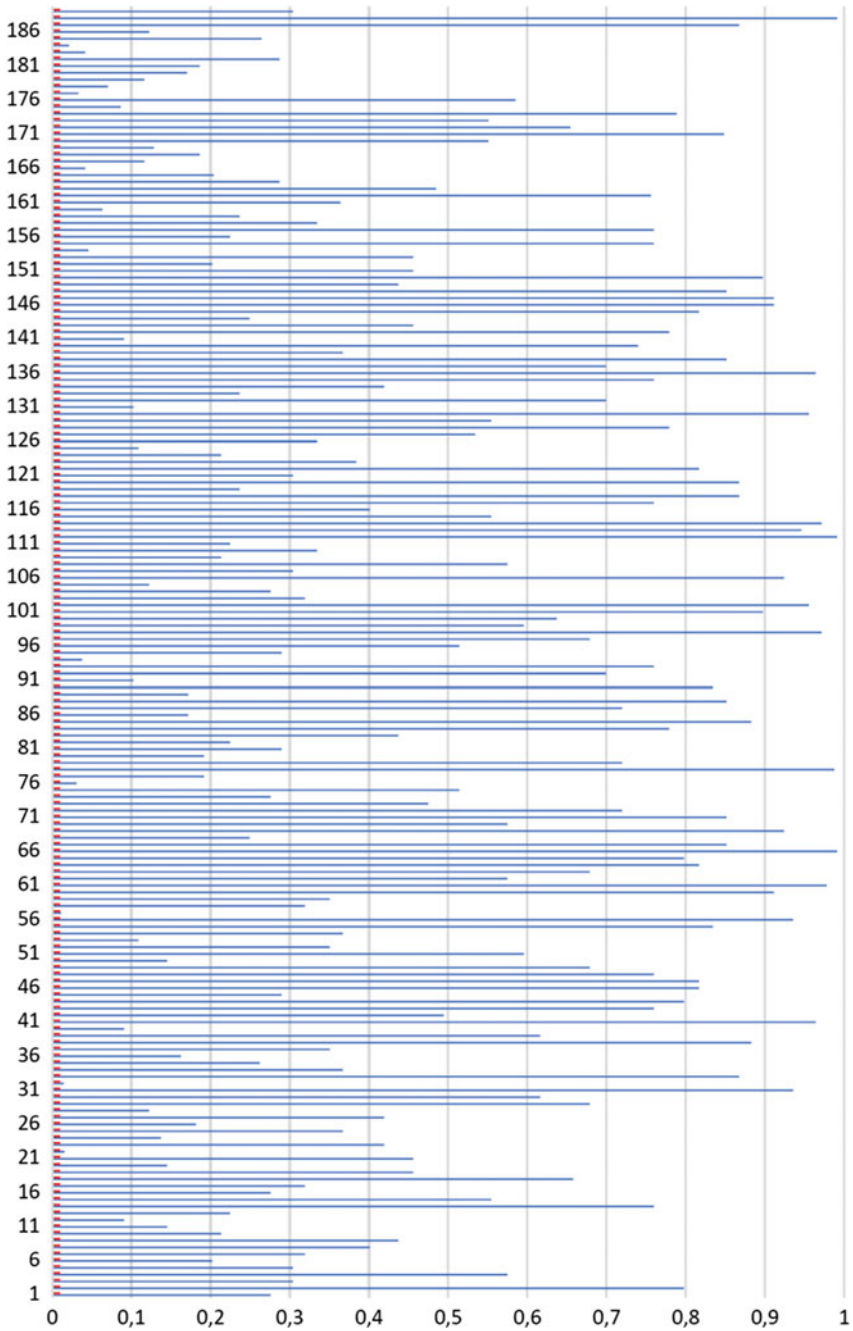
Fig. 3 Generator on elliptic curves
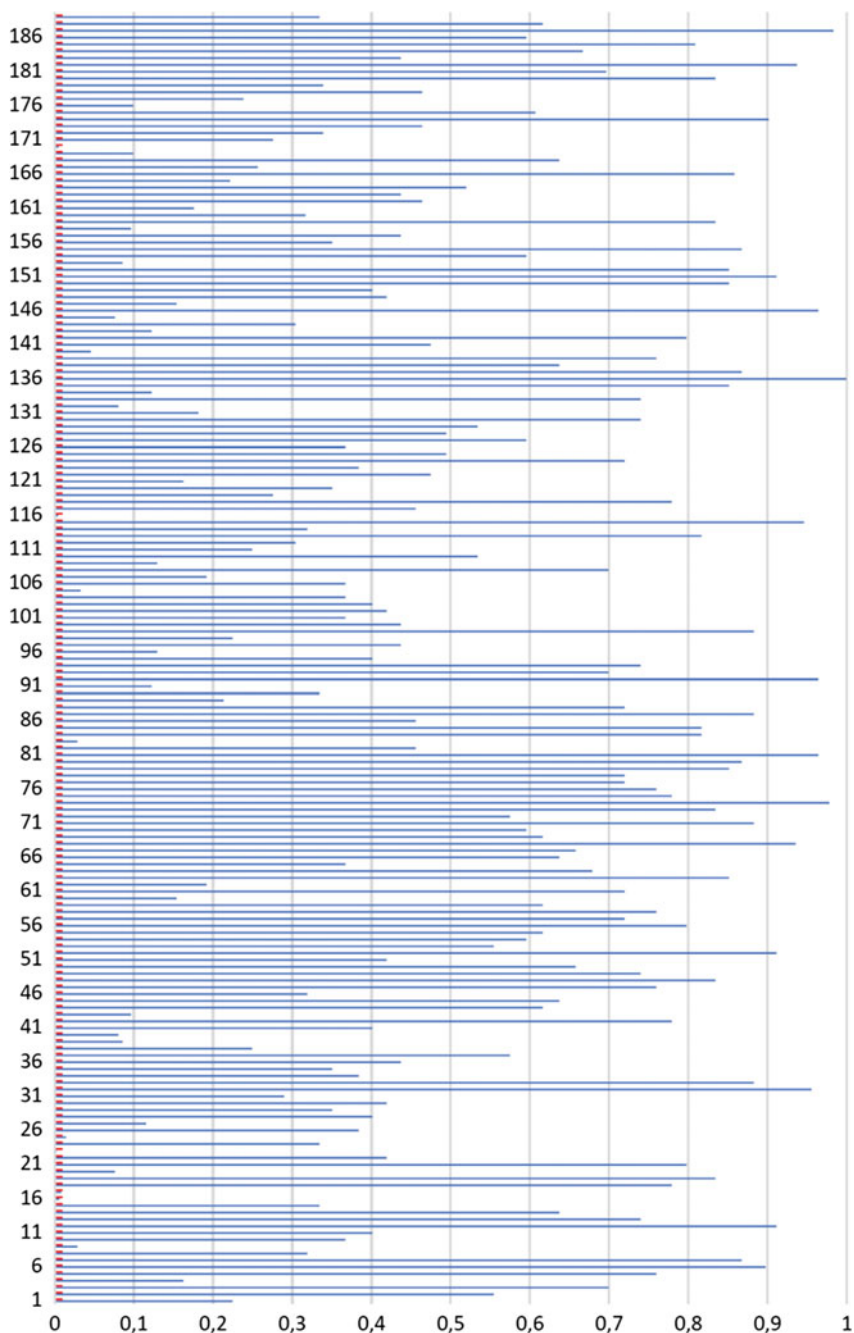
Fig. 4 Micali Shnorr generator
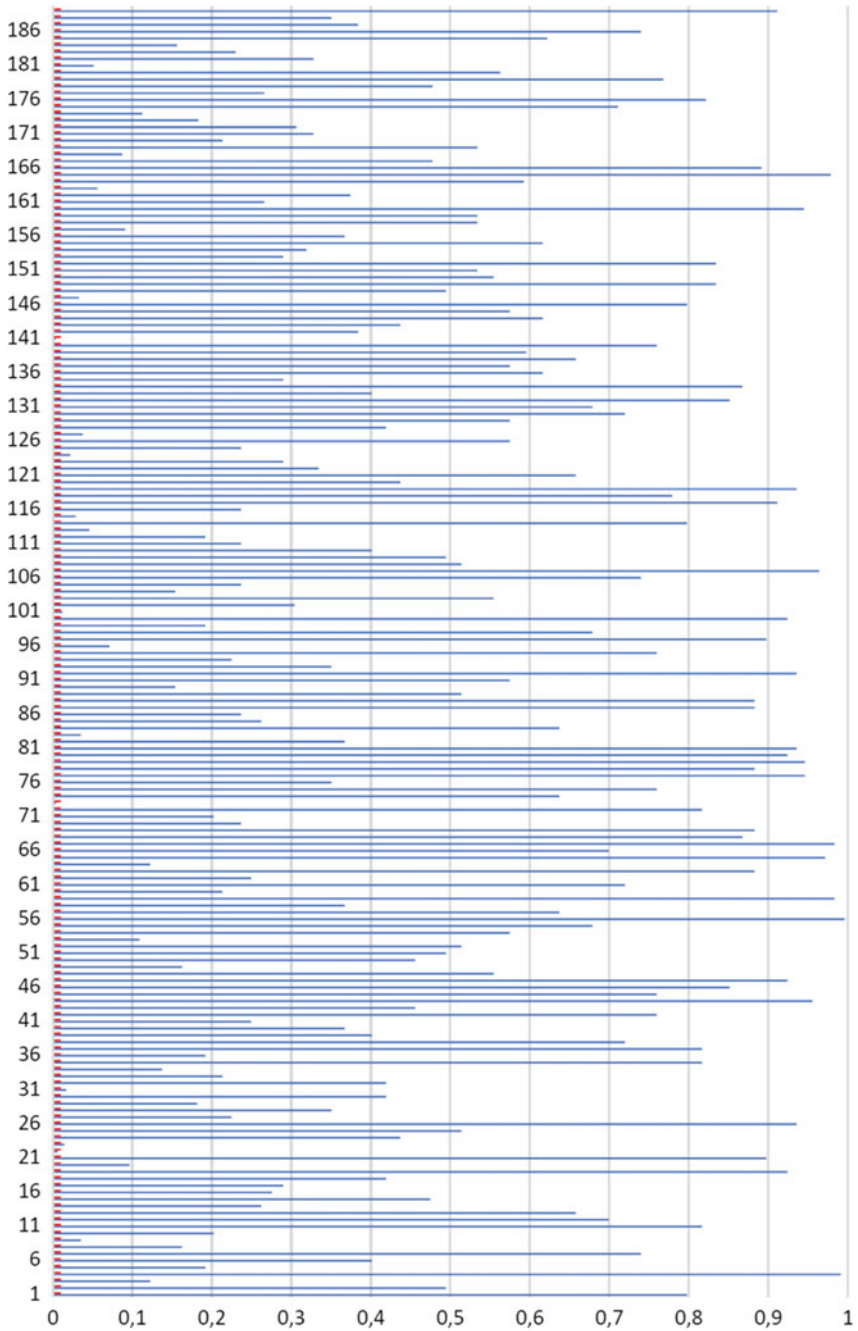
Fig. 5 ANSI X9.17 generator (3-DES)
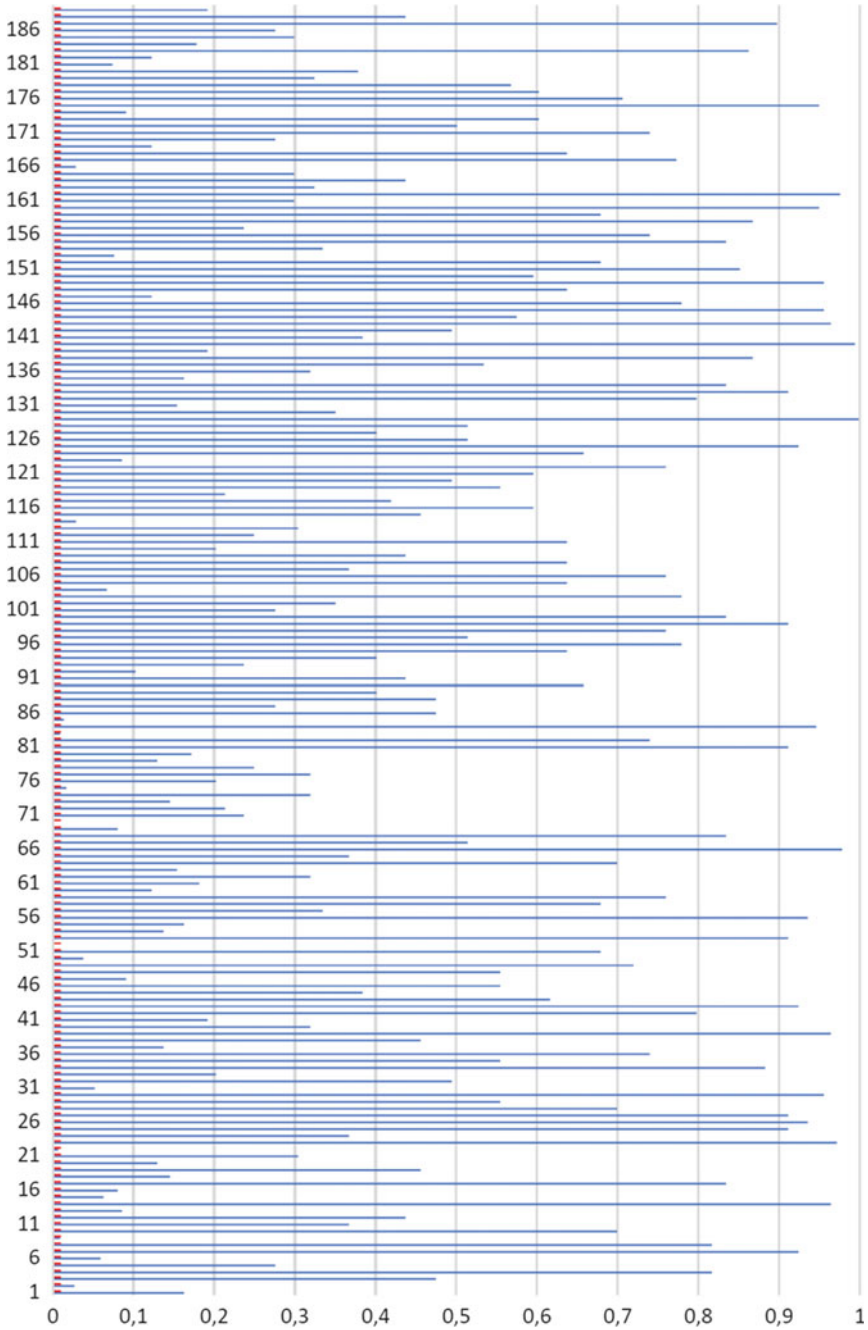
Fig. 6 Generator using SHA-1

sequences generated by the generators do not differ in a statistical sense from truly random ones.

5   Conclusions

The application of cryptographic transformations in a group of elliptic curves points allows to build efficient pseudo-random sequences generators. However, for the known Dual Elliptic Curve Deterministic Random Bit Generator (which is described in the previous version in the standard NIST SP 800-90) there are some drawbacks. In particular, the cyclic function of the generator does not provide the maximum period of the formed sequence of internal states and the corresponding points of the elliptic curves. There is an early looping, as a result, the real period of pseudo-random sequences is much smaller than expected. This paper proposes a new method that additionally uses recurrent transformations (implemented, for example, using linear recurrent registers with feedback). This allows you to generate sequences of states of the generator with the maximum period. The formed sequences of elliptic curve points also have a maximum period. Therefore, the new method removes certain shortcomings of the Dual Elliptic Curve Deterministic Random Bit Generator regarding the periodic properties of the generated pseudo-random sequences.

We also performed statistical testing of some generators. For this, we used NIST statistical tests. The results obtained convincingly show the indistinguishability of the generated sequences from the truly random ones, i.e. these generators can be used in various cryptographic applications.

References

1. Menezes AJ., van Oorschot, P.C., Vanstone, S.A., van Oorschot, P.C., Vanstone, S.A.: Handbook of Applied Cryptography. CRC Press (2018). https://doi.org/10.1201/9780429466335
2. Delfs, H., Knebl, H.: Introduction to Cryptography. Springer, Berlin, Heidelberg (2015). https://doi.org/10.1007/978-3-662-47974-2.
3. Rubinstein-Salzedo, S.: Cryptography. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94818-8.
4. Blum L, Blum M, Shub M (1986) A simple unpredictable pseudo-random number generator. SIAM J Comput 15:364–383. https://doi.org/10.1137/0215025
5. Blum, M., Micali, S.: How to generate cryptographically strong sequences of pseudo random bits. In: 23rd Annual Symposium on Foundations of Computer Science (SFCS 1982). pp. 112–117 (1982). https://doi.org/10.1109/SFCS.1982.72.
6. Blum M, Micali S (1984) How to generate cryptographically strong sequences of pseudo-random bits. SIAM J Comput 13:850–864. https://doi.org/10.1137/0213053
7. Shamir, A.: On the generation of cryptographically strong pseudo-random sequences. In: Even, S. and Kariv, O. (eds.) Automata, Languages and Programming. pp. 544–550. Springer, Berlin, Heidelberg (1981). https://doi.org/10.1007/3-540-10843-2_43.

8. Steinmetz, R., Dittmann, J., Steinebach, M. eds: Communications and Multimedia Security Issues of the New Century, Исходный текст: IFIP TC6/TC11 Fifth Joint Working Conference on Communications and Multimedia Security (CMS'01) May 21–22, 2001, Darmstadt, Germany. Springer US (2001). https://doi.org/10.1007/978-0-387-35413-2.

9. Reyad, O., Karar, M.E., Hamed, K.: Random Bit Generator Mechanism Based on Elliptic Curves and Secure Hash Function. arXiv:2002.09239 [cs]. (2020).

10. Payingat, J., Pattathil, D.P.: Pseudorandom Bit Sequence Generator for Stream Cipher Based on Elliptic Curves, https://www.hindawi.com/journals/mpe/2015/257904/, last accessed 2021/01/16. https://doi.org/10.1155/2015/257904.

11. Benssalah, M., Djeddou, M., Drouiche, K.: Pseudo-random sequence generator based on random selection of an elliptic curve. In: 2015 International Conference on Computer, Information and Telecommunication Systems (CITS). pp. 1–5 (2015). https://doi.org/10.1109/CITS.2015.7297719.

12. Brown, D.R.L., Vanstone, S.A.: Elliptic curve random number generation, https://patents.google.com/patent/US8396213B2/en, (2013).

13. Chevardin, V.: Deterministic random bit generator on elliptic curve transformations. In: Proceedings of International Conference on Modern Problem of Radio Engineering, Telecommunications and Computer Science. pp. 468–468 (2012).

14. Lee L-P, Wong K-W (2004) A random number generator based on elliptic curve operations. Comput Math Appl 47:217–226. https://doi.org/10.1016/S0898-1221(04)90018-1

15. Kuznetsov, A., Kavun, S., Panchenko, V., Prokopovych-Tkachenko, D., Kurinniy, F., Shoiko, V.: Periodic Properties of Cryptographically Strong Pseudorandom Sequences. In: 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S T). pp. 129–134 (2018). https://doi.org/10.1109/INFOCOMMST.2018.8632021.

16. Kuznetsov, A., Kiian, A., Smirnov, O., Cherep, A., Kanabekova, M., Chepurko, I.: Testing of Code-Based Pseudorandom Number Generators for Post-Quantum Application. In: 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT). pp. 172–177 (2020). https://doi.org/10.1109/DESSERT50317.2020.9125045.

17. Barker E, Kelsey J (2012) Recommendation for Random Number Generation Using Deterministic Random Bit Generators. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-90A

18. Barker E, Kelsey J (2015) Recommendation for random number generation using deterministic random bit generators. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-90Ar1

19. Canteaut, A.: Linear Feedback Shift Register. In: van Tilborg, H.C.A. and Jajodia, S. (eds.) Encyclopedia of Cryptography and Security. pp. 726–729. Springer US, Boston, MA (2011). https://doi.org/10.1007/978-1-4419-5906-5_357

20. Computer Security Division, I.T.L.: NIST SP 800–22: Documentation and Software - Random Bit Generation | CSRC | CSRC, https://content.csrc.e1c.nist.gov/projects/random-bit-generation/documentation-and-software, Accessed 2021/03/12

Studying the Dynamic Bottlenecks of a Load Balancer in Distributed Systems

Oleksandr Khoshaba ⓘ**, Viktor Grechaninov** ⓘ**, Anatoliy Lopushanskyi** ⓘ**, and Kostiantyn Zavertailo** ⓘ

Abstract This article describes a research system in a distributed network that consists of structural and functional components. Also, the paper proposes (describes) deterministic and probabilistic methods for determining the characteristics of objects in the research system: benchmark, pool, modules. The main focus is on deterministic and probabilistic methods for studying bottlenecks' dynamics in a load balancer. Deterministic methods for studying the dynamics of bottlenecks in a load balancer are considered examples of synchronous and asynchronous data exchange models. The use of probabilistic models to investigate the load balancer's dynamic bottlenecks consisted of Turki's method, which calculated percentiles and determining deviations from specified boundaries. A model of a vector representation of dynamic data exchange between a benchmark and a load balancer pool is proposed.

Keywords Dynamic bottlenecks · Benchmarks · Load balancers · Dynamic data exchange · Performance thresholds · Nodes of distributed systems · Deterministic and probabilistic methods · Turki method · Structural and functional components of the research system

1 Introduction

Cloud infrastructures and distributed systems are designed to provide high availability, scalability, and reliability by combining nodes (workstations, servers, and network equipment) into groups. Therefore, user and equipment requests should be evenly distributed across each node in the group to increase their performance [1–4].

In this regard, several areas are developing aimed at solving many problems. In particular, an adaptive and efficient load balancing algorithm for corporate server nodes are offered [5–8].

O. Khoshaba (✉) · V. Grechaninov · A. Lopushanskyi · K. Zavertailo
Institute of Mathematical Machines and Systems Problems of the Ukraine National,
Academy of Science, 42 Academician Glushkov Avenue, Kyiv 03187, Ukraine
URL: http://khoshaba.vk.vntu.edu.ua/

Besides, more and more companies are moving their data management applications from expensive high-performance servers to the cloud. Therefore, a cloud-based data management system becomes relevant due to scalability, fault tolerance, and the effective solution of large-scale data management tasks.

Based on this, the implementation of existing cloud-based data management systems represents a wide range of approaches, including storage architectures, data models, tradeoffs inconsistency, and availability.

Therefore, to assess the performance of computing and network systems, many methods and standards are proposed [9–16]. Nevertheless, studies of such results remain relevant that allow users to understand the impact of various approaches to implementing different methods on the performance of computing and network systems.

2　Purpose of Work

The aim of this work is:

- the study of the performance thresholds of load balancer bottlenecks in distributed systems;
- the creation of a research structure for a load balancer in distributed systems;
- the development of structural and functional components of the research system;
- the use of deterministic and probabilistic methods for determining the characteristics of the studied objects in the system: benchmark, pool, and modules;
- the creation of a model of synchronous and asynchronous data exchange based on deterministic methods for studying the dynamics of bottlenecks in the load balancer;
- determination of bottlenecks of the load balancer by the Turki method [17];
- investigation of dynamic data exchange between benchmark and load balancer pool.

3　Problem Formulation

The problem formulation considered in this article is focused on studying the dynamics of bottlenecks that appear in the load balancer in distributed systems. For this, several stages of the life cycle are created: system concept, design and development, operation, service and support, decommissioning, and disposal.

This life cycle is used when building a research system. Therefore, the main purpose of building a research system is to study the functioning of a load balancer in a distributed system.

Also, the study of bottlenecks that appear in the load balancer should cover such localizations in the research system as:

- from the benchmark to the pool of the load balancer;
- from the pool to the load balancer modules;
- from load balancer modules to distributed system nodes.

4 Structural and Functional Components of the Research System

4.1 Structural Components of the Research System

The research system's structural components include such basic parts as a benchmark, a pool, and load balancer modules, a monitoring system, and nodes of a distributed system (Fig. 1).

Let's consider the main structural components of the research system.

A benchmark is a structural component of a research system required to generate streams of requests that are executed according to a specific user scenario.

A module is a structural component of a load balancer research framework. Its main functions include:

- processing and exchange of synchronous data streams (requests);
- synchronization of requests between the pool and load balancer;
- synchronization of requests between the load balancer and nodes.
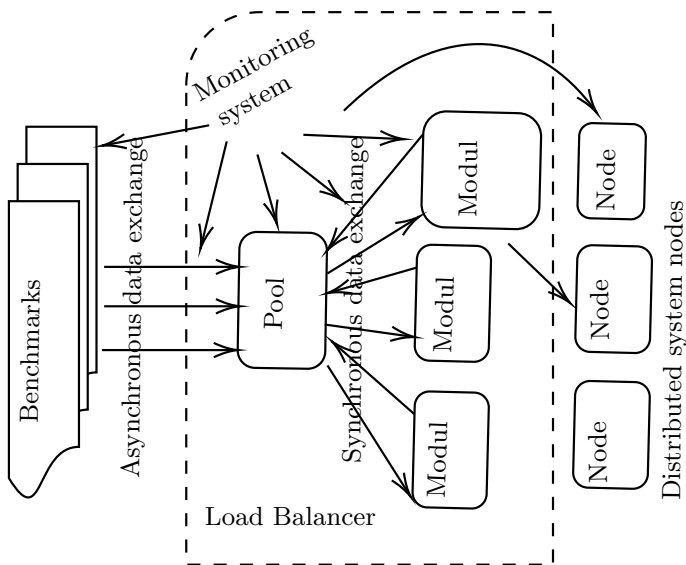
Fig. 1 Structural components of the research system

A pool is a structural component of a load balancer research framework that is required to:

- storage of request streams;
- synchronization of request streams between the benchmark and the load balancer.

In this research system, the pool is a dynamic bottleneck in the load balancer.

Distributed system nodes are a structural component of a research system. Their main function is to implement request streams that are generated by a benchmark using a script. Streams of requests are transmitted to the nodes of the research system using modules.

A monitoring system in a research system is a structural component of a load balancer, which is necessary to measure a benchmark, load balancer pool, modules, and communication channels.

Watchpoints are used to identify load balancer bottlenecks in the monitoring system. The observation points of the monitoring system of the research system can be set at a certain time interval of loading effects or the occurrence of an event.

Also, the structural components include communication channels that play an important role in the research system's work. Communication channels are the structural formations required to pass requests to the main components of the load balancer. In communication channels, synchronous and asynchronous request streams can be transmitted.

4.2 Functional Components of the Research System

The research system's functional components consist of request streams, time intervals between requests in communication channels, mechanisms for managing computing and information resources, a user scenario for load impact on research objects and their states. Let's consider some of them.

Request streams are functional components of the load balancer research framework that can be synchronous or asynchronous.

The time interval between requests in the communication channel is a functional component of the research system based on determining the time period between requests. The time interval in the communication channel can take on some values (4, 3).

Computational and information resource management mechanisms are a functional component of the research system that defines methods, models, and states of resource management for the research system: benchmark, load balancer pool, and modules (Table 1).

For example, there is a load balancer pool management. This pool management is based on a synchronization mechanism, locking access to the benchmark and modules to implement stream requests to remote nodes. In this case, the user or developer

Table 1 Computational and information resource management mechanisms

Characteristics	Objects of study		
	Benchmark	Pool	Module
Models	Deterministic model	Probabilistic model	Probabilistic model
States	Loaded and unloaded	Loaded and unloaded	Loaded and unloaded
Resource manageability	Controlled by a user	Not controlled by a user	Controlled by a user

cannot control the pool management mechanism. This is because the pool synchronization mechanism, written in Java, is controlled by the programming environment: system and application libraries of the virtual machine.

A custom load scenario is a functional component of a research system that describes the sequence in which requests flows to a load balancer are generated. The states of the research objects are also functional components of the research system, which describe the states of the benchmark, the load balancer pool, and the remote modules of the system based on certain criteria. Such criteria for the state of research objects can be obtained by expert advice or experimental research.

Let us define the concepts of the loaded and unloaded state of an object of the research system. Let us assume that the unloaded state of an object of the research system is a stable or steady state in which the research object does not determine the effect of the load on it.

The loaded state of an object of the research system is an unstable state in which the object of research remains at the moment of exposure or after some time has elapsed after exposure to the load.

Determination of the impact of the load on the object of research is based on criteria obtained by expert means (user experience) or experimental data (as a result of using methods).

For example, load states' criteria of the study objects can be such signs as queues, the appearance of parallel flows of requests, the loss of requests, and errors in data transmission in the communication channel.

5 Deterministic Models for Studying the Dynamic Bottleneck of a Load Balancer

The load balancer's performance is based on measurements of its pool using a monitoring system. In doing so, the most important measurements are determining the capacity of the load balancer pool, response time, and intervals between requests from the benchmark to the load balancer. Consider a deterministic model based on response time measurements and intervals between requests in the communication channel between the benchmark and the load balancer pool.

A deterministic model is a model in which the values of the research object's dependent variables are completely determined by the parameters necessary for the research process. In this case, the scenario of loading actions is used. This scenario determines the load at which the benchmark generates a sequence of requests with given characteristics.

These specified characteristics include the following parameters: the response time (RT_i) between requests in the communication channel (1); the time interval (In_i) between requests in the communication channel (2); the total number of requests; the resource address of the remote node of the distributed system.

Building a deterministic model to investigate a load balancer's dynamic bottleneck is based on the following conditions. Let for the current request in the communication channel, the time of its sending will be equal to t_i^b, and the time of its reception will be a t_i^e. Then the response time (RT_i) to the current request in the communication channel will be:

$$RT_i = t_i^e - t_i^b \tag{1}$$

The time interval (In_i) between the current and the next request in the communication channel will be equal to:

$$In_i = t_{i+1}^b - t_i^e \tag{2}$$

Since there can be one or several request streams in the communication channel, the values of the time interval (In_i) can take the following values:

$$In_i = 0; \; In_i > 0 \tag{3}$$

$$In_i < 0 \tag{4}$$

Based on the relationships (3, 4), all requests (5) that are generated by the benchmark in the communication channel can be synchronous (3) and asynchronous (4):

$$In_{ALL} = \{In_1, In_2, In_3, \ldots In_i, \ldots\} \tag{5}$$

According to properties (3) and (4), all queries between which there are intervals can be divided into sets:

$$In_{SYN} = \{In_{s1}, In_{s2}, In_{s3}, \ldots In_{sj}, \ldots\} \tag{6}$$

$$In_{ASN} = \{In_{a1}, In_{a2}, In_{a3}, \ldots In_{aj}, \ldots\} \tag{7}$$

where

$$In_{SYN} \bigcap In_{ASN} = \emptyset \tag{8}$$

and

$$In_{SYN} \bigcup In_{ASN} = In_{ALL} \tag{9}$$

In this regard, we have the following definitions of communication channels.

Definition 1 A synchronous channel (Ch_{SYN}) is a collection of request flows, where between each request all intervals have a property (3):

$$Ch_{SYN} := \forall I_n : In_i \in I_{SYN} \tag{10}$$

or

$$Ch_{SYN} := \forall I_n : In_i \notin I_{ASN} \tag{11}$$

Definition 2 An asynchronous channel (Ch_{ASN}) is a collection of request flows, where between each request there is at least one interval that has property (4):

$$Ch_{ASN} := \exists! I_n : In_i \in I_{ASN} \tag{12}$$

6 Probabilistic Models for Studying Dynamic Bottlenecks of a Load Balancer

A probabilistic model is a model in which the values of the dependent variables of the study object are defined as distributions of random variables. Using this model assumes that various results will be obtained at the output of the study's object that depends on random factors.

As mentioned above, the load impact process is formed by the benchmark and controlled using a user scenario. Therefore, such an impact on the object of study as the pool is deterministic. As a result of such an impact, a response is formed on the research objects: the pool, modules, and remote nodes. The reaction from these objects of research depends on their software and hardware resources and is different. Such a reaction has a probabilistic nature of the distribution of random variables and is studied by the corresponding models.

With insufficient resource capacity and high load impact, anomalies appear in the objects of the study's parameters. The determination of these anomalies is a separate area of knowledge. There are methods of varying degrees of difficulty in this area.

The probabilistic models for studying dynamic bottlenecks of a load balancer consist of applying the Turki [17] method where percentiles are calculated and deviations from the given boundaries are determined. This method is used to determine

anomalous data and determine the boundary states of the study object as follows. As a result of experimental studies, the following data are obtained:

- Median (Q_2/50th Percentile): the middle value of the dataset.
- First quartile (Q_1/25th Percentile): the lower quartile $q_n(0.25)$ is the median of the lower half of the dataset.
- Third quartile (Q_3/75th Percentile): the upper quartile $q_n(0.75)$ is the median of the upper half of the dataset.
- Interquartile Range (IQR): is the distance between the upper and lower quartile. The formula calculates IQR:

$$IQR = Q_3 - Q_1 = q_n(0.75) - q_n(0.25) \tag{13}$$

7 The Model of the Data Exchange Process

Quite often, the load impact on the web resources of remote nodes of distributed systems is performed using user scenarios. Such scenarios contain various addresses (localizations) of web resources, the frequency, and the number of requests to remote nodes of a distributed system. The results of user scenarios are often expressed in the monitoring system by the number of requests sent and received and the amount of data.

For the research system (Fig. 1), such places of exchange will be the benchmark–pool, pool–modules, modules–remote nodes. Thus, studying bottleneck dynamics in a load balancer involves considering pool-level data exchange. Therefore, let's look at the data exchange process between the benchmark and the load balancing pool. In this case, we define the dynamics of processing requests in the pool as receiving and sending data. Then we plot the function of the values of query processing dynamics, where along the X-axis, we denote the sent data and along the Y-axis the collected data.

After that, on the graph (Fig. 2), we indicate a random point A, which corresponds to the dynamic data exchange between the benchmark and the pool of the load

Fig. 2 Vector representation of the dynamic data exchange between the benchmark and the pool of the load balancer
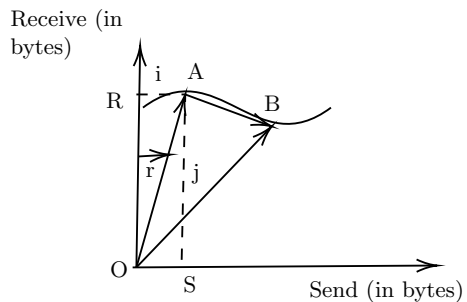
balancer. Also, we introduce the concept of unit vectors, which are shown on the graph as i and j for a random point of A.

For some time, the data exchange values may vary depending on the user scenario, the congestion of the pool and the load balancer module, and the distributed system's remote nodes. Such changes in the input (user scenario) and output (load of the pool and the load balancer module, remote nodes of the distributed system) parameters of the model of the load balancer pool will describe the trajectory of the points $(\overrightarrow{AB}$, Fig. 2) as a function. As a result of this, the unit vector $(\triangle \overrightarrow{r})$ will correspond to:

$$\overrightarrow{AB} = \triangle \overrightarrow{r} = \overrightarrow{r}(t + \triangle t) - \overrightarrow{r}(t) \tag{14}$$

where

$$\overrightarrow{r} = \overrightarrow{r}(t) = S(t) \cdot \overrightarrow{i} + R(t) \cdot \overrightarrow{j} \tag{15}$$

and

$$\overrightarrow{V} = \overrightarrow{V}(t) = \triangle \overrightarrow{r}/t, (\triangle t \to 0) \tag{16}$$

$$\overrightarrow{a} = \overrightarrow{a}(t) = \triangle \overrightarrow{V}/t, (\triangle t \to 0) \tag{17}$$

wherein,

$\overrightarrow{V}(t)$—the rate of the dynamic data exchange between the benchmark and the pool of the load balancer;

$\overrightarrow{a}(t)$—changing the dynamic data exchange rate between the benchmark and the pool of the load balancer.

8 Research Results

These results are the authors' first step in studying a load balancer's dynamic bottlenecks in distributed systems using the above models, deterministic, and probabilistic research methods. In the first step, the study of the interaction of dynamic data exchange between the benchmark and the balancer pool made it possible to identify the following characteristics of the performance thresholds (Table 2).

The formation of the performance thresholds (S0–S4) was as follows. The absence of load impact was attributed to the state of S0. With minimal load impact, the values that did not exceed the values for synchronous communication channels (3) were assigned to state S1. States S2 and S3 were determined based on the calculation of indicators (Q1, Q2, Q3, IQR) using the developed program in the language R:

Table 2 Characteristics of the performance thresholds

Performance thresholds	Characteristics		
	Load impact	Channel	Criterion
S0	No exposure	Synchronous channel	Low node load
S1	Low impact	Synchronous channel	Average node load
S2	Medium impacts	Asynchronous channel	Request queue
S3	High impact	Asynchronous channel	Abnormal requests
S4	Ultra high impact	Asynchronous channel	Errors or loss of requests

```
data = read.csv("filename", header = TRUE)
summary(data)
result.mean <- mean(data$sQu)
result.min <- min(data$sQu)
result.max <- max(data$sQu)
result.median <- median(data$sQu)
result.q1 <- quantile(data$sQu, 0.25)
result.q2 <- quantile(data$sQu, 0.5)
result.q3 <- quantile(data$sQu, 0.75)
sprintf("mean = %5.2f", result.mean)
sprintf("q1 = %5.2f", result.q1)
sprintf("q2 = %5.2f", result.q2)
sprintf("q3 = %5.2f", result.q3)
sprintf(" **** IQR **** = %5.2f",
result.q3 - result.q1)
sprintf("max = %5.2f", result.max)
sprintf(" **** IQR inner fence
(mild outlier) **** = %5.2f",
result.q2 + 1.5 * (result.q3 - result.q1))
sprintf(" **** IQR outer fence
(extreme outlier) **** = %5.2f",
result.q2 + 3 * (result.q3 - result.q1))
sprintf("min = %5.2f", result.min)
sprintf(" **** IQR inner fence
(mild outlier) **** = %5.2f",
result.q2 - 1.5 * (result.q3 - result.q1))
sprintf(" **** IQR outer fence
(extreme outlier) **** = %5.2f",
result.q2 - 3 * (result.q3 - result.q1))
```

As a result of the script, we got the result:

```
. . .                                    [1] " **** IQR **** =   4.09"
RIM modul:10098                          [1] "max = 21.03"
Min.    : 0.211                          [1] " **** IQR inner fence
1st Qu.: 1.967                           (mild outlier) **** = 11.57"
Median : 5.431                           [1] " **** IQR outer fence
Mean    : 4.416                          (extreme outlier) **** = 17.71"
3rd Qu.: 6.059                           [1] "min =   0.21"
Max.    :21.032                          [1] " **** IQR inner fence
[1] "mean =   4.42"                      (mild outlier) **** = −0.71"
[1] "q1 =   1.97"                        [1] " **** IQR outer fence
[1] "q2 =   5.43"                        (extreme outlier) **** = −6.84"
[1] "q3 =   6.06"
```

Cases when, as a result of load impacts, errors appeared during data processing or requests were lost were assigned to state S4.

When determining the S2–S4 states, the localization of the dynamic bottlenecks was determined. As a result of data processing, summary tables, or graphs (Fig. 3) were obtained, which were analyzed for localization and degree of dynamic bottlenecks.
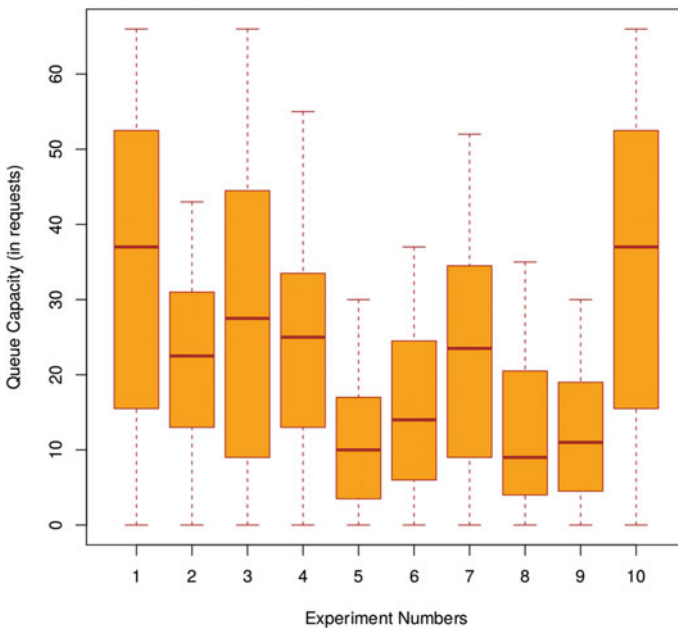
Fig. 3 Bottleneck capacity definition example

Figure 3 shows the results of measurements of bottlenecks' dynamics that appeared on the load balancer pool.

9 Further Research

The next step of further research will be:

- the study of the impact of synthetic benchmark tests on the load balancer in distributed systems;
- the study further statistical analysis of the data;
- the study of the work of the load balancer in an enterprise environment.

References

1. Vladimir Z, Oksyuta O, Dayub N (2020) Load balancing in cloud computing. Model Syst Processes 13:25–32. https://doi.org/10.12737/2219-0767-2020-13-1-25-32
2. Soundarabai P, Venugopal KR, Patnaik L (2014) Situation based load balancer for distributed computing systems
3. Khiyaita A, Zbakh M, El Bakkali H, Kettani D (2012) Load balancing cloud computing: state of art. In: Proceedings of the 2nd national days of network security and systems, JNS2 2012, pp 106–109. https://doi.org/10.1109/JNS2.2012.6249253
4. North D, Vasile A (2015) System and method for load balancing cloud-based accelerated transfer servers
5. Lin WC, Zhang L (2016) The adaptive load balancing algorithm in cloud computing. https://doi.org/10.2991/wartia-16.2016.94
6. Singh Manpreet, Gupta Vishal, Goyal SandipKumar (2011) An adaptive load balancing algorithm for computational grid. J Eng Technol 1:70. https://doi.org/10.4103/0976-8580.86636
7. Park G, Gu B, Heo J, Yi S, Han J, Park J, Min H, Piao X, Cho Y, Park C-W, Chung H, Lee B, Lee S (2006) Adaptive load balancing mechanism for server cluster, vol 3983, pp 549–557. https://doi.org/10.1007/11751632_60
8. Wu Y, Luo S, Li Q (2013) An adaptive weighted least-load balancing algorithm based on server cluster. In: Proceedings—2013 5th international conference on intelligent human-machine systems and cybernetics, IHMSC 2013, vol 1, pp 224–227. https://doi.org/10.1109/IHMSC.2013.60
9. Shi Y, Meng X, Zhao J, Hu X, Liu B, Wang H (2010) Benchmarking cloud-based data management systems, pp 47–54. https://doi.org/10.1145/1871929.1871938
10. Lee G, Bin L, OĹoughlin J (2014) Benchmarking cloud performance for service level agreement parameters. Int J Cloud Comput 3:3–23. https://doi.org/10.1504/IJCC.2014.058828
11. Sun B, Hall B, Wang H, Zhang D, Ding K (2014) Benchmarking private cloud performance with user-centric metrics. In: Proceedings—2014 IEEE international conference on cloud engineering, IC2E 2014, pp 311–318. https://doi.org/10.1109/IC2E.2014.74
12. Joel S, Philipp L, Jurgen C, Harald G (2014) Cloud WorkBench—infrastructure-as-code based cloud benchmarking. In: Proceedings of the international conference on cloud computing technology and science, CloudCom, 2015. https://doi.org/10.5167/uzh-98872
13. Bermbach D, Wittern E, Tai S (2017) Cloud service benchmarking: measuring quality of cloud services from a client perspective. https://doi.org/10.1007/978-3-319-55483-9

14. Rabl T, Frank M, Mousselly Sergieh H, Kosch H (2010) A data generator for cloud-scale benchmarking, vol 6417, pp 41–56. https://doi.org/10.1007/978-3-642-18206-8_4
15. Iosup A, Capota M, Hegeman T, Guo Y, Ngai W, Varbanescu A, Verstraaten M (2015) Towards benchmarking IaaS and PaaS clouds for graph analytics, pp 109–131. https://doi.org/10.1007/978-3-319-20233-4_11
16. Oussama S, Karim A (2018) impact live migration on cloud performance. SSRN Electron J. https://doi.org/10.2139/ssrn.3186348
17. Tukey JW (1977) Exploratory data analysis. Addison-Wesley Longman, p 688. ISBN-13: 9780201076165

An Automatic Optimization Method for BPG Compression Based on Visual Perception

Fangfang Li , **Sergey Krivenko** , **and Vladimir Lukin**

Abstract Better Portable Graphics (BPG), which is based on the High Efficiency Video Coding (HEVC), outperforms many other traditional lossy compression methods in terms of compression ratio and visual quality expressed in standard metrics. It might be necessary to provide a desired visual quality in some conditions where visual perception is of prime importance. In this paper, we propose an automatic optimization method for BPG coder intended to provide a desired visual quality, keeping in mind peculiarities of image perception in real-life applications. A two-step method is involved in the design; PSNR-HMA metric is employed to characterize the visual quality of a color image. The Kodak image set is used in our experiment to get an average rate-distortion curve and verify the coder performance. The results show that our approach provides the desired visual quality with a deviation less than 1 dB in terms of PSNR and PSNR-HMA. This method is simpler than recently proposed learn-based algorithms; less hardware and computer resources are required. Thus, it is expected that the proposed approach can be useful in various applications.

Keywords Lossy image compression · BPG · Visual perception · Two-step method

1 Introduction

The development of imaging system and technology has enabled to achieve high visual quality with rich image information. A huge number of photographs are acquired by smartphones and digital cameras every day; most of them are then

F. Li (✉) · S. Krivenko · V. Lukin
National Aerospace University, Chkalova Str. 17, Kharkov 61070, Ukraine

S. Krivenko
e-mail: krivenkos@ieee.org

V. Lukin
e-mail: lukin@ai.kharkov.com

F. Li
Nanchang Hangkong University, Fenghe South Ave. 696, Nanchang 330063, China

uploaded and shared on Facebook and other social media platforms. Besides, images are widely used in remote sensing [1–3], medicine [4, 5] and other areas [6]. Thus, people deal with the explosive increasing of amount of image data and necessity to store and transfer them effectively. In order to reduce the images' size for their transmission and/or storage, compression technology is more important than ever [7].

Compression can be performed in two ways, namely using lossless and lossy techniques. Lossless compression is carried out by removing redundancy in an image, and it is often used in medical imaging [5], bioinformatics [8], and other applications where it is required that a decompressed image does not differ from the corresponding original one; this takes place for such formats as PNG, GIF, etc. [9]. Lossy compression, on the other hand, introduces some distortions but provides larger compression ratios; as example, it is possible to mention such compression techniques as JPEG2000 [10], SPIHT [11, 12], AGU [13], BPG [4, 14, 15]. Due to high compression efficiency, lossy compression is widely used in numerous fields, i.e., multimedia, IoT, and so on.

As one knows, the main goal of image compression is to reduce the storage space and to meet time and bandwidth restrictions if images are transmitted via communication lines. Another aspect that should be taken into consideration is providing a desired quality of reconstructed (decompressed) images [16, 17]. For example, it might be desired to have them indistinguishable from the original ones if images are viewed (analyzed) by a human observer [18]. Such demands mean that the lossy compression process also needs to take visual quality into account. So, visual quality metrics have to be employed in the design and application of techniques used in lossy compression. Peak-Signal-to-Noise Ratio (PSNR) is the classic metric, but it is not adequate enough. Because of this, it becomes popular to employ some other metrics based on Human Vision System (HVS), which allows meeting the visual perception better [19–21].

Concerning visual perception in lossy compression, intensive studies have been focused on perceptual compression [7, 18, 22] that aims to provide an acceptable visual quality in terms of desired values of used metrics. Prediction-based methods allow predicting the visual quality characterized by a given parameter. Such a prediction is fast enough, but its accuracy still needs to be improved [23, 24]. Learning-based image compression schemes have been proposed recently [25, 26], they consist of encoder-decoder progress with a loss function and a trained neural network. Such schemes improve the compression efficiency and outperform JPEG2000, but go beyond the BPG [27]. Although for the extended hybrid method [28], the performance is slightly better, it has limitations as well. All these learning-based methods demand high hardware equipment and computations. Another direction of current research deals with optimizing traditional compression methods aiming at providing a desired visual quality. A two-step method [29] is a fast and efficient choice, and it has proved to work well for several gray-scale image compression coders [22, 29]. Having a preliminary stage that employs simulations for a set of test images with getting averaged dependences, this approach provides fast processing being applied

to a given image [30–32]. In view of its success in compressing gray-scale images, it is worth promoting this method used for color images.

The goal of this paper is to propose and analyze an automatic optimization of a desired quality providing for the state-of-the-art BPG compression applied to color images. The organization of the paper is as follows. The BPG coder is introduced in Sect. 2. Two-step method principle is described in Sect. 3. Then, the automatic optimization method for BPG tested on Kodak color images is proposed. In Sect. 5, we analyze the method accuracy for the metrics PSNR and PSNR-HMA. Finally, the conclusions are given.

2 BPG Coder

Better portable graphics (BPG) is an HEVC-based compression coder [33], which produces a higher compression ratio compared to JPEG for the same quality and even outperforms the novel learning-based algorithm in terms of PSNR and reconstructed image size [27]. BPG is designed in JavaScript that leads to its easy use for web browsers and all typical portable applications such as mobile devices [34]. Another advantage of BPG is open assess [33], this enables numerous researchers to participate in the optimization and improvement of the algorithm. These advantages result in superior compression characteristics of BPG and have led to its wide application.

Aggregating the above, BPG can be considered as a state-of-the-art lossy compression coder. It has a parameter Q responsible for varying and controlling the compressed image quality. However, the use of the same Q leads to quality that depends on image complexity ([27], see also examples in the next sections). So, it is important for BPG to enable an option of providing a desired quality, in particular, visual quality to meet human perception since human beings are often the terminal customers of images, especially color ones. In our previous work [34], BPG was enabled to provide a desired visual quality for gray-scale images due to the designed and optimized two-step method. In this paper, color image compression experiment is conducted to extend the universality of our method.

3 Related Work on Two-Step Method

The task of providing a desired quality in lossy compression is challenging. There are two issues needed to be overcome. Firstly, some adequate metrics are needed to assess the visual quality and characterize distortions (between the original image and the reconstructed one). Secondly, any lossy compression method has a parameter controlling compression (PCC), but such a parameter also needs some method (algorithm) to ensure a desired visual quality. We have recalled three approaches earlier where the two-step method seems to be the best choice for BPG encoder taking into account both time efficiency and precision.

A two-step method was proposed recently, and its feasibility has been verified for the DCT-based encoders [22] and DWT-based encoder SPIHT. This method operates as follows.

Firstly, there is a preliminary stage of data processing that has to be done once and offline to have averaged data till the moment of compressing a given image. At this stage, a large number of test images that simulate a variety of properties of other images to be compressed are used as an image set. Compression and decompression for each image are conducted on a series of PCC values, which covers a range we are interested in. Then, a rate-distortion curve (dependence of a considered metric on PCC) is built up from these data. The average rate-distortion curve is then obtained by averaging the metric values corresponding to the same PCC for all images in the set. The rate-distortion curves for the metric PSNR are presented in Fig. 1. They have been obtained for 12 images that were selected as the basic (training, simulation, preliminary) image set. As one can see, the trends for all curves are similar, including the average curve. However, PSNR for a given Q might vary by up to 10 dB (see data for $Q = 40$).

The obtained average rate-distortion curve is important since this characteristic can be used to determine an approximate PCC for providing a desired quality characterized by a given metric M. For example, as it follows from analysis of data in Fig. 1, one has to use PCC = $Q \approx 28$ to produce compression with PSNR about 40 dB. Note that this curve can be presented and saved in Table form as $M_{ave\ i}(PCC_i)$ where values of PCC_i cover the entire range of possible variation of this parameter (from 1 to 51 for the parameter Q for BPG).
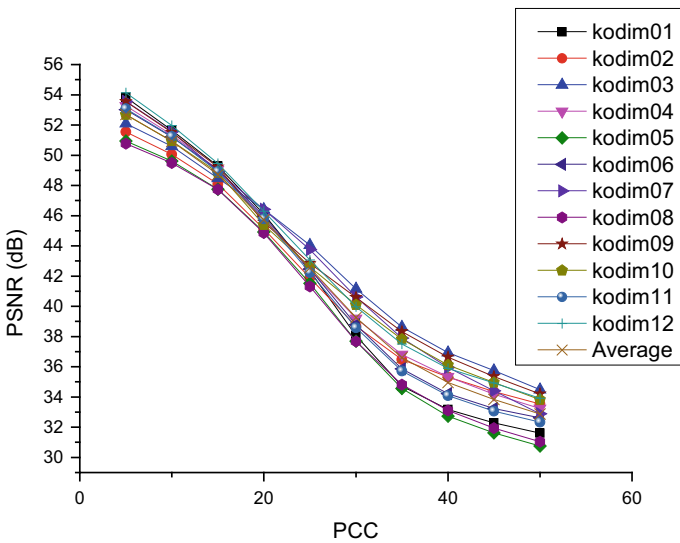
Fig. 1 The PSNR versus PCC (Q) dependences for 12 test images

In the first step of the two-stage procedure, the initial PCC_{init} has to be determined. Linear interpolation for Table data can be used. Then, PCC_{init} is calculated using the following equation

$$PCC_{init} = PCC_{est} + \frac{M_{des} - M_{ave}}{M'} \tag{1}$$

where M_{ave} is the value closest to the desired M_{des} at the right end of the corresponding interval of the average rate-distortion data array. PCC_{est} is the value corresponding to this M_{ave}. M' is the curve derivative for the corresponding PCC_{est}.

The next operation at the first step is compressing and decompressing the target image with PCC_{init}. After decompression, the metric value M_{init} has to be calculated between the original image and the decompressed one. For some images, this metric value can be appropriately close to the desired value [22], then the second step is not needed. If the accuracy characterized by $|M_{des} - M_{init}|$ is not appropriate, the PCC at the second step needs to be corrected using Eq. (2)

$$PCC_{des} = PCC_{init} + \frac{M_{des} - M_{init}}{M'}. \tag{2}$$

After this, the image is compressed using PCC_{des} and compression with providing a desired quality is considered completed.

4 Automatic Optimizing Method for BPG on Color Image

Color images contain richer visual information than the corresponding gray-scale ones. Color images are perceived by humans as reflections of the surrounding real world. However, it is more complex to deal with color images than with grayscale ones [35]. Even though the two-step method works well for BPG on grayscale images, we still need to take into account certain peculiarities of color image perception by humans and quality assessment.

4.1 Visual Quality Metrics for Color Images

First, we need to choose a visual quality metric able to characterize well human perception of color images. Of course, PSNR is the classic metric that can be left for further statistical analysis, but it has been shown that large PSNR values do not consistently mean a high perceptual quality, some color banding and blackness are not well-captured by PSNR [36].

PSNR-HMA [37, 38] is a modified image visual quality metric for HVS taking into account. In particular, it takes into consideration the perception of mean shift and contrast change. Advantages of PSNR-HMA are the following: (a) it is expressed in
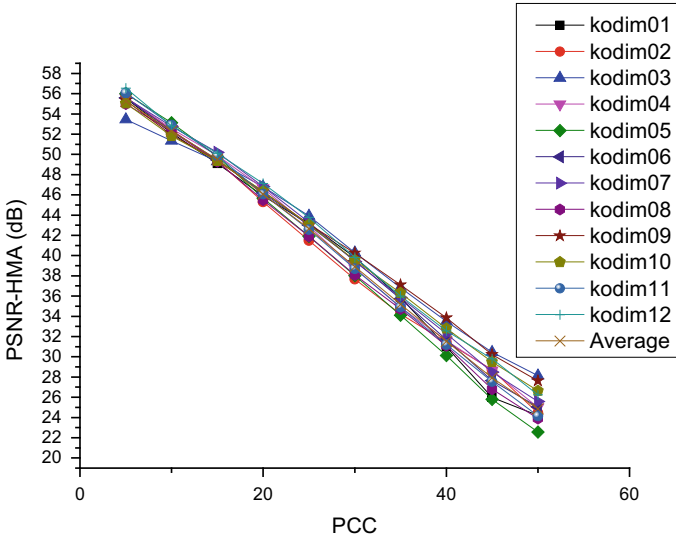
Fig. 2 The PSNR-HMA versus PCC (Q) dependences for 12 test images

dB and covers a wide range of values starting from 15 to 20 dB (annoying distortions, very bad quality) to > 40 dB (with high probability, invisible distortions); (b) its properties have been quite thoroughly studied [31]; (c) it can be easily and quickly calculated, its open access code is available at [39]. One more advantage is rate-distortion curve behavior close to linear (Fig. 2)—this is important for two-step procedure based on linear interpolations (1) and (2).

Figure 2 shows the rate-distortion curves PSNR-HMA vs PCC (Q for BPG coder) for 12 test images which come from Kodak image set. The plots are similar to those of PSNR vs PCC in Fig. 1. Despite of similar behavior, the plots are individual for each test image.

4.2 Parameter Controlling Compression (PCC) in BPG

In each lossy compression technique, there is a PCC. Mostly, it is quantization step or scaling factor [19, 21, 34]. For BPG, it is parameter Q [26]. Small Q results in less quantization and, thus, better visual quality of decompressed images [40]. According to [33], this value ranges from 1 to 51.

The parameter Q and a considered metric should be set into the Eqs. (1) and (2), it is displayed in Eqs. (3)–(6).

$$Q_{init} = Q_{est} + \frac{PSNR_{des} - PSNR_{ave}}{M'} \tag{3}$$

Img #1   Img #2   Img #3   Img #4   Img #5   Img #6

Img #7   Img #8   Img #9   Img #10   Img #11   Img #12
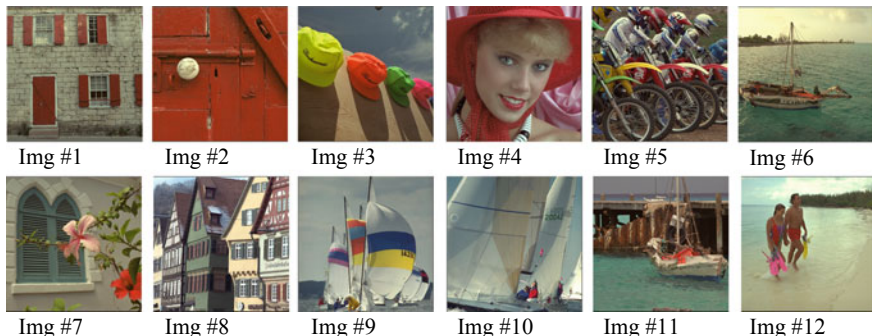
Fig. 3 Basic image set using images 1–12 of the Kodak test set

$$Q_{des} = Q_{init} + \frac{PSNR_{des} - PSNR_{init}}{M'} \qquad (4)$$

$$Q_{init} = Q_{est} + \frac{PSNR\text{--}HMA_{des} - PSNR\text{--}HMA_{ave}}{M'} \qquad (5)$$

$$Q_{des} = Q_{init} + \frac{PSNR\text{--}HMA_{des} - PSNR\text{--}HMA_{init}}{M'} \qquad (6)$$

A specific feature is that Q can be only integer. Thus, determined values of Q in (3)–(6) have to be rounded-off to the nearest integer.

4.3 Image Set

In our experiment, 12 images (see Fig. 3) from Kodak image dataset have been chosen to obtain the average rate-distortion curve off-line, the dependences of visual quality on the control parameter Q are displayed in Figs. 1 and 2. This set simulates a variety of data that can be met in practice and contains various types of images including simple structure, textural, colorful, etc. All of these images have the size of 512×512 pixels.

Other 12 images (see in Fig. 4) from Kodak image set are used to verify the proposed method, there are images of different scenes and people, the complexity of the images varies. This verification set can be treated as a tool to understand have we chosen the original image simulation set correctly or not.

Img #13   Img #14   Img #15   Img #16   Img #17   Img #18

Img #19   Img #20   Img #21   Img #22   Img #23   Img #24
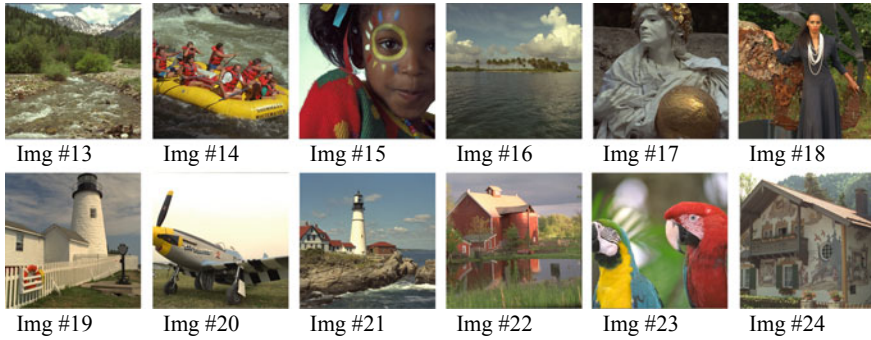
Fig. 4 Verification image set using image 13–24 of the Kodak test set

5 Analysis of the Accuracy

We focus our attention on lossy compression with either invisible distortions or with distortions that are visible but are not annoying. This approximately corresponds to the range from 33 to 38 dB in terms of PSNR. Experiments performed for the database TID2013 have shown that PSNR about 35 dB can be considered as distortion invisibility threshold. In our experiment, four values in this region have been set as desired ones, namely 40, 38, 35, and 33 dB. The same desired values have also been used for the metric PSNR-HMA. The obtained results are given in Table 1.

Here M_{des} is the desired value of the PSNR or PSNR-HMA. VAR_{fir} denotes the variance of providing visual quality at the first step, VAR_{sec} is variance of providing the desired quality at the second step. These variance values are determined by processing data for all 12 test images.

It can be also interesting to look at compression ratio and its variation for a given quality. To get imagination, we present the values of ΔCR that denote the maximal difference between CR values corresponding to different images for the same M_{des}.

Table 1 The obtained statistical data (training set)

Quality metric	M_{des}	VAR_{fir}	VAR_{sec}	ΔCR
PSNR	40	1.0067	0.0966	23.65
PSNR	38	1.298	0.1253	46.126
PSNR	35	1.511	0.0804	144.477
PSNR	33	1.0347	0.2058	737.175
PSNR-HMA	40	1.0615	0.0606	19.395
PSNR-HMA	38	1.1131	0.0651	28.45
PSNR-HMA	35	1.0861	0.0442	45.293
PSNR-HMA	33	1.1514	0.0575	72.542

Table 2 Statistical data for the desired PSNR-HMA$_{des}$ = 33 dB (verification set)

Test image	Qinit	PSNR-HMA_first	Qdes	PSNR-HMA_prov	CR
kodim13	38	32.373	37	33.243	20.133
kodim14	38	31.493	36	33.086	26.605
kodim15	38	32.352	37	32.936	62.066
kodim16	38	33.882	39	33.088	73.192
kodim17	38	33.23	38	33.23	56.905
kodim18	38	31.893	36	33.384	29.041
kodim19	38	33.692	39	32.981	59.072
kodim20	38	34.965	41	32.686	92.675
kodim21	38	34.104	40	32.61	53.549
kodim22	38	32.212	37	32.776	49.855
kodim23	38	33.92	39	33.236	86.915
kodim24	38	32.284	37	32.965	32.213
Variance		1.1614		0.0575	

It can be seen from data in Table 1 that our optimization scheme can indeed provide visual quality for BPG color image compression, at least, in the human perception region under interest. The max VAR$_{sec}$ go beyond 0.21; it has considerably (by about one order) reduced due to the second step. This indicates that the provided visual quality is enough close to the desired visual quality which can be set by a user. In some cases, even one (the first) step is enough to produce the desired quality for BPG applied to color image (the examples are the test image #17 in Table 2 and test images #15, 17, 19 in Table 3).

Analysis of ΔCR shows that, for the same desired PSNR, the compression ratio can be sufficiently different, especially for small values (33 dB) of the desired metric. The differences ΔCR for PSNR are larger than for the same PSNR-HMA.

Let us use the images #16 and #13 as examples. The original images are shown in Fig. 5a and d, the decompressed images with the desired PSNR-HMA = 33 dB are shown in Fig. 5b and e, and the reconstructed images with the desired PSNR = 33 dB are given in Fig. 5c and f. Comparing them, we can find that the distortions are practically invisible when the desired PSNR-HMA = 33 dB, the CR values are equal to 73.192 and 20.133, respectively. But when the desired PSNR = 33 dB, the distortions for the test image #13 are still invisible, while the distortions for the test image #16 are annoying, the CR values are equal to 768.731 and 31.575, respectively. The detailed data (for PSNR-HMA$_{des}$ = 33 dB) are given in Table 2. In turn, the detailed data (for PSNR$_{des}$ = 33 dB) are presented in Table 3.

In Table 3, the image #16 has the largest CR (768.731) while the highly textural image #13 has the smallest CR (31.575). The provided PSNR values are equal to 33.255 dB and 32.408 dB, respectively, both of them have the error less than 0.6 dB. According to CR for the test image #16, the obtained result is amazing. However,

Table 3 Statistical data for the desired PSNR$_{des}$ = 33 dB (verification set)

Test image	Qinit	PSNR_first	Qdes	PSNR_prov	CR
kodim13	49.000	31.267	40.000	32.408	31.575
kodim14	49.000	31.481	41.000	32.902	54.828
kodim15	49.000	33.034	49.000	33.034	385.121
kodim16	49.000	33.524	51.000	33.255	768.731
kodim17	49.000	33.078	49.000	33.078	301.304
kodim18	49.000	31.813	43.000	32.780	84.758
kodim19	49.000	33.030	49.000	33.030	241.745
kodim20	49.000	34.790	51.000	34.299	370.091
kodim21	49.000	33.194	50.000	33.041	303.233
kodim22	49.000	32.770	48.000	32.902	306.416
kodim23	49.000	33.905	51.000	33.426	367.731
kodim24	49.000	32.256	45.000	32.945	103.047
Variance		1.0347		0.2058	

a.Original image #16

b. Compressed image #16 (PSNR-HMA$_{des}$=33dB)

c. Compressed image #16 (PSNR$_{des}$=33dB)

d. Original image #13

e. Compressed image #13 (PSNR-HMA$_{des}$=33dB)

f. Compressed image #13 (PSNR$_{des}$=33dB)
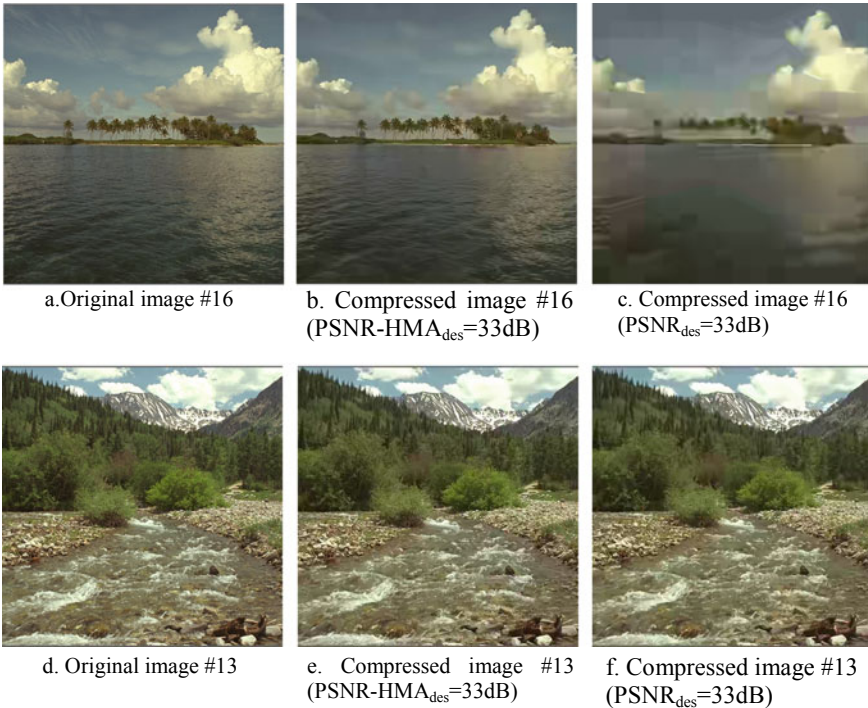
Fig. 5 Two examples with the desired metric value equal to 33 dB

the visual quality of decompressed image is unsatisfactory (see Fig. 5c), the distortions are obvious compared to the original image. Meanwhile, the visual quality for decompressed image #13 is acceptable to human perception.

In Table 2, the analyzed metric is PSNR-HMA with the desired value equal to 33 dB. The provided values of PSNR-HMA for test images #16 and #13 are equal to 33.088 dB and 33.243 dB, respectively. And the visual quality of both images is satisfactory, the distortions are almost invisible. Analysis of variance values in Tables 2 and 3 shows that accuracy after the second step has radically improved, and this shows expedience of using the two-step procedure. In addition, it shows that the simulation set has been chosen correctly.

Thus, some statements can be drawn. Firstly, there are distortions due to lossy compression that cannot be characterized by PSNR adequately even if a desired PSNR is provided and it is in the considered range. Because of this, we propose to use the metric PSNR-HMA to meet the human visual perception better and since it is more "reliable" than the metric PSNR. Secondly, by adopting the automatic optimization scheme for BPG, one can increase (maximize) the compression ratio while the provided visual quality meets the requirement.

The lower limit of $PSNR_{des}$ is about 33 dB whilst it is about 28 dB for the metric PSNR-HMA, this is because the PCC Q for BPG cannot be larger than 51 (there is the test image kodim20 in Table 3 for which the provided PSNR is equal to 34.3 dB although the maximal possible Q has been used). This also means that if the values of Q_{des} calculated according to (4) or (6) are larger than 51, then compression with $Q = 51$ has to be carried out at the second step.

As one can see, in some cases, the metric value provided at first step is approximately equal to the desired value, then the procedure can quit after the first step. This makes it possible to skip the second step and, thus, to accelerate the two-step procedure. For this purpose, a precision (maximal allowed error) can be preset - if the error of metric providing at the first step is less than a certain threshold value, the procedure is supposed to quit. According to data in Tables 2 and 3, variance after the second stage is about 0.06 dB^2. Then, the standard deviation σ is about 0.25 dB. Then, supposing the Gaussian distribution of errors and requiring the provided metric values to be within the limits from $M_{des} - 2 * \sigma$ to $M_{des} + 2 * \sigma$, the following stopping rule is possible: if the error $|M_{des} - M_{init}|$ is less than 0.5 dB, then quit.

6 Conclusions

It is the trend now to provide a desired visual quality that meets human perception in lossy compression. Meanwhile, this task is challenging since the coder performance characteristics of each image are different. Peculiarities of both the compression encoder and used metric need to be taken into account. Fortunately, the recently proposed two-step method has proved that fast providing of the desired quality is possible, at least for gray-scale images. This paper gives the automatic optimization method for BPG encoder for color images, the presented results of experiments

prove that this method works well enough. The metrics employed in our experiment are the classic metric PSNR and HVS-based metric PSNR-HMA. Comparison of decompressed images shows that the use of the metric PSNR-HMA is preferable since it takes into account peculiarities of human perception of images better. Our method allows setting the parameter Q employed as PCC in BPG encoder depending on a given image to be compressed. A possible range of setting the desired values of the considered metrics is established taking into account possible range of the parameter Q variation in BPG encoder.

References

1. Jinju J, Santhi N, Ramar K, Bama BS (2019) Spatial frequency discrete wavelet transform image fusion technique for remote sensing applications. Eng Sci Technol Int J 22(3):715–726
2. Villafranca AG, Corbera J, Martín F, Marchán JF (2012) Limitations of hyperspectral earth observation on small satellites. J Small Satell 1(1):19–29
3. Li D, Shao Z, Zhang R (2020) Advances of geo-spatial intelligence at LIESMARS. Geo-spatial Inf Sci 23(1):40–51
4. Yee D, Soltaninejad S, Hazarika D, Mbuyi G, Barnwal R, Basu A (2017) Medical image compression based on region of interest using better portable graphics (BPG). In: 2017 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, Piscataway, pp 216–221
5. Reddy MP (2018) The lossless medical image compression for telemedicine applications with delimiter. J Adv Res Dyn Control Syst 10(3):74–79
6. Ding Y (2018) Visual quality assessment for natural and medical image. Springer, Berlin
7. Zhai G, Min X (2020) Perceptual image quality assessment: a survey. Sci China Inf Sci 63:1–52
8. Barnouti NH (2016) Fingerprint recognition improvement using histogram equalization and compression methods. Int J Eng Res Gen Sci 4(2):685–692
9. Syahrul E (2011) Lossless and nearly-lossless image compression based on combinatorial transforms. Université de Bourgogne
10. Skodras A, Christopoulos C, Ebrahimi T (2001) The JPEG 2000 still image compression standard. IEEE Signal Process Mag 18(5):36–58
11. Doss S, Pal S, Akila D, Jeyalaksshmi S, Jabeen TN, Suseendran G (2020) Satellite image remote sensing for identifying aircraft using SPIHT and NSCT. IEEE Signal Process Mag 7(5):631–634
12. Jamel ALEM (2011) Efficiency SPIHT in compression and quality of image. J Coll Educ Women 22(3):627–637
13. Ponomarenko N AGU download page. http://ponomarenko.info/agu.htm. Accessed 6 June 2021
14. Albalawi U, Mohanty SP, Kougianos E (2015) A hardware architecture for better portable graphics (BPG) compression encoder. In: 2015 IEEE international symposium on nanoelectronic and information systems. IEEE, Piscataway, pp 291–296
15. Lin C, Yao J, Chen F, Wang L (2020) A spatial RNN codec for end-to-end image compression. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Piscataway, pp 13269–13277
16. Klinger T (2003) Image processing with LabVIEW and IMAQ Vision. Prentice Hall Professional
17. Vu CT, Phan TD, Banga PS, Chandler DM (2012) On the quality assessment of enhanced images: a database, analysis, and strategies for augmenting existing methods. In: 2012 IEEE southwest symposium on image analysis and interpretation. IEEE, Piscataway, pp 181–184

18. Zhang R, Isola P, Efros AA, Shechtman E, Wang O (2018) The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, Piscataway, pp 586–595
19. Nafchi HZ, Shahkolaei A, Hedjam R, Cheriet M (2016) Mean deviation similarity index: efficient and reliable full-reference image quality evaluator. IEEE Access 4:5579–5590
20. Deepsita SS, Sk NM (2019) Energy efficient binary adders for error resilient applications. In: 2019 IEEE conference on modeling of systems circuits and devices (MOS-AK India). IEEE, Piscataway, pp 23–28
21. Qi F, Zhao D, Fan X, Jiang T (2016) Stereoscopic video quality assessment based on visual attention and just-noticeable difference models. Signal Image Video Process 10(4):737–744
22. Li F, Krivenko S, Lukin V (2020) Adaptive two-step procedure of providing desired visual quality of compressed image. In: Proceedings of the 2020 4th international conference on electronic information technology and computer engineering. ACM, New York, pp 407–414
23. Ayoobkhan MUA, Chikkannan E, Ramakrishnan K (2017) Lossy image compression based on prediction error and vector quantisation. EURASIP J Image Video Process 2017(1):1–13
24. Krivenko S, Zriakhov M, Lukin V, Vozel B (2018) MSE and PSNR prediction for ADCT coder applied to lossy image compression. In: 2018 IEEE 9th international conference on dependable systems, services and technologies (DESSERT). IEEE, Piscataway, pp 613–618
25. Mentzer F, Agustsson E, Tschannen M, Timofte R, Van Gool L (2018) Conditional probability models for deep image compression. In: Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, Piscataway, pp 4394–4402
26. Lee J, Cho S, Beack S-K (2018) Context-adaptive entropy model for end-to-end optimized image compression. arXiv preprint arXiv:1809.10452
27. Patel Y, Appalaraju S, Manmatha R (2019) Human perceptual evaluations for image compression. arXiv preprint arXiv:1908.04187
28. Fu H, Liang F, Lei B (2020) An extended hybrid image compression based on soft-to-hard quantification. IEEE Access 8:95832–95842
29. Li F, Krivenko SS, Lukin VV (2020) Analysis of two-step approach for compressing texture images with desired quality. Aerosp Tech Technol 161(1):50–58
30. Fieguth P (2010) Statistical image processing and multidimensional modeling. Springer, Berlin
31. Grgic, M, Kunt, M, Mrak, M (2010) High-quality visual experience: creation, processing and interactivity of high-resolution and high-dimensional video signals. Springer, Berlin
32. Chen B, Wu Y, Coatrieux G, Chen X, Zheng YJCV (2020) JSNet: a simulation network of JPEG lossy compression and restoration for robust image watermarking against JPEG attack. Comput Vis Image Underst 197:103015
33. Bellard F BPG image format. https://bellard.org/bpg/. Accessed 15 March 2021
34. Li F, Krivenko S, Lukin V (2020) An approach to better portable graphics (BPG) compression with providing a desired quality. In: 2020 IEEE 2nd international conference on advanced trends in information theory (ATIT). IEEE, Piscataway, pp 13–17
35. Koschan A, Abidi M (2005) Detection and classification of edges in color images. IEEE Signal Process Mag 22(1):64–73
36. Mier JC, Huang E, Talebi H, Yang F, Milanfar P (2021) Deep perceptual image quality assessment for compression. arXiv preprint arXiv:2103.01114
37. Xiao R, Yeh M (2017) A new method with saliency detection for image quality assessment. In: 2017 IEEE international conference on consumer electronics—Taiwan (ICCE-TW). IEEE, Piscataway, pp 75–76
38. Prabhushankar M, Temel D, AlRegib G (2017) Ms-unique: multi-model and sharpness-weighted unsupervised image quality estimation. Electron Imaging 2017(12):30–35
39. Ponomarenko N PSNR-HMA download page, http://www.ponomarenko.info/psnrhma.htm. Accessed 27 March 2021
40. Mentzer F, Gool L.V, Tschannen M (2020) Learning better lossless compression using lossy compression. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Piscataway, pp 6638–6647

Agent Functionals in Monitoring Information Systems

Serhii Holub (ID)**, Svitlana Kunytska** (ID)**, and Viktor Grechaninov** (ID)

Abstract The processes of construction and use of agent functionalities in monitoring information systems (MIS) are investigated in the work. Several data-converting agents are combined by the monitoring information system into global and local multi-agent structures to provide the ability to perform monitoring tasks while reducing the informativeness of the input data arrays. The relevance of these studies is determined by the fact that there is a need for methods to identify the properties of monitoring objects in the process of changing their states. At the same time, the informativeness of the results of already conducted observations decreases in the process of discovering new properties of the monitored objects. There is a need to make new observations and to identify new properties of objects based on the results of already observed observations. The definition of basic concepts and principles of construction of agent functionalities is given. Hypotheses are formulated to increase the efficiency of monitoring by building agent functionalities. An experimental test of the formulated hypotheses was performed. To do this, the process of constructing an agent functional in monitoring the individual patient's response to the use of a standard treatment regimen was studied. Blood glucose was a simulated measure of health. It is experimentally proved that the method of ascending synthesis of elements allows ensuring the integrity of the agent functional. In this case, the SLE of the simulation results from the actual values decreases by more than 34%.

Keywords Intelligent monitoring · Agent functionality

S. Holub (✉) · S. Kunytska
Cherkasy State Technological University, Cherkasy, Ukraine
e-mail: s.holub@chdtu.edu.ua

S. Kunytska
e-mail: kunitskaya@gmail.com

V. Grechaninov
Institute of Mathematical Machines and Systems Problems of the NASU, Kiev, Ukraine
e-mail: grechaninov@mns.gov.ua

1   Introduction

Monitoring information systems (MIS) are the implementation of information technology for intelligent monitoring. They are used to reduce uncertainty in decision-making in the relevant subject area. Uncertainty reduction is achieved through the use of a process to predict the consequences of using each of the available strategies. To identify the properties of objects and processes involved in the selected strategies, ITIM intelligent monitoring information technology is built. The results of ITIM monitoring are used to calculate the situation—to determine the consequences of the each the existing strategies. If these consequences are known, the decision will be made under conditions of certainty.

Intelligent agents are used as one of the realizations forms the knowledge gained in the monitoring process Agents contain information about the properties of the object or group of objects to be monitored, about the patterns change of these properties, and the trends in the development of their interactions. They can predict the consequences of the strategy for which they were built, assess the impact of the external and internal factors, classify the state of the monitoring objects, to form conclusions based on the results of intellectual analysis of the results of the observation. For cases where the properties of the input data set change, the monitoring results become less informative, the MIS should provide the opportunity to adaptively increase the diversity and capacity of information retrieval and knowledge retrieval tools. This is achieved by decomposing the task of transforming the observation results, using agents to solve local problems of transforming the form of information and coordinating their interactions to perform the global task of MIS. As a result, a global multi-agent structure (GMS) is formed, which allows providing a separate result of MIS monitoring. The study the processes of the formation the synergetic connections between intelligent agents to create emergent GMS is an urgent task. Solving this problem will help eliminate the problems that arise when changing the profile of the MIS by the change of the next order of the decision-maker.

2   Results of Information Search of the Analogs

Agent functionalities (AF) are obtained by effectively combining agents in order to form new properties of global multiagent structures (GMS). GMS is a reflection of knowledge about the processes and objects ordered for monitoring, and at the same time serves as an algorithm for converting the results of new observations. With the help of GMS there is a profiling of information technology of intellectual monitoring in accordance with the individual orders of the decision maker in the relevant subject area [1].

The methodology of GMS construction involves the consistent formation of models, agents, agent functionalities and global multi-agent systems at certain stages of the organization of monitoring processes and processing of their results.

This paper describes the processes of formation of effective interactions between data-converting agents (DCA), which allow obtaining AF for given purposes.

Intelligent Monitoring Information Technology (IMIT) [1] is used to provide knowledge of decision-making processes in accordance with the individual order of the subject area. IMIT is implemented in the form of an intelligent monitoring system (IMS). Monitoring information technology is built each time individually at the request of the decision maker (ATS). Therefore, the intelligent monitoring system (IMS) has the ability to implement various information technologies without changing the software. Changing the monitoring task is formalized by changing the tasks of the simulated indicators, changing the list of the informative features, setting up the forming processes an array of the input data, and model synthesis.

The central processes implemented in GMS are modeling of monitoring objects. When changing the monitoring task, it becomes impossible to obtain additional results of observations that would coincide in time with the existing data and would provide a new task. The informativeness of the array of numerical characteristics when solving a new problem decreases. To build a utility model for a new task, it is necessary to increase the variety of means of model synthesis [2].

The use of the agent approach in IMIT is caused by the need to expand their capabilities to more fully reflect the properties of the monitored objects in the conditions of insufficient informativeness of the observation results. Possibility to apply the results of using traditional methods of group decision making [3–5] in solving problems of object recognition, methods of multilevel modeling [2], generalization of results obtained by several models, predictor complexation [6] and other existing methods for the formation of synergetic relationships between intelligent agents as elements of the methodology of multi-agent data processing in MIS, allows to increase the diversity of information modeling technologies. This paper investigates the processes of forming relationships with intelligent data processing agents in IMIT.

A special position in the review of methods and means of forming relationships between agents is occupied by methods of building hierarchical structures. In [7] the process of hierarchical clustering is described. A hierarchy of clusters based on agglomerate and separate strategy is built. The agglomerate strategy (inductive, bottom-up) involves the first stage of building clusters at the lower level. This is followed by the grouping of these clusters. According to a separate strategy (deductive, top-down), the hierarchy of clusters is formed by dividing existing clusters into subclusters.

Processes of using intelligent agents in monitoring systems have already become widespread. In [8] the use of intelligent agents (IA) in environmental monitoring technology is investigated. An agent is an information system that ensures the performance of certain monitoring tasks—monitoring of water agents, soils, atmosphere, emissions, generalization of results at the regional and national levels, agents of departmental services and so on. These agents have a typical structure, means of implementing functions. To ensure the gradual implementation of tasks by intelligent agents, the author proposes to provide additional functions for a multi-agent system: development of rules of conduct for agents in the network, protection of the

network from unauthorized actions of agents, detection of contradictions in the structure and actions of agents. It was not possible to find information about the methods of coordinating the interactions of agents and the types of models that determine the behavior of these agents.

In [9] proposes a list of indicators to assess the interaction of agents in the network, in particular the compatibility of goals and competence. It is proposed to provide agent interaction management with a separate network of intelligent agents, which operates on the basis of self-restraint and self-adjustment.

The use of genetic algorithms to form a strategy for the behavior of intelligent agents is described in [10]. The agent makes decisions after solving a local problem. The best strategy is formed by crossing several vectors of numerical characteristics of the parameters of available decision-making strategies (chromosomes), mutations of individual parameters (genes) and evolutionary selection of chromosomes with mutated genes that provide better values of quality function.

In [11] the structure of methods of coordination presented in Fig. 1 is offered.

The author proposes to form the interaction of agents on the basis of game approaches. Relationships between agents are proposed to be formed on the basis of the prefix form of counting predicates of the first degree of the Knowledge Interchange Format (KIF) [12].

The application of the predictor complexation method [6] allows increasing the reliability of forecasting by combining the results of using several models. With a

Fig. 1 Structure of methods of coordination of interaction of agents [12]
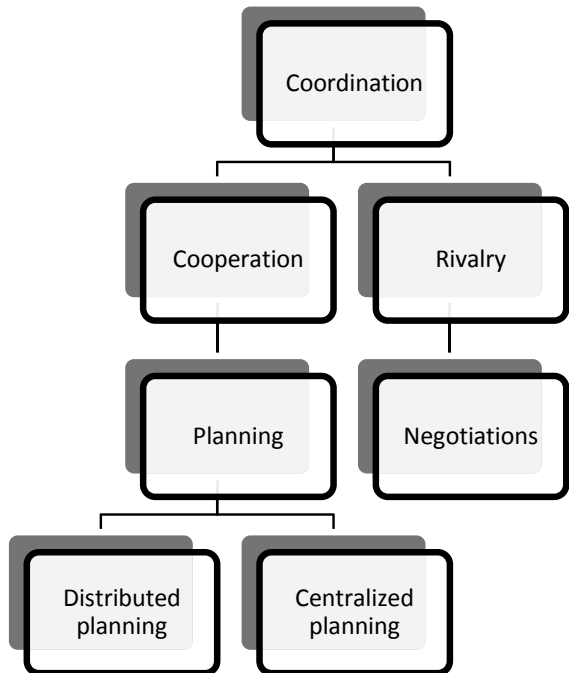
successful combination of models, the effect of emergence is formed—the efficiency of the combined models is higher than the efficiency of any of these models [13–16]. This system-wide property underlies successful collective solutions used by technical systems [4, 17–19]. Complexation is performed by a one-level combination of output signals of models that have a common simulated index:

$$y* = f[y(1), y(2), \ldots, y(n), X], \tag{1}$$

where $y(1)$, $y(2)$, …, $y(n)$ are predictive predictor models; X is the set of independent variables; $y*$ is the result of modeling several predictors.

The authors propose the implementation of the complexation process by sequential solution of problems [6]:

1. Evaluation of the quality of models according to predefined criteria.
2. Designing a team of complex methods by aggregating self-forecasts.
3. The use of a team of complex methods to obtain a comprehensive forecast.

The team of predictors is presented in the form of a linear combination of the most informative models:

$$w = \sum_{i=1}^{m} a_i y_i, \tag{2}$$

where w is the collective predictor, y_i is the vector of predicted values for the ith model; a_i is a vector of unknown coefficients.

The quality of the models is evaluated by the root mean square error, the criterion of regularity, the mean and maximum modulus of error, the correlation coefficient, the Darbin-Watson criterion [20]

To obtain a comprehensive forecast the following items are used [6]:

- the arithmetic mean of the signals obtained from each predictor;
- averaging the signals of the group of the best models [21];
- representation of the predictor in the form of regression with the calculation of the values of the coefficients by the method of least squares [22];
- model assault based on the use of a multi-row GMDH algorithm [24.]
- factor analysis. The complete list of reference predictors is replaced by a limited list of main components, for which the complexation algorithm is then applied [24];
- solution of the problem of minimization of the objective function in the space of predetermined priority coefficients of reference predictors [25];
- ensuring minimization of the variance of the error of the output signal of the collective predictor [26, 27].

All these methods are united by the fact that the complexation of models is aimed at improving the forecast of one, common to all models, indicator. Complexation is

limited to one layer, is the results of the collective forecast are not used for the next complexation.

In the case when the results of collective forecasting remain unsatisfactory, the methods of synthesis of multilevel models are used [2], and information technologies of multilevel intellectual monitoring are built on their basis. To improve the results of collective forecasting of the common indicator, multilayer modeling technologies have been developed. Methods of synthesis of multilevel models, duplication of levels, stratification of the structure of global functional dependences are used to take into account the mutual influences of several objects of monitoring and forecasting the development of the situation [2]. To optimize the description of the object of monitoring, clustering of observation points according to the characteristics of the models is used [28].

Similar approaches are used by the authors to construct multilayer neural networks in order to solve the problems of prediction [29], identification, classification [30] and clustering [31].

Thus, today there are methods that can be used to form links between agents, which provide the process of monitoring both individual objects and the situation as a whole. a comprehensive forecast the following items are used [6]: the arithmetic mean of the signals obtained from each predictor; averaging the signals of the group of the best models [21]; representation of the predictor in the form of regression with the calculation of the values of the coefficients by the method of least squares [22]; model assault based on the use of a multi-row GMDH algorithm [23], factor analysis.

3   Identification of the Problem and Formulation of the Purpose of Research

One of the problems hindering the development of information technology for intelligent monitoring is the existence of a contradiction between the need to reduce uncertainty in the decision-making process and the limited methods and means of intelligent monitoring while calculating the consequences of each strategy of varying complexity. Simultaneous calculation of the consequences of the application of several strategies with structures of different complexity requires the ability of the MIS to form several GMS in parallel, adaptable to the properties of each of these strategies. The processes of forming effective connections between intelligent agents in the construction of agent functionalities, which are elements of GMS, remain unexplored. Links that allow you to obtain new properties or improve existing MIS performance on specified quality indicators without a significant increase in the number of resources used are considered effective. Therefore, the purpose of this research is to develop a strategy for coordinating the interactions of intelligent agents in HMS to ensure the ability of MIS to calculate the consequences of the application of strategies by the decision maker.

4 Hypotheses

1. Improving the adequacy of HMS is achieved in the form of a systemic effect of a hierarchical combination of intelligent agents in agent functionalities. To ensure the system-wide properties of agent functionalities, in particular emergence, it is necessary to ensure the formation of synergistic connections between its elements.

2. Synergistic will be the relationship between agents, as a result of which one of the systemic effects is achieved, in particular emergence. The emergence of AF will be considered the property of increasing the adequacy or accuracy of the description of the properties of the object of the output signal of the agent structure or improving the performance of another function, which is achieved by coordinating the interactions of intelligent MIS agents. The construction of a multilevel structure of AF is achieved by forming synergistic relationships between agents by combining the method of ascending synthesis of elements [2] with methods of data mining—MSUA, neural network, neural network, and others.

3. The integrity of the functionality is achieved by combining a minimum number of agents that provide the emergence of the emergence effect.

5 Research results

The agent functionality determines the mapping of the nonlinear space of the output signals of the intelligent APD MIS on the scalar space of the results of observations—indicators of the state of the monitored object:

$$y_i = y[f_{1i}(x_1, x_2, \ldots, x_n), f_{2i}(x_1, x_2, \ldots, x_m), f_{3i}(x_1, x_2, \ldots, x_k),$$
$$f_{zi}(x_1, x_2, \ldots, x_l)], \tag{3}$$

where f_i is a nonlinear operator that reflects the structure of the model of the ADF monitoring object; z is the number of agents that form the structure of the agent functional y_i; i—the number of agent functionals that form the structure n, m, k, l—the number of features whose characteristics are included in the structure of the model of the object of monitoring the ADF.

Depending on the role of AF in MIS in its structure, agents of the appropriate purpose are combined. This paper describes the process of constructing information converters. The main quality criteria for such AF are accuracy, adequacy and stability. By AF accuracy we mean its ability to correctly display the properties of an object at a single point. Accuracy is traditionally characterized by error—the difference between the simulated value of the studied indicator of the state of the monitored object and its actual value at a given point of the multidimensional feature space, which is used in the form of absolute and relative values [32]:

$$\Delta_i = y_i - y_i^*, \tag{4}$$

$$\delta_i = \frac{y_i - y_i^*}{y_i^*} \cdot 100, \tag{5}$$

where Δ_i is the absolute error of modeling the studied indicator at the i point of the feature space; δ_i is the relative error of modeling the studied indicator at the ith point of the feature space; y_i is the simulated value of the studied indicator at the ith point of the feature space; y_i^*—the actual value of the studied indicator in the ith point of the feature space.

Adequacy refers to the ability of an AF to correctly reflect a change in the properties of a monitored object over a given period of time. One of the characteristics of AF adequacy is the standard deviation Sy of the simulation results y_i from the actual values y_i^* of the studied indicator [32]:

$$S_y = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - y_i^*\right)^2}{n}} \tag{6}$$

AF stability is characterized by its ability to ensure the value of accuracy and adequacy of the studied object is not worse than the results of observations set on the arrays, which were not used in the process of AF construction.

Functional dependence of blood glucose in patients of Cherkasy City Hospital №3 on the patient's health. The results of general blood and urine tests of patients and the results of biochemical analyzes were used as modeling variables. The list of symptoms was formed by an expert method with the involvement of experienced physicians who work in this medical institution and have significant experience in

Table 1 The list of signs of the initial description of a condition of patients

#	Features	Range of change of numerical characteristics
1	Hemoglobin, g/l	74–165
2	Leukocytes, g/l	4.2–23.4
3	Neutrophils, %	1–27
4	Lymphocytes, %	2–38
5	Total protein, g/l	46.80–87.10
6	Urea, mmol/l	2.80–18.20
7	Creatinine, mmol/l	0.060–0.446
8	Active partial thromboplastin time, seconds	19.50–58.40
9	Prothrombin index	20–100
10	Aspartate aminotransferase, μm/tsp	0.10–117.00
11	Alanine aminotransferase, μm/tsp	0.10–93.60

treating patients with such diseases. Table 1 lists the features and the range of changes in their values.

The initial description contained 79 observation points (patients). Of these, 59 points were used to train agent models. 20 points formed an examination sequence and did not participate in the training of models.

16 agents were generated who trained their models and optimized their structure according to their own patterns of behavior. Each of the agents solved the problem of identifying functional dependencies. After completing the formation of their own structure, the agents began to form relationships, forming AF by the method of ascending synthesis of elements. Figure 1 shows the results of studies of the influence of the number of models in AF on the quality of identification of the amount of glucose in the patient's blood.

As can be seen from Fig. 2, with the increasing structure of the agent functional, the value of the standard deviation of the simulation results from the actual values of the concentration of glucose in the patient's blood decreases. Thus, we observe the effect of emergence in the form of a decrease in the value of SLE with an increasing number of agents that have formed bonds between them. When the number of agents increases by more than 9, the value of SLE does not change, the integrity of the AF is achieved by combining 9 agents.
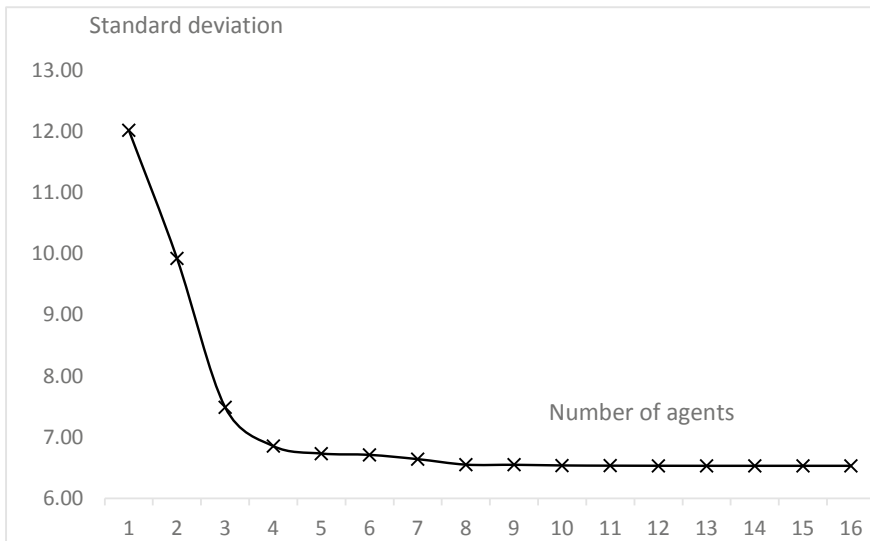
Fig. 2 The dependence of the standard deviation of the results of modeling the concentration of glucose in the patient's blood on the number of agents in the functional

6   Conclusion

The use of agent functionalities allows to improve the results of processing the results of observations in the intelligent monitoring system. The definition of agent functionals is given and the approach to their construction is described.

When agents are combined into functionals, a system-wide effect of emergence is formed in the form of a decrease in the value of the standard deviation of the simulation results from the actual value of the patient's blood glucose concentration by 34.03%. This suggests that the formation of inter-agent bonds by the ascending synthesis method provides emergence by multi-agent formation. The ability to optimize the number of agents in the functionality allows you to determine the minimum number of agents that provide the maximum emergent effect. It is experimentally proved that the integrity of the agent functional under experimental conditions is achieved by combining 9 agents.

References

1. Kunytska S, Holub S. Multi-agent monitoring information systems. 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. Springer, Cham, pp 164–171
2. Holub SV (2007) Bahatorivneve modelyuvannya v tekhnolohiyakh monitorynhu otochuyu-choho sere-dovyshcha. Vyd. vid. CHNU imeni Bohdana Khmel′nyts′koho, Cherkasy, 220 P
3. Rastrigin LA, Erenshteyn RKH (1981) Metod kollektivnogo raspoznavaniya. Energoizdat, Moscow, 80 p
4. Barabash YUL (1983) Kollektivnyye statisticheskiye resheniya pri raspoznavanii. Radio i svyaz', Moscow, 224 p
5. Zhukov VG, Bukhtoyarov VV (2014) Model' sinteza kollektivov intellektual'nykh informat-sionnykh tekhnologiy resheniya zadachi obnaruzheniya intsidentov informatsionnoy bezopas-nosti. Programnyye produkty i sistemy, no 1, pp 20–25
6. Rozenberg GS, Shitikov VK, Brusilovskiy PM (1994) Ekologicheskoye prognozirovaniye (Funktsional'nyye prediktory vremennykh ryadov). Tol'yatti, 182 p
7. Rokach L, Maimon O (2005) «Clustering methods.» Data mining and knowledge discovery handbook. Springer, USA, pp 321–352
8. Labzhyns′kyy VA (2013) Rozrobka mul′tyahentnoyi systemy obrobky danykh dlya zabezpechennya ekolohichnoho monitorynhu. Visnyk NTU "KHPI" 26(999):113–123
9. Novakivs′kyy II (2011) Zastosuvannya shtuchnoho intelektu dlya upravlinnya innovat-siynymy protsesamy u lantsyuhakh vartosti. In: Novakivs′kyy II, Rachyns′ka HV (eds) Visnyk Natsional′noho universytetu "L′vivs′ka politekhnika", vol 720. Menedzhment ta pidpryyemnytstvo v Ukrayini: etapy stanovlennya i problemy rozvytku, pp 303–309
10. Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning. Addison-Vesley
11. Kravets′ PO, Malanyuk YUV (2008) Modelyuvannya vzayemodiyi u hri mobil′nykh ahentiv. Visnyk Natsional′noho universytetu "L′vivs′ka politekhnika". № 610. Informatsiyni systemy ta merezhi, pp 167–174
12. Knowledge interchange format. http://logic.stanford.edu/kif/kif.html
13. Khrabrov YUB (1960) K voprosu o sostavlenii prognozov pogody kompleksnym metodom. Tr. tsentr. in-ta pogody 89:122–126

14. Bagrov NA (1962) O kompleksnom metode prognozov. Meteorologiya i gidrologiya, no 4, pp 14–21
15. Zhukovskiy YeYe, Brunova TI (1978) Statisticheskiye metody optimal'nogo kompleksovaniya al'ternativnykh prognozov. Primeneniye statisticheskikh metodov v meteorologii. Gidrometeoizdat, Moscow, pp 40–50
16. Brusilovskiy PM (1987) Kollektivy prediktorov v ekologicheskom prognozirovanii. Izd-vo Sarat. un-ta, Saratov, 104 p
17. L'yus RD, Rayfa KH (1961) Igry i resheniya. Izd-vo inostr. lit, Moscow, 642 p
18. Rastrigin LA, Erenshteyn RKH (1981) Metod kollektivnogo raspoznavaniya. Energoatomizdat, Moscow, 80 p
19. Bestuzhev-Lada IV (1982) Rabochaya kniga po prognozirovaniyu. Mysl', Moscow, 430 p
20. Suslov VI, Ibragimov NM, Talysheva LP, Tsyplakov AA (2005) Ekonometriya. SO RAN, Novosibirsk, 744 p
21. Makridakis S, Winkler RL (1983) Averages off forecasts: same empirical results. Manag Sci 29(9):987–996
22. Dayitbegov DM, Kalmykova OV, Cherepanov AI (1984) Programmnoye obespecheniye statisticheskoy obrabotki dannykh. Finansy i statistika, Moscow, 192 p
23. Brusilovskiy PM, Rozenberg GS (1983) Model'nyy shturm pri issledovanii ekologiche-skikh sistem. Zhurn obshch biol 44(2):254–262
24. Gorelik NA, Frenkel' AA (1983) Statisticheskiye problemy ekonomicheskogo prognozirovaniya. Statisticheskiye metody analiza ekonomicheskoy dinamiki. Nauka, Moscow, pp 9–48
25. Bronshteyn YeM, Brusilovskiy PM (1984) Protsedura formirovaniya kollektivnogo pro-gnoza. Primeneniye metodov teorii informatsii. Sov.radio, Moscow, pp 66–67
26. Dikinsen JR (1975) Some comments of the combination of forecasts. Opl Res Q 26(1):205–210
27. Yershov EB (1975) Ob odnom metode ob"yedineniya chastnykh prognozov. Statisticheskiy ana-liz ekonomicheskikh vremennykh ryadov i prognozirovaniye. Nauka, Moscow, pp 87–105
28. Holub SV, Burlyay IV (2014) Zastosuvannya klasteryzatsiyi pry formuvanni modeley informatsiynykh system bahatorivnevoho monitorynhu pozhezhnoyi bezpeky. Systemy obrobky informatsiyi: Zbirnyk naukovykh prats′. KH.: Kharkivs′kyy universytet povitryanykh syl imeni Ivana Kozheduba 4(120):203–208
29. Synehlazov VM, Chumachenko EY, Horbatyuk VS (2012) Metod reshenyya zadachy prohnozyrovanyya na osnove kompleksyrova-nyya otsenok. 29 Induktyvne mo-delyuvannya skladnykh system: zb. nauk. pr. – K.: MNNTS IT·S NAN ta MON Ukrayiny 4:214–223
30. Zhang GP (2000) Neural networks for classification: a survey. IEEE Trans Syst Man Cybern Part C (Appl Rev) 30(4), 451–462. https://doi.org/10.1109/5326.897072
31. Bock HH (1998) Clustering and neural networks. In: Rizzi A, Vichi M, Bock HH (eds) Advances in data science and classification. Studies in classification, data analysis, and knowledge organization. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-72253-0_37
32. Matematicheskaya statistika: Uchebnik. Ivanova VM, Kalinina VN, Neshumova LA (1981) i dr. Vyssh. Shkola, Moscow, 371 p. Author F (2016) Article title. Journal 2(5):99–110

Performance Model for Convolutional Neural Networks

Dmytro Rahozin and **Anatoliy Doroshenko**

Abstract The paper describes the method for computational performance evaluation of convolutional neural network-based inference for object recognition tasks in embedded and automotive areas. The convolutional neural network pipeline includes hundreds of computationally intensive stages, each has some basic stage type. We propose to analyze the memory traffic generated by neural network-based application to predict its performance running on an embedded system. Computationally intensive inference stages are analyzed to get resource consumption information for simulated model description, the model is used for coarse-grained simulation of the inference pipeline. The pipeline is represented as a finite state automaton model of a computing process and the automaton is annotated with number of consumed resources. The pipeline is simulated using the previously gathered resource consumption information at the analyzed network stages. The simulation result allows to determine if the inference pipeline fits the computational platform, to find out if multiple pipelines can fit the parallel computational platform and analyze computing resource consumption to choose the proper hardware configuration, that is important for embedded applications. This approach may be used for any other computationally intensive processes.

Keywords Convolutional neural networks · High performance computing · State automaton · Simulation

1 Introduction

Today convolutional neural network (CNN) inference is a challenging task for high-performance computing. Image/object recognition employing CNN becomes more and more popular in automotive and embedded market segments. The main

D. Rahozin (✉)
Software Systems Institute of NASU, Acad. Glushkow Ave, 40/5, Kyiv 03187, Ukraine

A. Doroshenko
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Peremohy Ave, 37, Kyiv 03056, Ukraine

effort for hardware platform computational performance improvements goes to massive parallel computations and computation unit's architecture optimizations for "neural networking"—for example, newer mobile video cards and specific neural network/artificial intelligence processors. Modern CNNs have variable complex computational profiles [1] and it is practically unreal to map a CNN on some massive parallel computational platform (e.g. video card) without excessive resource waste. One of the key problems is the performance of memory subsystem (cache memory subsystem or video memory hierarchy) so we need some appropriate tools to analyze it—for example, simulation of CNN computational resource consumption on a hardware platform.

2   Problem Definition

There are only several popular CNN-based methods used for object recognition which appeared on some digital image today. The methods can be separated into two large groups: the first group is based on finding an object or a part of it in some limited image area, possibly previously marked for further analysis, the second is based on segmenting the whole image in parts, each part refers to some semi-standard environment type (e.g. "road", "car", "bike", "building", "sky", etc.). Real practice shows us that the computing performance of the first group is much better, so these methods are the most widely used in current industrial applications. The big part of object recognition software has some modification of popular Yolo network family inside, for example Yolo-v3 [2], Yolo-v4 [3], Poly-Yolo [4], which allows to fine tune object recognition precision for customer tasks. Yolo-type networks has better performance than another type networks, for example R-CNN [5], as Yolo makes only one pass over the image, but R-CNN needs two passes—the first to extract image areas possibly containing objects, and the second to detect objects inside these areas. Yolo-type networks and another network types are highly profitable for effective parallelization at massive parallel processors, as multiple Yolo feature extractors (FE) [2] are run over the small parts of source image simultaneously, but at the end of computation the results of all FE runs should be joined into the list of found objects with probabilities. The Yolo-v3 FE structure is rendered on the top of Fig. 1, other network types FE may differ in structure. Practically all the neural networks have similar structure of joined big blocks which process large 4-dimensional tensors (stacks of 3D matrices) and the number of possible computational blocks is more than 150. So, the neural processing tasks is an application with high demand for computing power and memory traffic, both naturally is a challenge for effective algorithm parallelization and prediction of possible speed-up which can be achieved if the neural network is run on massive parallel processor.

Usually the Yolo-v3 FE chain (Fig. 1) processes quite small 32 by 32 pixels square image areas, finding a part of some object type inside this area. So, if the source image resolution is 1280×1024 pixels, the $32 * 40 = 1280$ FE copies should be run over the image to get the final result. Several years ago, the resolution of processed network
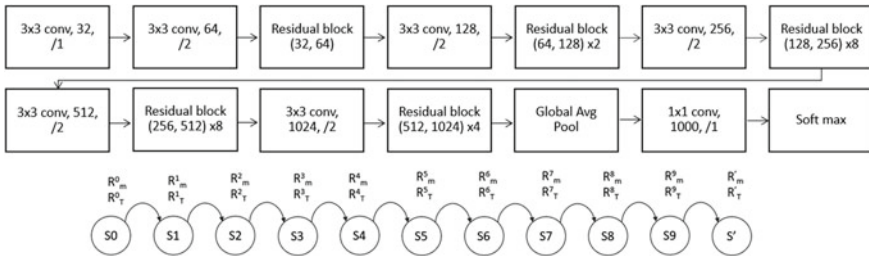
Fig. 1 Darknet-53 feature extractor and corresponding state automaton

was limited to near 600×400 pixel resolution, but this resolution in era of Full HD is out of interest as many details are lost on image. Off-the-shelf camera resolution constantly increases, for example, switching from 1280×1024 resolution to Full HD (1920×1080) doubles the number of pixels on the image, doubles the computing power necessary to process the image and leads to continuously increasing demand for high-performance computations and increasing the computing power of new platforms.

The severe problem for practical or industrial object recognition is that the CNN-based object recognition should run on some specialized hardware with performance accelerators: video processing cards, accelerations engines (for example Qualcomm Artificial Intelligence Engine), or another specialized computer. There is no way to develop some general hardware: different network types have different architectures and use different FE chains.

3 Performance Model

The basic problem of running CNN on any computational platform is the limited computational resources. No case was detected when a CNN was perfectly fit the computational platform. Therefore, researchers should always consider the tradeoff between accuracy and performance of CNN and consumed resources. For many application cases there is no way to change platform parameters, so the only way to fit a CNN on a platform is the optimization of CNN structure and application parameters.

The key question is the ability of the underlying hardware to run a CNN under defined time frame and computational resources. The performance model answers the question if the hardware platform is able to perform the CNN, calculating the final timings and resource consumption. In order to achieve the goal, the model should be supplied with actual resource consumption data of different FEs, and these data should be extracted from actual CNN runs on real data streams. So, in next chapters we consider (1) how to model platform performance for a CNN run; (2) which data

should be supplied to parametrize the performance model; (3) how to gather data necessary for modeling preferably in automatic way.

3.1 Performance Model Basics

Here only the inference stage on CNN is considered for simplicity, as the learning stage can be performed on "big" computers where the platform limitations are loose-fitting. We consider massive parallel computing platforms, for example video cards or special functional units such as Neural Processing Engine in Qualcomm system-on-chips [6], as older acceleration technologies (single CPUs and SIMD engines) performance improvements are not considered any more due to insufficient computing performance.

The most cases for CNN applications are related to various computer vision tasks for different environments and tasks, but usually for object recognition. If the object recognition is performed by unmovable cameras, the requirements for CNN computers are loosen, as the platform power envelope is not restricted severely. But the demand for robotic/automotive vision raises now, as a modern vehicle is equipped with a dozen of color/infrared cameras which should be processed in real time with limited power envelope. The typical modern CNN engineering task is to map the developed and evaluated CNN into limited embedded resources. The mapping process requires a lot of handwork including the optimization an initial CNN structure previously developed in environment such as Tensorflow (www.ten sorflow.org) or Caffe (caffe.berkeleyvision.org). The multistage CNN processing should be fit into restricted resources of a system-on-chip and system performance should be analyzed before moving to field tests. The mapping process may fail in at least two cases: (1) CNN resources cannot fit into platform resources, so more computing power should be added; (2) there are non-used computing resources, so less expensive computing system can be used.

Therefore, the cornerstone problem is a cheap and fast evaluation of computational system parameters which is able to run the CNN without paying extra penalty (usually dozens of dollars per an installed system) for using the platform. It's clear that the precise simulation of a CNN run is not possible, but even semi-fast one (taking a week or more) and semi-precise (10% accuracy) helps significantly to evaluate the final system performance. This evaluation should be done without actual CNN performance check at target platform: the goal is to simplify the analysis process for CNN application performance and it is necessary to evaluate the performance with the use of the CNN instance running on a general-purpose computer (for example, x86 platform) and correct the evaluation result which employs the performance model reflecting the actual architecture of an CNN accelerator. The thoughtful reader should claim here that the x86 platform is a doubtful case for simulating a CNN as the most CNN runs are made by video cards. Still the main difference between x86-based system and video card is the different parallelism degree (e.g., 16 processing at CPU units vs 2000–4000 units at GPU), but the memory throughput is not scaled so much,
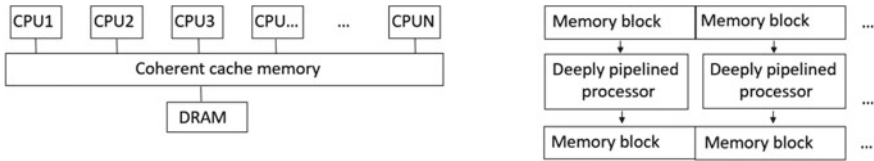
Fig. 2 Hardware structures for performance simulation

so the memory interface become the limiting factor for CNN runs due to nature of data processing in CNN pipeline. X86-based CNN runs showcase the memory performance very well, so it is a good starting point for performance evaluation. Figure 2, left, shows the typical system-on-chip architecture for performance simulation.

Of course, our model cannot properly evaluate all possible computing architectures, for example data-driven deep pipelines (systolic computing pipelines) at Fig. 2, right. But it is not true that the performance model, described below in text, does not work for such type of computing architecture, or other architecture types—but the parametrization of the model changes significantly, and this is the subject for future work.

A good example of the complexity of the CNN mapping task is the business story of Israeli Mobileye company, which has spent near 10 years to provide a handy and useful road traffic analysis solution to a mass market. Modern researches have not so long timeframe and need some useful tools to predict the final performance of a CNN running on some resource restricted hardware.

Previously [7] a simplified general model for parallel application runs was considered. The model aim was somehow different—the simulation was dedicated to prove the deadlock-free processes execution on a computational platform with limited memory and time resources. The parallel running programs are represented as an abstract model, which is represented as a graph of automaton states and transitions between the states. The resource marks R_M^i and R_T^i were introduced for each state, each mark corresponds to consumed amount of memory and minimal time resource necessary to execute code associated with some state. Transitions are assumed to take very short time, so the resource consumption is considered only for states. Additional states allow to make a lock on shared resources. Let's consider the FE chain at Fig. 1 and its performance model. Here each state reflects some FE, and each FE at Fig. 1 has its own R_M^i and R_T^i marks, so the initial assumption is that the total FE chain execution time is the sum of corresponding R_T^i times. So basic consumed FE resources are considered as the sum of corresponding time and memory consumptions, but here we have several inconsistencies. First, any computing accelerator uses a sophisticated memory hierarchy, which affects the execution time, so R_T^i is $f(R_M^i)$. Second, memory hierarchy may differ from platform to platform so $f(R_M^i)$ differs. Third, the memory hierarchy may be controlled by the FE stages execution, and this directly affects $f(R_M^i)$. All these inconsistencies (but, in reality, they are peculiarities of CNN execution and a hardware platform architecture) require us to improve the automaton state markings.

3.2   Model Parametrization

We should make several assumptions about the performance model. Previously we have said that massive parallel computing architectures are considered, and they get necessary data from the memory—Fig. 2, left—shared among multiple processors or execution units. A memory hierarchy, which use several levels (or zero) of cache memory is assumed to be used in computing system, as rendered at Fig. 3. The non-essential part (which can be omitted while simulation parametrization) are enclosed in dashed lines.

In order to execute the FE chain roughly two sets of operations should be executed: $C = \{c_i\}$, the set of computing operations, performed at processor/accelerator functional units and $M = \{m_i\}$, the set of memory access operations, performed by memory hierarchy and memory controller. Total R_T^i is the function $R_T^i = f_t(C, M)$, which can be approximated as $f_t(C, M) = k_c\|C\|+k_m\|M\|$, where k_c and k_m are the weight coefficients, reflecting the relative input of computational and memory instructions into total FE execution time. $\|C\|$ and $\|M\|$ are the total time necessary to perform C and M operations on a platform. An important model assumption is that $k_c \ll k_m$, so that the memory performance and memory hierarchy structure input is the basic values for R_T^i calculation on *both* platforms—x86 platform for simulation and target platform for final execution. This assumption is very well illustrated by thread execution in common video cards, where the thread is stopped if the current load/store operation goes to memory, so another thread, which is not blocked by waiting memory load/store operation, continues execution. If any computing instruction, including multiply-and-add, is executed during one processor clock cycle, on-chip memory access takes several clocks, but the external memory load/store operation takes dozens of nanoseconds and all video card architecture solutions are developed to hide this high external memory latency.

To reflect the memory hierarchy an additional resource parameter R_i^D is introduced—it defines the memory bandwidth which is necessary to perform FE operations (or some operation sequence). R_i^D depends mostly on (1) memory hierarchy configuration, which include timings for different cache memory levels access times, cache level miss time and main DRAM memory latency time; (2) static configuration of tensors, handling actual data for FE execution. Memory hierarchy timings are well defined for the target platform; tensor configuration is static for real life applications, so R_i^D may be approximated from computing platform and CNN tensor configurations.
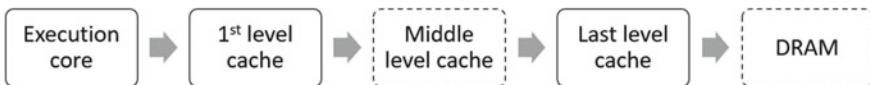
Fig. 3 Simulated memory hierarchy for the performance model

3.3   Getting Numerical Parameters for Performance Model

Finally, after defining the performance model as state automaton and resource consumption annotation concept we need to supply the model with real data—so to measure resource consumption of a CNN inference run. There are at least two ways to do the measurement—the first, do actual runs and catch processor performance counters for the whole program run; the second way—to run the real program under a processor simulation tool and make a full performance profile for a CNN run, including clock-accurate cache memory simulation. There are drawbacks for both ways—the actual run way is inaccurate but fast, the simulated run is slow but precise. We have chosen the second way—and use a cache profiler to analyze a CNN run.

Well-known tool Valgrind [8] was used for performance analysis—particularly its sub-tool Cachegrind. Cachegrind emulates the memory hierarchy of a x86 compatible processor and is able to collect all events related to cache memory state changes and emulate cache content handling. The tool is able to execute x86 binary instructions without a recompilation or additional instrumentation and store all cache memory events in a log file. After the simulated program run is finished, these events are mapped to source lines of original code and form a listing, which includes the source code lines and the numbers of cache events of different types "caught" at the corresponding source code line during the simulation. As the cache memory configuration is a parameter of a simulated program run, this tool is used to analyze the workload behavior for different cache sizes, configurations, and also to check various cached data handling politics.

The actual CNN run details are described in the next section.

Figure 4 renders the sample results of the run. Rendering full details is not possible for this paper, but Cachegrind is able to get information to create ‖C‖ and ‖M‖ values. So, the FE chain (Fig. 1) memory and computing resources consumption may be predicted from platform description and CNN input configuration. Values of C, M, k_c, k_m allow to compute $f_t(C, M) = k_c‖C‖+k_m‖M‖$ for each automaton state and to represent an FE chain or CNN complex pipeline (employing multiple FE chains) as a parametrized state automaton, which reflect the coarse-grain structure of CNN inference. It should be noted that sometimes for simplicity we can omit ‖C‖ values

```
      .                      .            void gemm_nn(int M, int N, int K, float ALPHA,
      .                      .                         float *A, int lda,
      .                      .                         float *B, int ldb,
      .                      .                         float *C, int ldc)
      0                      0            {
      .                      .                int i, j, k;
    203,502 ( 0.00%)         0                for (i = 0; i < M; ++i) {
 128,693,815 ( 0.33%)  6,173,888 ( 0.11%)       for (k = 0; k < K; ++k) {
  64,296,032 ( 0.17%)  6,912,576 ( 0.12%)           PUT_IN_REGISTER float A_PART = ALPHA * A[i * lda + k];
 128,659,898 ( 0.33%)  3,081,313 ( 0.05%)           for (j = 0; j < N; ++j) {
 32,167,924,337 (83.51%) 5,680,796,154 (99.02%)        C[i*ldc + j] += A_PART*B[k*ldb + j];
      .                      .                       }
      .                      .                   }
      .                      .               }
    237,419 ( 0.00%)      33,917 ( 0.00%)    }
```
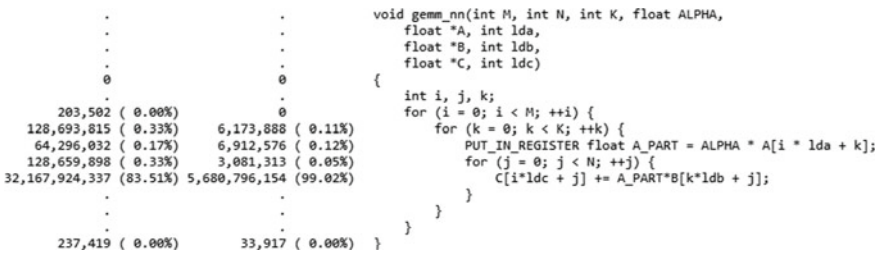
Fig. 4 Example profile tool output for basic CNN GEMM function

and use the got $\|M\|$ value for simplicity, employing it as a R^i_D for the state automaton. This greatly simplifies the simulation.

The simple data profiling works well for 1-threaded FE execution, where the only thread uses cache memory of all levels. For multiple thread execution the first level cache is used exclusively for a thread, but the last level cache is divided among all executed thread. For two threads, resource consumption is R^1_D and R^2_D. Let $R^C_D = R^1_D \cap R^2_D$. R^{12}_D is the common data shared by threads. Local thread data $R^{Li}_D = R^i_D \backslash R^C_D$, so the necessary resource allocation for both threads is $R^C_D + R^{L1}_D + R^{L2}_D$, that may be significantly less than initial $R^1_D + R^2_D$.

To start the simulation [7] each state (that corresponds to some FE) is annotated with resource consumptions R^i_T, R^i_M, R^i_D. Actually, R^i_D is united with R^i_T marking. These markings allow to simulate efficiently the parallel execution. As the number of execution processors is static for a CNN run, corresponding automatons for parallel executed FEs are just duplicated. Although the pipeline and the corresponding automaton looks simple, the simultaneously working pipelines may trap into deadlock struggling for resource R^i_M (on the other hand R^i_T just changes the execution time). The simulation system uses a greedy method to simulate parallel execution of state automatons, including avoiding possible deadlock states—it is possible to change this greedy politic, but it is well enough for our simulation purposes, as we need final timing value to check CNN inference run time and performance. Changing basic cache operational times and main memory access times (memory may have static and/or dynamic type) allows to map simulated CNN run performance to some target architecture and this is the goal of our research.

4   Simulation Sample

In order to simulate the CNN run, we have implemented a simplest possible finite state automaton simulation model, which uses R^i_M and R^i_T annotations in the simplest form. The automaton architecture is very similar to Fig. 1, but has more states inside, reflecting complex Yolo-v4 computing structure.

We consider modern *Darknet* [9] framework, implemented in C programming language. The list of supported layers in *Darknet* is comprehensive and can be easily extended, it has several code forks for research work and became an industry standard very fast. *Darknet* C-based implementation supports both multicore x86 CPUs and Nvidia video cards and enables accurate performance analysis even for complex CNNs. For performance analysis we use Yolo-v4 inference [3] running on 1200 × 800 pixels image under *Darknet* [9] environment, it runs well for 6x core Ryzen platform with 16 GB RAM and for Nvidia RTX 2070 video card with 8 GB memory on board.

Specially for simulation runs *Darknet* was compiled for single processor run to get the accurate cache memory hierarchy performance log. Valgrind profiling is used in cache simulation mode [10] to gather memory resource consumption information. Despite of huge amount of code, big memory footprint and extreme complexity of

the code Cachegrind simulator was able to profile Yolo-v4 inference without any kind of software recompilation from start to end and made a performance log for *Darknet* source tree. This reaffirms the quality of Valgrind simulation tool and ability of Cachegrind to support heavy workloads.

The sample Cachegrind output is shown at Fig. 4. The original output was stripped to fit into the paper format: Cachegrind simulates 2 cache memory levels, providing accounting for cache memory hits and misses for read or write operations. Cache memory performance log on Fig. 4 is stripped and accounts 1st level cache misses summary for all neural network stages.

This function generates more than 83% of real memory references. The video card memory traffic is practically the same for this function.

After checking the results of Fig. 4 we clearly see that basic information for CNN configuration, e.g. M, N, K tensor dimensions—as they are *gemm_nn* function parameters are insufficient for performance modeling, as for nested loops the memory traffic is mostly defined by cache memory hierarchy behavior and theoretical evaluation of the traffic volume is not possible. As Valgrind tool simulates parametrized cache memories and hierarchies, it is possible to check resource consumption for different cache configurations (sometimes changing only cache memory size is well enough) and analyze memory traffic changes at least by hand for simulation models.

It should be noted that we have limited this discussion with only one showcase at Fig. 4. Due to similar basic computational kernels the performance accounting refers to functions such as *gemm_nn*. Following Fig. 5 highlights this (functions provide different convolution types).

The simulation is performed by a simple tool, written from scratch and parametrized with simplest C preprocessor definitions. It handles simple automaton descriptions with R_T^i, R_M^i and R_D^i. Basic greedy methods are used for resource allocation with default assumption that one CPU core is necessary to execute an automaton. The simulation may be implemented also using other high-level tools, e.g. Simulink or AnyLogic, but now we are focusing mostly on defining the performance model, measuring memory performance to get actual data and map the simulated CNN inference evaluation to a target platform, as these are the hardest things to do.

Fig. 5 Performance report result generated by Intel Vtune Amplifier Tool

Considering results, the simulated memory traffic out of 16 MB 2nd level cache memory for *gemm_nn* function (Fig. 2) is near 1.5 billion cache misses per an inference cycle or 24,000 MB (24 GB) per an inference cycle. We are not considering first-level cache performance for simplicity. The maximum off-the-shelf DDR4 DRAM memory access speed is 3200 MB/s, real speed for such tasks is up to 2000 MB/s (this is a good suggestion for video card memory). So, the simple simulation result shows that a single-channel DDR4 memory for usual multicore processor limits CNN inference speed to 24,000/2000 = 12 s at multicore CPU, and this result could be projected well for video cards. Sure, the picture looks somehow different for video cards due to different memory access methods and higher bandwidth, but the principle is the same so the results of basic simulation can be used to evaluate the CNN performance on a target platform.

So, what are the benefits we achieve with the use CNN performance modeling using cache profiler? The basic benefit is that we can evaluate the real memory traffic, as all existing computing platforms are employing some kind of cache memory or use some register memory instead of cache as video cards do. Second, as the platform peak memory traffic is well known for engineering specifications, the initial memory subsystem performance considerations allow to check if the hardware platform has enough performance to fit CNN execution into some timeframe. Third, the performance model simulates process synchronizations in multiprocessor system, so enables the evaluation for even a distributed computation platform (running CNN inference), for its performance limits and possible deadlocks due to lack of resources. The evaluation is profitable for modern ARM-based platforms with more than 8 CPUs and several neural processing units. Fourth, the performance model allows to determine how many identical FE chains or CNN pipelines can be run simultaneously without resource deadlocks—as running only one FE chain cannot utilize all video card resources. Also, the simulation evaluates the peak resource consumption and time to run multiple pipelines.

Before moving to the approach shortcomings discussion, we render similar (if compare the result outline) performance report got by Intel OpenVino toolkit for a CNN run (Fig. 5).

This sample was got from Intel site docs.openvinotoolkit.org. It highlights that most of the time is spent in internal convolution functions. Intel Vtune Amplifier tool is used to gather performance information over a CNN network run (using OpenVino toolkit). The performance information at Fig. 5 was gathered with use of internal performance counters, which register internal processor events, such as cache hits, cache misses, instruction issues, branch prediction hits, etc. Vtune tool renders practically the same information as Cachegrind simulator and performs much faster as it does not simulate each processor instruction. Vtune is an excellent tool for fast evaluation of "hot spot" at software runs, still simulator allows to analyze internal cache memory and get any statistics we need, and the benefits of this approach are discussed in the next section.

Author's experience shows that even basic digits—such as total memory amount necessary for running CNN on a platform—is useful for early platform evaluation, as its late evaluation may cause the bad surprise for developers at testing stage.

5   Parameter Extraction Shortcomings and Ways to Avoid It

The proposed approach has some drawbacks and shortcomings, and their source is the mismatch of a simulated architecture (the Cachegring cache memory model) and actual target architecture and simulation simplifications. The simplification is made intentionally, but there are major issues that affect the final computations. Looking back at Fig. 3 we see that two important objects are not considered—(1) intermediate cache level—as most modern processors have 3 cache memory levels and (2) DRAM. Extra cache level does not change the performance significantly, as two-level cache is enough for semi-precise simulation, but dropping DRAM simulation is a possible issue. The fact is that fast DRAM access is possible inside DRAM pages—they have at least 4 KB size, and only several pages per DRAM chip are active simultaneously (page data is stored in output amplifiers). Accessing data in non-active pages introduce DRAM access delays up to 70 ns, it needs to be accounted. The simplest way to cover this delay is to use some mean value for memory access accounting. The hard way is to employ a DRAM memory simulator, similar to [11]. The physics of DRAM chip cannot overcome the old problem of opening memory pages in short time—since 1990s opening a DRAM page requires 60–70 ns, and the modern improvements in memory bus technology for DDR5 DRAM only allow very fast transfer a DRAM block to cache memory. Another way to hide page opening latency is the multibank memory structure. So, if the memory access is localized among several memory pages the memory access speed should not degrade significantly. But the size of 4D tensor for FullHD camera (2 M pixels) is several megabytes, so basic processing uses hundreds of pages still the mean memory access time can be predicted. Still, the question of employing more precise DRAM controller simulation is questionable and should be reviewed in case of sophisticated memory controller, or in the case if the compiler which enables "deep" software prefetch to hide the memory access latency.

Another shortcoming is the result of an older cache simulation paradigm inside cache simulators. Taking back to Fig. 4 we see that function *gemm_nn* incapsulated memory accesses from all FEs in CNN run. Let us consider Fig. 6.

On Fig. 6 multiple FEs share the common low-level procedures to provide computation library. So "summary" values for memory traffic and cache misses for kernels A, B and C on Fig. 6 do not reflect the real memory traffic generated after instructions execution in FEs #1, #2, #3. The example of such "kernel" is *gemm_nn* function (Fig. 4), which is annotated with memory accesses from all the "high level" code. Such "summary" simulation hides the real memory traffic and makes cache accesses evaluation incorrect for simulations for multicore processors which use one common 2nd or 3rd level cache memory. Also, it prevents analysis for memory traffic per FE, and hardens analysis of FE work with "hot" (previously loaded with reusable data) and "cold" (without reusable data) cache states.

In order to overcome this limitation, we are considering to implement the API extension for Cachegrind to handle several copies of memory traffic summary information and switch them depending on FE type or CNN pipeline type. This work looks
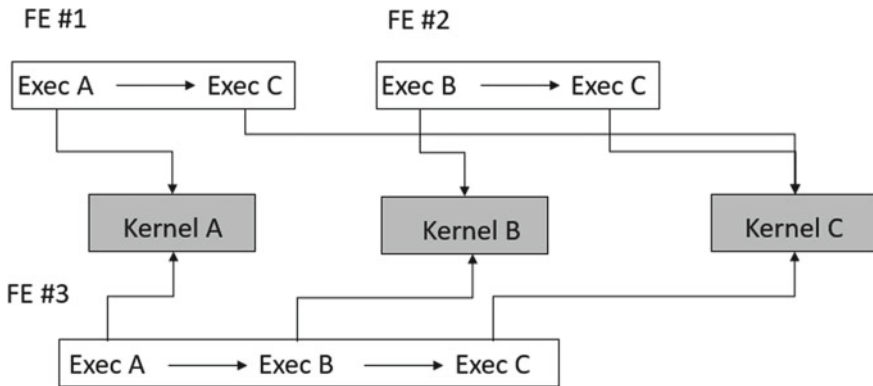
Fig. 6 Typical call graph for FE chain execution

to be quite technical, but requires us to pay attention to details, so we have considered to review Cachegrind tool extension in a separate paper for memory hierarchy analysis techniques in simulator. Also, we should render a set of approaches which should be applied to analyzed program to get useful information about software structures similar to rendered at Fig. 1.

Another point of consideration is if our attention to CNNs looks too narrow for modern workloads and if our application study is really useful. First, internal CNN computational kernels (Figs. 4 and 5) look quite similar to linear algebra matrix operations still with real difference—high memory traffic with complex memory access patterns. Even inside Yolo-family of object recognition CNNs we observe absolutely different computational structures. Memory amount required for CNN runs—8 GB or more—challenges every computational platform and provide enough cases for platform analysis. Second, despite of similar execution structures where CNN is sequence of FEs, the performance patterns of different CNN structures are different. For example, Yolo9000, Yolo-v3, Yolo-v4 and Yolo-v5 [2–4] have different structures, very different workload and definitely different demand for platform computational resources. Anyway, useful performance observations for different CNN structures also require a separate study, as the observation itself cannot study much about particular applications. An application study requires analysis of changed CNN dimensions, batching parameters and different cache sizes. The result of this study should be the platform performance requirements and their considerations will be our next study.

6 Conclusions and Future Work

We have described a useful simulation methodic and scenario for analysis of highly loaded computational platforms. Quite simple simulation system is used to evaluate

the application performance, and the simulated system description is derived from real world computational workloads using memory profiling tool. Our technique can be applied for other computational math and physics problems, for example rigid body modeling or weather forecasts. Illustrative and simple models improve the quality of applications analysis and parallelization degree.

In near future we are planning to modify Valgrind/Cachegrind tool so that simulation system can be improved with more detailed information about program runs. Also, we are going to evaluate performance requirements of different CNNs types.

References

1. Li Z, Yang W, Peng S, Liu F (2020) A survey of convolutional neural networks: analysis, applications, and prospects. ArXiv, abs/2004.02806
2. Redmon J, Farhadi A (2018) YOLOv3: an incremental improvement. ArXiv, abs/1804.02767
3. Bochkovskiy A, Wang C, Liao H (2020) YOLOv4: optimal speed and accuracy of object detection. ArXiv, abs/2004.10934
4. Hurtík P, Molek V, Hula J, Vajgl M, Vlasánek P, Nejezchleba T (2020) Poly-YOLO: higher speed, more precise detection and instance segmentation for YOLOv3. ArXiv, abs/2005.13243
5. Girshick R (2015) Fast R-CNN. In: 2015 IEEE international conference on computer vision (ICCV), Santiago, Chile, 2015, pp 1440–1448
6. Ignatov A et al (2019) AI benchmark: running deep neural networks on android smartphones. In: Leal-Taixé L, Roth S (eds) Computer vision—ECCV 2018 workshops. ECCV 2018. Lecture notes in computer science, vol 11133. Springer, Cham
7. Rahozin D (2020) Models of concurrent program running in resource constrained environment. In: Proceedings—UkrProg-2020, Kyiv, Ukraine, pp 149–156
8. Weidendorfer J, Kowarschik M, Trinitis C (2004) A tool suite for simulation based analysis of memory access behavior. In: Proceedings of the 4th international conference on computational science (ICCS 2004), Krakow, Poland, June 2004, 8 p
9. Redmon J. Darknet: open source neural networks in C. http://pjreddie.com/darknet. Accessed 1 March 2021
10. Nethercote N, Walsh R, Fitzhardinge J (2006) Building workload characterization tools with Valgrind. In: IEEE international symposium on workload characterization (IISWC 2006), San Jose, CA, USA, Oct 2006, 102 p
11. Mirosanlou R, Guo D, Hassan M, Pellizzoni R (2020) MCsim: an extensible DRAM memory controller simulator. IEEE Comput Archit Lett 19(2):105–109

Adaptive Selection of Turbo Code Parameters in Wireless Data Transmission Systems

Sergei Zaitsev , **Vladyslav Vasylenko** , **Yuliia Tkach** , **Yurii Posternak** , **and Svitlana Lytvyn**

Abstract The article proposes a method for multilevel parametric adaptation of the turbo code codec in wireless data transmission systems. The choice of the actual parameters of the turbo code depends on the values of the normalized number of changes of the sign in the posterior-a priori the log likelihood ratio (LLR) for transmitted data bits of the turbo decoder. The results of simulation showed that the application of the method of multilevel parametric adaptation of the turbo code codec reduces the number of errors by 2–3.3 times, and also makes it possible to increase the reliability of information transmission in comparison with the known results, for example, the fourth generation 4G LTE-Advanced mobile communication.

Keywords Turbo codes · Wireless networks · Posterior-a priori the log likelihood ratio · Reliability of information transmission

1 Introduction

Mobile communications and wireless data transmission systems are widespread in our time and continue to develop intensively.

S. Zaitsev · Y. Tkach (✉) · Y. Posternak · S. Lytvyn
Chernihiv Polytechnic National University, 95 Shevchenka Str., Chernihiv 14035, Ukraine
e-mail: tkachym79@gmail.com

S. Zaitsev
e-mail: serza1979@gmail.com

Y. Posternak
e-mail: posternak21051976@meta.ua

S. Lytvyn
e-mail: chdtu.fld@gmail.com

V. Vasylenko
Institute of Telecommunications and Global Information Space of NAS of Ukraine, 13 Chokolivskiy bulv., Kyiv 02000, Ukraine
e-mail: vladvasilenko9@gmail.com

One of the main and urgent tasks is to increase the reliability of information transfer. It is possible to achieve an increase in reliability through the use of error-correcting codes, such as: convolutional codes, low-density parity-check codes (LDPC codes), Hamming codes, polar codes, turbo codes (TC), Reed-Solomon codes, etc. The most effective are TC and LDPC codes.

TCs are adopted by the mobile communication standards of the third generation 3G UMTS [1], the fourth generation 4G (LTE-Advanced) [2] and the fifth generation 5G [3]. They are inferior in energy efficiency to the Shannon boundary of 0.5 dB for a channel with additive white Gaussian noise at a coding rate $R = 1/3$ [4, 5].

LDPC codes [6] are accepted by the 10GBase-T [7], WiMax [8], Wi-Fi [9], DVB-S2 [10] standards. They are somewhat inferior to TC at a low signal-to-noise ratio.

It should also be noted that polar codes are gaining popularity [11]. They are currently being considered for possible adoption in future sixths generation 6G mobile standards.

In wireless data transmission systems of the fourth generation 4G [12–16] and the fifth generation 5G, single-layer parametric adaptation schemes for TC codecs are used for adaptation, in which only the coding rate R changes.

Consequently, it becomes necessary to implement schemes for multi-level parametric adaptation of TC codecs.

2   Analysis of Research and Publications

The work [17] presents an adaptive algorithm for discrete optimization of signal structures and coding rate of error-correcting code for DS-CDMA systems. Depending on the signal-to-noise ratio values, the error probability value is calculated for various signal modulation schemes and compared with the specified value. The required parameters are selected depending on the comparison results. In this case, channels with Additive White Gaussian noise (AWGN) and Rayleigh fading are considered.

In [16], an algorithm is presented for discrete optimization of the coding rate using a pseudo-random interleaver depending on the value of the error probability.

Work [18] presents a HARQ (hybrid automatic repeat request) system with soft/hard decision-making during decoding. Soft decoding decision making is limited by signal-to-noise ratio of 1.4 dB. The adaptation process is as follows: the size of the information block and the coding rate change depending on the signal-to-noise ratio. In this case, a channel with an AWGN is considered.

In [19], LDPC codes are used and an AWGN channel is considered. The main idea of this work is the real-time evaluation of the signal-to-noise ratio and further adaptation depending on this value. As an indicator of the reliability of information in the work, the average probability of a bit error of decoding $P_{B\ \text{dec}}$ was chosen.

In [20], LDPC codes are also used and one-level parametric adaptation of the coding rate R is applied. The choice of the optimal coding rate is done by comparing the current bit error with table values.

3 Formulating the Goals of the Article

The purpose of the article is to develop a method for multilevel parametric adaptation of a turbo code codec in wireless data transmission systems depending on the values of the normalized number of changes in the sign of a posteriori-a priori LLR about the transmitted data bits of the TC decoder.

4 Presentation of the Main Material

In Fig. 1 a block diagram of a modified TC encoder is shown. The encoder refers to schemes PCCC—parallel concatenated convolutional code. A block of data K is fed to the block generator. In the block generator additional service information is added to the data and the sequence $\overline{X}^C$ is obtained. Then $\overline{X}^C$ is fed to the first recursive systematic convolutional encoder (RSCE-1) and a serially connected interleaver and the second encoder (RSCE-2). Depending on the normalized value of the LLR decisions Q are made about changing the number of iterations, the type of interleaver, and the RSCE polynomials. $Q = \left\{ \overrightarrow{Q_1}, \overrightarrow{Q_2}, \overrightarrow{Q_3} \right\}$, where $\overrightarrow{Q_1}$ is the change of the number of iterations, $\overrightarrow{Q_2}$ is the change of the interleaver, $\overrightarrow{Q_3}$ is the change of the RSCE polynomial.

The input of each decoder gets the information received from the output of the corresponding RSCE and taking into account the passage of the channel with AWGN.
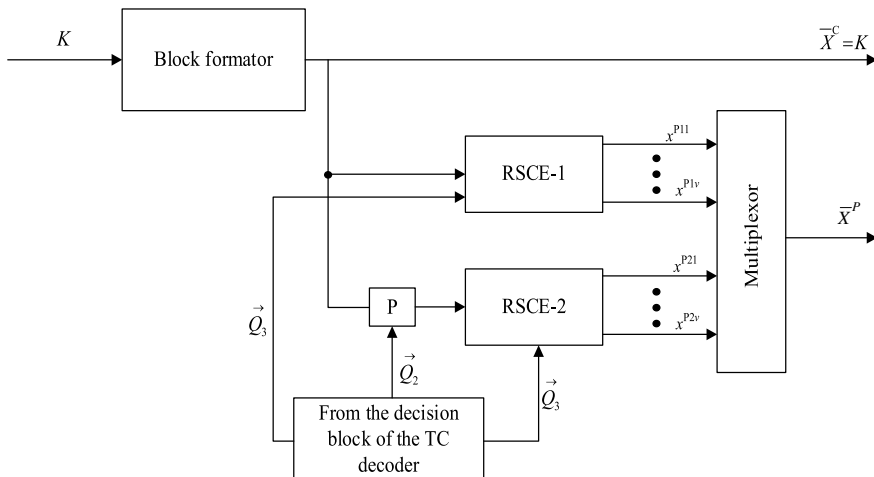
Fig. 1 Block diagram of the modified TC encoder

An iterative decoder consists of interconnected component decoders. One iteration includes two serially connected component decoders, two interleavers and deinterleavers. The interleaver and deinterleaver in the TC decoder circuit converts the error packets at the output of the current decoder into single errors, that greatly facilitates and improves the operation of the next decoder [21].

The block diagram of a modified iterative turbo code decoder with a decision maker block, which contains a module for calculating the normalized number of changes in the sign of a posteriori-a priori LLR F^*, is shown in Fig. 2.

Logarithmic relation likelihood function $L(x_t)$ about transmitted bit x_t in general, is defined as follows [22, 23]:

$$L(x_t) = L_c(x_t) + L_a(x_t) + L_e(x_t), \tag{1}$$

where $L_c(x_t)$—channel information, $L_a(x_t)$ and $L_e(x_t)$ a priori and the posterior LLR data bit x_t respectfully. If $L\left(x_t^C\right) \geq 0$, it is considered that bit $x_t^C = 1$ was transmitted, else $x_t^C = 0$.

In Fig. 3 a block diagram of module for calculating the number of changes in the sign of a posteriori-a priori LLR $F^{d,j}$ for a d decoder, $d \in \overline{1, 2}$ and j decoding iterations, $j \in \overline{1, I}$.
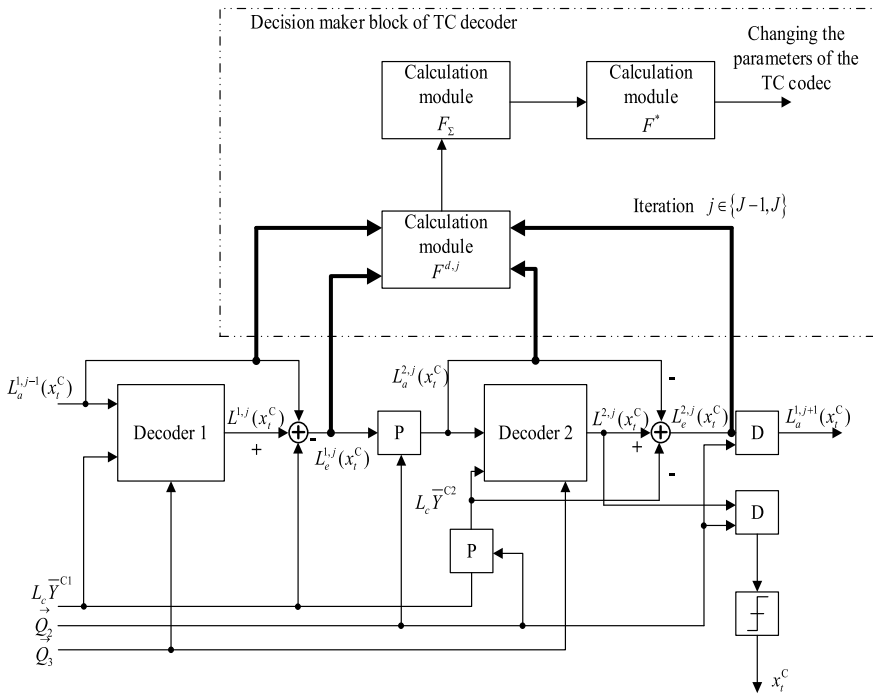
Fig. 2 Block diagram of the modified iterative TC decoder
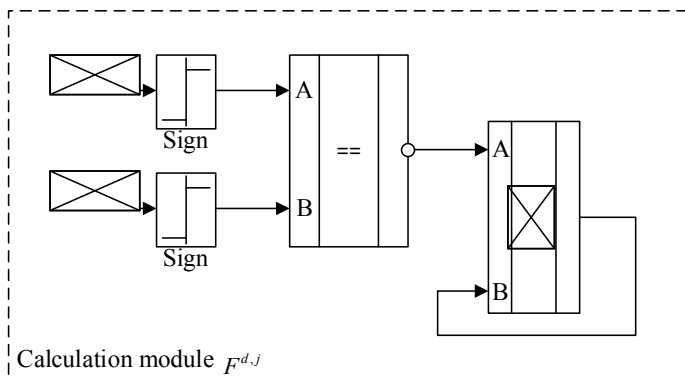
Fig. 3 Block diagram of the module for calculating the number of changes in the sign of a posteriori-a priori LLR $F^{d,j}$

The input of calculation module $F^{d,j}$ receives a priori LLR $L_a^{d,j}(x_t^C)$ and a posteriori LLR $L_e^{d,j}(x_t^C)$. Then their signs are determined and if they are not equal, then the value of $F^{d,j}$ is incremented:

$$F^{d,j} = F^{d,j} + 1, \quad \text{if } \text{sign}(L_a^{d,j}(x_t^C)) \neq \text{sign}(L_e^{d,j}(x_t^C)), t \in \overline{1, N}, \qquad (2)$$

where N is a number of bits in a block.

The more often the values of $F^{d,j}$, the more often incorrectly decoded bits appear, which leads to a deterioration in the reliability of information reception.

The total indicator of the number of changes in the sign of a posteriori-a priori LLR for all decoding iterations F_Σ is determined:

$$F_\Sigma = \sum_{j=1}^{I} \sum_{d=1}^{2} F^{d,j}. \qquad (3)$$

If the transmission channel does not affect the transmitted information, then the value of the number of changes in the sign of the a posteriori-a priori LLR will be minimal and equal to the number of transmitted information bits:

$$F_{\min} = N \qquad (4)$$

If the transmission channel affects the transmitted information in such a way that normal decoding becomes impossible, then the maximum value of the number of changes in the sign of a posteriori-a priori LLR will be equal to:

$$F_{\max} = \frac{N(2I + 1)}{2}. \qquad (5)$$

In practice, for a qualitative assessment, it is better to use the normalized value of the number of changes in the sign of a posteriori-a priori LLR:

$$F^* = \frac{F_\Sigma - N}{N(I - 0.5)}. \tag{6}$$

The values of F^* are used to select the parameters of the turbo code codec.

The algorithm for implementing the method of multilevel parametric adaptation is presented below.

Step 1. Formation of the set of values of systematic information bits X^C of size $1 \times N$, produced by the turbo code encoder:

$$X^C = \{x_1^C, x_2^C, \ldots, x_N^C\}, \tag{7}$$

where N is the length of the input.

Step 2. The initial parameters of the data transmission system are set: polynomial generator $(1.7/5)$, Log-Map decoding algorithm, TC coding rate $R = 1/2$, regular interleaver (deinterleaver), the number of bits in the transmitted (received) block $N = 1000$.

Step 3. Formation of the set of a priori values of the about the transmitted data bits at the 2nd decoder of the jth iteration

$$LA = \left[L_a^{2,j}\left(x_1^C\right) \; L_a^{2,j}\left(x_2^C\right) \; \ldots \; L_a^{2,j}\left(x_N^C\right) \right]. \tag{8}$$

Step 4. Formation of the set of a posteriori values of the LLR about the transmitted bits

$$LE = \left[L_e^{2,j}\left(x_1^C\right) \; L_e^{2,j}\left(x_2^C\right) \; \ldots \; L_e^{2,j}\left(x_N^C\right) \right]. \tag{9}$$

Step 5. Cycle execution: if $\text{sign}\left(L_a^{d,j}\left(x_t^C\right)\right) \neq \text{sign}\left(L_e^{d,j}\left(x_t^C\right)\right)$, then $F^{d,j} = F^{d,j} + 1$, $F_\Sigma = \sum_{j=1}^{I} \sum_{d=1}^{2} F^{d,j}$, $F^* = \frac{F_\Sigma - N}{N(I-0.5)}$, $t \in \overline{1, N}$ for all bits of a block of length N, decoders d, $d \in \overline{1, 2}$, decoding iterations j, $j \in \overline{1, I}$.

Step 6. Calculation of the average normalized number of changes in the sign of a posteriori-a priori LLR based on the results of receiving K data blocks:

$$\tilde{F}^* = \frac{\sum_{i \in 1}^{K} F_i^*}{K}. \tag{10}$$

Step 7. Depending on the value of the average normalized number of changes in the sign of the a posteriori-a priori LLR, the parameters of the turbo code codec will change. If the calculated value is less than the specified value, then it is considered that the data transmission system meets the conditions and the parameters do not change. If it is greater, then depending on the value, the parameters of the turbo

Table 1 Interleaver types depending on the average normalized number of sign changes of a posteriori-a priori LLR

$\tilde{F}^*$	Interleaver types
[0.25−0.3]	Regular
(0.3−0.35]	Pseudo-random
(0.35−0.4]	S-random

Table 2 Values of the encoder polynomials depending on the average normalized number of changes in the sign of a posteriori-a priori LLR

$\tilde{F}^*$	Polynomial
[0.0−0.05]	(1, 17/15)
(0.05−0.1]	(1, 35/23)
(0.1−0.15]	(1, 75/53)
(0.15−0.2]	(1, 171/133)
(0.2−0.25]	(1, 371/247)

code codec change.

$$\begin{cases} \tilde{F}^* > 0.4, \ I = I + 1 \\ 0.25 < \tilde{F}^* \leq 0.4, \ P \Leftrightarrow \\ \tilde{F}^* \leq 0.25, \ RSCE \, Polynomial \uparrow \end{cases} \tag{11}$$

If value $\tilde{F}^* > 0.4$, then the number of decoding iterations is increased by 1. When $0.25 < \tilde{F}^* \leq 0.4$ is within these limits, the interleaver type is changed in accordance with Table 1. When $\tilde{F}^* \leq 0.25$, then the encoder polynomial is changed in accordance with Table 2.

Table 2 shows the values of the encoder polynomials depending on the average normalized number of changes in the sign of the posterior-a priori LLR for the coding rate TC $R = 1/2$.

5   Analysis of the Results

Evaluation of the characteristics of the information transmission reliability using the proposed method of structural adaptation of the encoder and decoder of the turbo code was carried out using the method of simulation. To compare the proposed results, the fourth generation 4G LTE-Advanced mobile communication standard was chosen as an analogue. The simulation was carried out in the Visual Studio 2019 environment. The simulation results were obtained based on the reliability $\alpha = 0.95$, $t_\alpha = 0.95$ (the argument of the Laplace function), relative accuracy $d = 0.1$.

A turbo code similar to the 4G LTE-Advanced standard was used with two component encoders, a Log-Map decoding algorithm, a regular interleaver (deinterleaver), the number of bits in the transmitted (received) block $N = 1000$. The coding rate
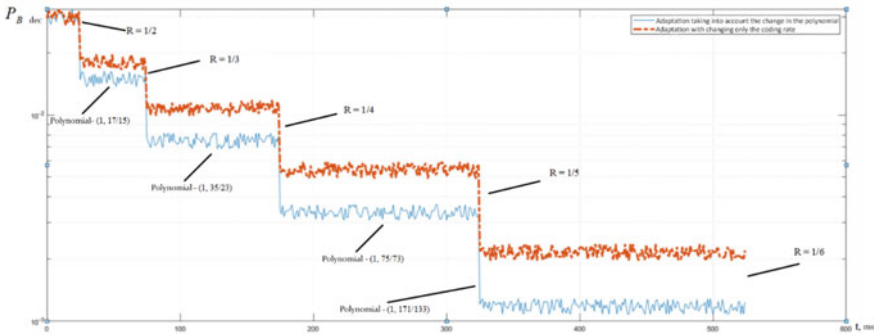
Fig. 4 Simulation results with an additional encoder polynomial change

changed from $R = 1/2$ to $R = 1/6$. The signal-to-noise ratio was chosen to be 0.4 dB. The average value of the decoding bit error probability $P_{B \text{ dec\_ pre}} = 10^{-3}$ is selected as the preset value of the information reliability.

In Fig. 4 is shown a graph obtained by means of simulation modeling, depending on $P_{B \text{ dec}}$ on the operating time of the TC codec in milliseconds when changing only the coding rate from $R = 1/2$ to $R = 1/6$ and a graph with changing the coding rate and additionally encoder polynomial.

The analysis of the simulation results shows an increase in the reliability of information transfer with an additional change in the TC polynomial, while the number of errors decreased by 0.2–0.7 times.

In Fig. 5 is shown a graph of the dependence of $P_{B \text{ dec}}$ on the operating time of the TC codec in milliseconds when only the coding rate is changed from $R = 1/2$ to $R = 1/6$ and a graph with the proposed parametric adaptation algorithm (the coding rate, the number of iterations and the interleaver are changed).

The analysis of the simulation results shows an increase in the reliability of information transfer using the proposed algorithm of multilevel parametric adaptation,
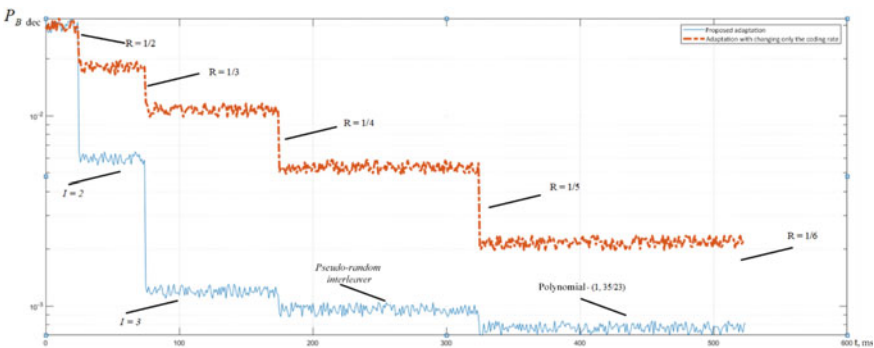
Fig. 5 Results of modeling the proposed method of multilevel parametric adaptation

while the number of errors has decreased by 2–3.3 times. This indicator is achieved by adding additional decoding iterations in the adaptation process, using more complex RSCE polynomials and changing the interleaving method.

6 Conclusions

1. The article proposes a method for multilevel parametric adaptation of the turbo code codec in wireless data transmission systems depending on the values of the normalized number of changes in the sign of a posteriori-a priori LLR about the transmitted data bits of the TC decoder.
2. The results of simulation showed that the application of the method of multilevel parametric adaptation of the turbo code codec reduces the number of errors by 2–3.3 times, and also makes it possible to increase the reliability of information transmission in comparison with the known results, for example, the fourth generation 4G LTE-Advanced mobile communication.

References

1. Hawwar Y, Farag E, Vanakayala S, Pauls R, Yang X, Subramanian S, Sadhanala P, Yang L, Wang B, Li Z, Chen H, Lu Z, Clark D, Fosket T, Mallela P, Shelton M, Laurens D, Salaun T, Gougeon L, Aubourg N, Morvan H, Henaff NL, Prat G, Charles F, Creach C, Calvez Y, Butel P (2006) 3G UMTS wireless system physical layer: baseband processing hardware implementation perspective. IEEE Commun Mag 44(9):52–58
2. Shen Z, Papasakellariou A, Montojo J, Gerstenberger D, Xu F (2012) Overview of 3GPP LTE-advanced carrier aggregation for 4G wireless communications. IEEE Commun Mag 50(2):122–130
3. Arora K, Singh J, Randhawa YS (2020) A survey on channel coding techniques for 5G wireless networks. Telecommun Syst 73:637–663
4. Berrou C, Glavieux A, Thitimajshima P (1993) Near Shannon limit error-correcting coding and decoding: turbo-codes. In: Proceedings of the international conference on communications, ICC-93, Geneva, May 1993, pp 1064–1070. https://doi.org/10.1109/ICC.1993.397441
5. Berrou C, Glavieux A (1996) Near optimum error correcting coding and decoding: turbo-codes. IEEE Trans Commun 44(10):1261–1271. https://doi.org/10.1109/26.539767
6. MacKay DJC, Neal RM (1996) "Near Shannon limit performance of low density parity check codes. Electron Lett 32(18):457–458
7. Westra JR, Mulder J, Ke Y, Vecchi D, Liu X, Arslan E, Wan J, Zhang Q, Wang S, van der Goes F, Bult K (2014) Design considerations for low-power analog front ends in full-duplex 10GBASE-T transceivers. In: Proceedings of the IEEE 2014 custom integrated circuits conference, Sept 2014, pp 1–8
8. Teo KH, Tao Z, Zhang J (2007) The mobile broadband WiMAX standard [standards in a Nutshell]. IEEE Signal Process Mag 24(5):144–148
9. IEEE 802.11-2012 Standard for Information Technology—Telecommunications and Information Exchange between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) (2012)

10. ETSI EN 302 307-1 Digital Video Broadcasting (DVB); Second Generation Framing Structure, Channel Coding and Modulation Systems for Broadcasting, Interactive Services, News Gathering and other Broadband Satellite Applications; Part 1: DVB-S2, V1.4.1 edn. (2014)
11. Arikan E (2009) Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. IEEE Trans Inform Theory 55(7):3051–3073
12. Dahlman E, Parkvall S, Skold J (2011) 4GLTE/LTE-advanced for mobile broadband. Academic, Oxford, 431 p
13. Sesia S, Toufik I, Baker M (2009) LTE—the UMTS long term evolution. From theory to practice. Wiley, West Sussex, 626 p
14. Hanzo L, Akhtman Y, Wang L (2011) MIMO-OFDM for LTE, WiFi and WiMax. Coherent versus non-coherent and cooperative turbo-transceivers. Wiley, New York, 658 p
15. Varda M, Badiu M, Bota V (2015) Link adaptation algorithm for distributed coded transmissions in cooperative OFDMA systems. Telecommunication systems. Springer, USA, pp 477–489
16. Ghazisaeidi A, Fernandez I, Schmalen L et al (2016) Submarine transmission systems using digital nonlinear compensation and adaptive-rate forward error correction. IEEE/OSA J Lightwave Technol 34(8):1886–1895
17. Sadjadpour HR, Sloane NJA, Salehi M, Nebe G (2006) Interleaver design for turbo codes. IEEE J Sel Areas Commun 19:831–837
18. Kumar S, Dalal H (2014) Performance comparison of turbo codes and modified turbo codes with different rate. Int J Sci Eng Technol Res (IJSETR) 3(5)
19. Wang S, Cui L, Cheng S, Zhai Y, Yeary M, Wu Q (2011) Noise adaptive LDPC decoding using particle filtering. IEEE Trans Commun 59(4):913–916
20. Mahalakshmi R, Bhuvaneshwari PV, Tharini C et al (2020) A novel algorithm to design rate-adaptive irregular LDPC codes. Wirel Pers Commun 113:453–468
21. Zaitsev SV, Kazymyr VV, Vasilenko VM, Yarilovets AV (2018) Adaptive selection of parameters of s-random interleaver in wireless data transmission systems with turbo coding. Radio-electronics and communications systems, vol 61. Allerton Press, Inc., New York, pp 13–21. https://doi.org/10.3103/S0735272715050039
22. Berrou C (2010) Codes and turbo codes. Springer, 415 p
23. Vucetic B, Yuan J (2000) Turbo codes. Principles and applications. Springer, 307 p

Neural Network Models Ensembles for Generalized Analysis of Audit Data Transformations

Tetiana Neskorodieva and **Eugene Fedorov**

Abstract We consider the problem of generalized analysis of audit data based on transformation models ensemble, which are based on artificial neural networks with associative memory and structural and parametric identification through clustering. This allows to increase the efficiency of the audit data analysis by reducing the computational complexity of structural and parametric identification (using the parallel processing technology CUDA and clustering in batch mode), increasing the accuracy of the analysis (using the adaptive structure of neural network models of transformation and joint clustering of training input and output data), expanding the capabilities of neural networks with associative memory (using a different number of hidden layers, corresponding to the stages of the analysis), and increasing the efficiency of the audit data analysis. The structure of the neural network model corresponds to the structural model of data transformations of the audit subject area. Developed software that implements the proposed method through the Matlab package. The developed software is used to solve the problem of studying the indicators and their relationships in the task of audit data analyzing.

Keywords Generalized analysis · Audit data · Mapping · Neural network · Associative memory

1 Introduction

Currently, the analytical procedures used during the audit are based on the methods of statistical [1] and data mining [2–5]. These methods are used to solve problems, to classify, to obtain estimated or predicted values of indicators in the tasks of auditing

T. Neskorodieva (✉) · E. Fedorov
Donetsk National University Named Vasyl Stus, 600-richchia str., 21, Vinnytsia 21021, Ukraine
e-mail: t.neskorodieva@donnu.edu.ua

E. Fedorov
e-mail: fedorovee75@ukr.net

E. Fedorov
Cherkasy State Technological University, Shevchenko blvd 460, Cherkasy 18006, Ukraine

at different levels and in relation to different objects of the audit. These methods can be used in automated audit systems, which are designed on the basis of intelligent technologies [6, 7]. In this case, it is necessary to reduce the computational complexity and increase the accuracy of the methods used in automatic mode.

The development of methods of estimation and prediction [8], clustering [9], formation of generalized associative relationships [10] are described in the works of the authors of this article. The goals of creating these methods: reducing the computational complexity for simple tasks (a single mapping of elements or sub-elements of the audit subject area), automatic structural identification, increasing the accuracy for complex tasks (compositions of mappings of elements or sub-elements of the audit subject area) and the possibility of applying these methods for the generalized analysis of elements and sub-elements of the audit subject area (see Table 1). Procedures used during the audit are based on the methods of statistical [1] and data mining [2–5]. These methods are used to solve problems, to classify, to obtain estimated or predicted values of indicators in the tasks of auditing at different levels and in relation to different objects of the audit. These methods can be used in automated audit systems, which are designed on the basis of intelligent technologies [6, 7]. In this case, it is necessary to reduce the computational complexity and increase the accuracy of the methods used in automatic mode.

The methods development of estimation and prediction [8], clustering [9], formation of generalized associative relationships [10] are described in the works of the authors of this article. The goals of creating these methods: reducing the computational complexity for simple tasks (a single mapping of elements or sub-elements of the audit subject area), automatic structural identification, increasing the accuracy for complex tasks (compositions of mappings of elements or sub-elements of the audit subject area) and the possibility of applying these methods for the generalized analysis of elements and sub-elements of the audit subject area (see Table 1).

Existing statistical methods have one or more of the following disadvantages, which make it difficult to automate the data analysis for audit objects with different characteristics:

- it is required to establish relationships between factors;
- the process of determining the structure of the model is not automated;
- assumptions about the distribution of factors are required, a priori information about the factors is required;
- input data should not be highly correlated, incomplete or noisy;
- difficult to analyze systems with a high degree of nonlinearity;
- the development of the model is carried out more slowly than in the case of the neural network approach;
- difficult to analyze systems with a large number of factors.

Therefore, a neural network approach was chosen in the work. The audit data transformation model can be implemented through neural networks with associative memory. Traditional neural networks with associative memory are:

1. Neural networks with heterassociative memory only [11–14].

Table 1 Comparative analysis of intelligent analysis methods in audit tasks (based on works [8–10])

The economic content of the display	Model of processing elements of the subject area, Features of the model or method	Purpose of processing elements of the subject area	Advantages and disadvantages of the model or method
Payment—delivery of raw material	Modified Liquid State. Machine, one-dimensional hidden layer, parameter identification based on matrix pseudo reversion	Evaluation and prediction of features of raw material supplies (by type) based on the values of payment features in a direct check of the display	Reducing computational complexity, and improving the forecast accuracy for simple tasks
Settlements with suppliers–customer settlements (a composition of mappings between a set of input and output data)	A neural network model based on a gateway recurrent unit. For parametric identification of this model is used adaptive cross entropy [8] (a combination of random and directional search)	Evaluation of features of settlements with customers on the basis of values of features of settlements with suppliers in a direct verification of mapping	Reducing computational complexity, improving the forecast accuracy for complex tasks This model is faster to learn but less accurate than in [8] because the pseudoreversal is not paralleled
Payment - delivery of raw materials	Spectral clustering [9]. Automatically determining the number of clusters based on the explained and sample variance rule; automatic determination of the scale parameter based on local scaling (the rule of K-nearest neighbors is used)	Formation of the appropriate number of sub-elements of the subject area of the audit in the elements that are mapped	Automatically determining the number of clusters, automatic determination of the scale parameter, resistance to noise and random outliers
Settlements with suppliers—settlements with customers (a composition of mappings)	Forward-only counterpropagating neural network, which is a nonrecurrent static two-layer ANN [10], assumed that the audit features are noisy with Gaussian noise	Construction of generalized associative relationships for generalized analysis tasks	Automating the formation of generalized features of audit sets and their mapping by means of a forward-only counterpropagating neural network. The number of pairs is set manually

2. Neural networks with autoassociative memory only [15–24].
3. ANNs with heterassociative and autoassociative memory [25–27].

Most associative memory neural networks have one or more disadvantages:

1. not used for both the original and the other sample reconstruction;
2. do not work with real data;
3. do not have a high capacity of associative memory.

In this regard, it is relevant to create a neural network, which will eliminate these disadvantages. Clustering techniques can be used to summarize audit data and train neural networks.

1. Partition-based method or center-based [28].
2. Model mixture or distribution-based or model-based methods [29].
3. Density-based methods [30].
4. Methods are hierarchical [30]: agglomerative, divisive or top-down.
5. Spectral methods [30].

Existing clustering methods have one or more of the following disadvantages:

- have high computational complexity;
- do not allow noise and accidental emissions to stand out;
- clusters cannot have different shapes and sizes;
- require the number of clusters to be set;
- require the determination of parameter values.

In this regard, it is relevant to create a method of clustering, which will eliminate these disadvantages.

The aim of the work is to improve the efficiency of generalized data analysis in SPPD auditing based on an ensemble of neural network transformation models.

To achieve the goal it is necessary to solve the following tasks:

- propose a structural model for transforming the data of the audit subject;
- to propose an ensemble of neural network transformation models;
- choose a criterion for evaluating the effectiveness of an ensemble of neural network transformation models;
- propose a method for structural and parametric identification of an ensemble of neural network models of transformations;
- perform numerical studies.

2   Problem Statement

The problem of increasing the efficiency of auditing by ensemble of neural network models of transformation $\big(g_1(S, \Theta_1), \ldots, g_K(S, \Theta_K)\big)$, where S-set of input data, Θ_k-set of parameters of k-th neural network model of transformations, is represented

as the problem of finding such vector of parameters $(\Theta_1^*, \ldots, \Theta_K^*)$, which satisfies the criterion

$$\Phi = \frac{1}{I} \sum_{i=1}^{I} \min_{k \in 1, K} g_k \left(S_i, \Theta_k^* \right) \to \min. \tag{1}$$

3   Structural Model of Subject Area Sub-elements Transformations of the "Completeness" Prerequisite Audit

Verification of prerequisite "Completeness" assumption consists in checking the consistency of the sequential operations sets $on_{i_1} \prec \ldots \prec on_{i_m} \prec \ldots \prec on_{i_M}$. These sets define the sub-elements of the audit subject area. Therefore, the interrelationships of the area subject sub-elements, which are subject to the procedure of checking the "Completeness" prerequisite, are presented as sequences of mappings of these sequential operations sets.

$$Z_{i_1} \to Z_{i_2} \to Z_{i_m} \ldots \to Z_{i_M}, \; i_m = \overline{1, I}, \; (i_1, \ldots i_m, \ldots i_M) \in A\left(\tilde{I} \right), \; \tilde{I} = \overline{1, I}, \tag{2}$$

where

$(i_1, \ldots i_m, \ldots i_M)$ is a combination of types of sequential operations of the set $\tilde{I} = \overline{1, I}$,
$on_{i_1} \prec \ldots \prec on_{i_m} \prec \ldots \prec on_{i_M}$ is the sequential operations,
$A\left(\tilde{I} \right)$ is a set of possible combinations on the set $\tilde{I} = \overline{1, I}$.

The set of possible combinations of sequential operations types $(i_1, \ldots i_m, \ldots i_M)$ defined in (2) includes checks in the forward and reverse direction, complete and incomplete.

The interrelationships of the lower-level domain sub-elements, which are subject to the procedure of checking the "Completeness" prerequisite, will be represented as a subsets mappings sequence of sequential operations:

$$\bigcup_{l_{i_1}=1}^{L_{i_1}} Z(l_{i_1}) \to \ldots \to \bigcup_{l_{i_m}=1}^{L_{i_m}} Z(l_{i_m}) \to \ldots \to \bigcup_{l_{i_M}=1}^{L_{i_M}} Z(l_{i_M}),$$

where
l and L are the number and quantity of sets, respectively.

Let's form a structural model of subelements transformation of the audit subject area of the "Completeness" premise by the example of direct material costs audit (see Fig. 1). This model can be represented as a graph, in which each graph vertex corresponds to a subelement, and an edge—a mapping that describes the relationship between the corresponding subelements.

To do this, we use the formalization of the direct material costs set in the form of a graph $G^{(1)} = (Z^{(1)}, R^{(1)})$, where the vertices are the accounts in which these current assets are accounted, and the edges are the operations that result in their transformation. Then the model of transformation of the subject area subelements of the prerequisite audit "Completeness" for direct full verification $((i_1, \ldots i_m, \ldots i_M) = (1, 2, 3, 4))$ is a mapping of the data subsets of settlements with suppliers $Z_{\Re}^{(r_1)}(i_1) \in \Re(Z_{i_1})$ $\Re$-algebra of sets into data subsets of the raw materials stocks $Z_{\Re}^{(r_2)}(i_2) \in \Re(Z_{i_2})$, then into data subsets of the production $Z_{\Re}^{(r_3)}(i_3) \in \Re(Z_{i_3})$ and finished products $Z_{\Re}^{(r_4)}(i_4) \in \Re(Z_{i_4})$, $T \in \{t_{j_m}, T_m, j = \overline{1, J_m}, m = \overline{1, M}, T\}$.

In the particular case, if the partitioning of the sets was carried out on the basis of logical conditions, characterizing the belonging to one of the subspecies of the accounting object, then the model of the subject area sub-elements transformation of audit prerequisite "Completeness" at direct full check is the mappings sequence of the settlement operations data sets by supplier type into subsets of operations data by type of raw material, then into subsets of operations data by production type and finished products type.
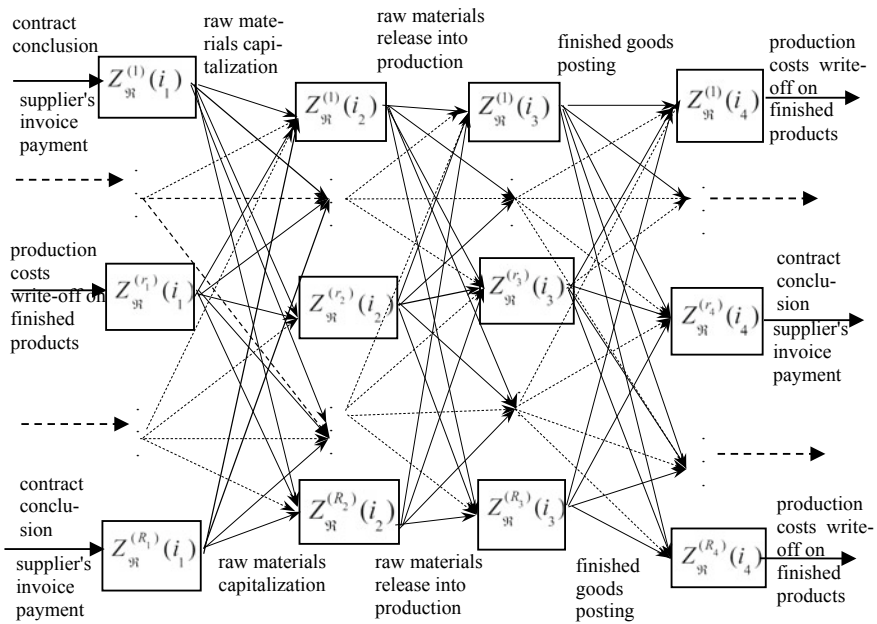
Fig. 1 Structural model of transformations of sub-elements of the subject area

In the generalized analysis in the audit system it is necessary to form a generalized relationship between the sub-elements of the audit subject area. For this purpose, on the basis of the proposed transformations structural model (see Fig. 1) the corresponding transformations neural network model is developed.

4 Ensemble of Transformation Neural Network Models

Figure 2 shows the structure of transformations neural network model, which is based on the proposed transformation model and contains an input layer, one or more hidden layers and an output layer. Each hidden layer corresponds to a step in the transformation model. As in the transformation model, input and hidden layer data are fed to all layers.

This paper proposes to use an ensemble of neural network transformation models with the parameters of each neural network transformation model identified based on its subset of training data.

Let the set of input, output and structural data be represented as

$$S = \left(\mathbf{x}, \mathbf{y}, \mathbf{V}^{(0,1)}, \ldots, \mathbf{V}^{(L-1,L)}\right).$$

Let the set of parameters k of the k-th neural network transformations mode be represented as $\Theta_k = \left(\mathbf{m}_k^{(0)}, \mathbf{m}_k^{(L)}, \mathbf{W}_k^{(0,1)}, \ldots \mathbf{W}_k^{(L-1,L)}\right).$

Then the selection of the neural network transformations model number from the ensemble for the specified data a S is performed.

$$k^* = \arg \min_{k \in 1, K} g_k(S, \Theta_k)$$
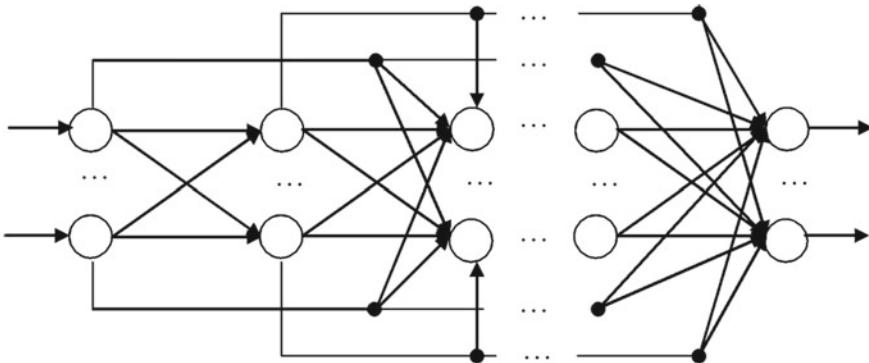
Fig. 2 Structure of the neural network model of transformations

$$= \arg \min_{k \in 1, K} \left(\left\| \mathbf{x} - \mathbf{m}_k^{(0)} \right\|^2 + \left\| \mathbf{y} - \mathbf{m}_k^{(L)} \right\|^2 + \sum_{l=1}^{L} \left\| \mathbf{V}^{(l-1,l)} - \mathbf{W}_k^{(l-1,l)} \right\|^2 \right),$$

where

$\mathbf{x}$ is the input vector of the transformation model, $\mathbf{x} = (x_1, \dots, x_{N^{(0)}})$,

$\mathbf{y}$ is the transformation model output vector, $\mathbf{y} = (y_1, \dots, y_{N^{(L)}})$,

$\mathbf{m}_k^{(0)}$ is the k-th input vector cluster centroid $\mathbf{m}_k^{(0)} = \left(m_{k1}^{(0)}, \dots, m_{kN^{(0)}}^{(L)} \right)$,

$\mathbf{m}_k^{(L)}$ is the k-th output vector cluster centroid, $\mathbf{m}_k^{(L)} = \left(m_{k1}^{(L)}, \dots, m_{kN^{(L)}}^{(L)} \right)$,

$\mathbf{V}^{(l-1,l)}$ is associative links weights matrix between the $(l-1)$-th and l-th stages of the transformation model, $\mathbf{V}^{(l-1,l)} = \left[v_{cu}^{(l-1,l)} \right] c \in \overline{1, N^{(l-1)}}, \ u \in \overline{1, N^{(l)}}$,

$\mathbf{W}_k^{(l-1,l)}$ is the k-th centroid of the associative links weights cluster between the $(l-1)$-th and l-th stages, $\mathbf{W}_k^{(l-1,l)} = \left[w_{icu}^{(l-1,l)} \right] c \in \overline{1, N^{(l-1)}}, \ u \in \overline{1, N^{(l)}}$.

$N^{(l)}$ is the elements quantity in the l-stage of the transformation model,

L is the quantity of the transformation model stages,

K is quantity of clusters (corresponds to the number of neural network transformation models).

Losses obtained as a result of selecting a neural network transformations model from the ensemble for the specified data S, calculated as

$$loss = \min_{k \in 1, K} g_k(S, \Theta_k)$$

$$= \min_{k \in 1, K} \left(\left\| \mathbf{x} - \mathbf{m}_k^{(0)} \right\|^2 + \left\| \mathbf{y} - \mathbf{m}_k^{(L)} \right\|^2 + \sum_{l=1}^{L} \left\| \mathbf{V}^{(l-1,l)} - \mathbf{W}_k^{(l-1,l)} \right\|^2 \right).$$

5   Choosing the Criterion for Evaluating the Effectiveness of the Neural Network Transformations Model

In this paper, to identify the parameters of the neural network transformations model, the loss function is selected, which means the choice of such a clustering function $\phi : \{1, \dots, I\} \to \{1, \dots, K\}$, which delivers a minimum of intra-cluster sums of squares

$$F(\phi) = WCSS^{(0)}(\phi) + WCSS^{(L)}(\phi) + \sum_{l=1}^{I} WCSS^{(l-1,l)}(\phi) \to \min_{\phi},$$

$$WCSS^{(0)}(\phi) = \sum_{i=1}^{I} \sum_{k=1}^{K} [\phi(i) = k] \left\| \mathbf{x}_i - \mathbf{m}_k^{(0)} \right\|^2,$$

$$WCSS^{(L)}(\phi) = \sum_{i=1}^{I} \sum_{k=1}^{K} [\phi(i) = k] \left\| \mathbf{y}_i - \mathbf{m}_k^{(L)} \right\|^2,$$

$$WCSS^{(l-1,l)}(\phi) = \sum_{i=1}^{I} \sum_{k=1}^{K} [\phi(i) = k] \left\| \mathbf{V}^{(l-1,l)} - \mathbf{W}_k^{(l-1,l)} \right\|^2,$$

$$[\phi(i) = k] = \begin{cases} 1, \phi(i) = k \\ 0, \phi(i) \neq k \end{cases},$$

where

$\mathbf{x}_i$ is the i-th transformation model input vector, $\mathbf{x}_i = (x_{i1}, \ldots, x_{iN^{(0)}})$,

$\mathbf{y}_i$ is the i-th transformation model output vector, $\mathbf{y}_i = (y_{i1}, \ldots, y_{iN^{(L)}})$,

$\mathbf{m}_k^{(0)}$ is the k-th centroid of the input vector cluster, $\mathbf{m}_k^{(0)} = \left(m_{k1}^{(0)}, \ldots, m_{kN^{(0)}}^{(L)} \right)$,

$\mathbf{m}_k^{(L)}$ is the k-th centroid of the output vector cluster, $\mathbf{m}_k^{(L)} = \left(m_{k1}^{(L)}, \ldots, m_{kN^{(L)}}^{(L)} \right)$,

$\mathbf{V}_i^{(l-1,l)}$ is the i-th matrix of associative links weights between the $(l-1)$-th and l-th stages of the transformation model, $\mathbf{V}_i^{(l-1,l)} = \left[v_{icu}^{(l-1,l)} \right]$, $c \in \overline{1, N^{(l-1)}}$, $u \in \overline{1, N^{(l)}}$,

$\mathbf{W}_k^{(l-1,l)}$ is the k-th centroid of the cluster of associative links weights between the $(l-1)$ and l stages, $\mathbf{W}_k^{(l-1,l)} = \left[w_{icu}^{(l-1,l)} \right]$, $c \in \overline{1, N^{(l-1)}}$, $u \in \overline{1, N^{(l)}}$.

$N^{(l)}$ is the quantity of elements in the l-th stage of the transformation model,

L is the quantity of stages of the transformation model,

K is quantity of clusters,

I is the set power.

6 Method for Structural and Parametric Identification of a Neural Network Transformations Model in Batch Mode

1. The initial number of clusters is set $K = 2$. The maximum number of clusters is set $K^{\max}$, $K^{\max} < I$.
2. Clustering function calculation and centroids of input and output vector clusters.
2.1 A training set is set $\{(\mathbf{x}_i, \mathbf{y}_i)\}$ $i \in \overline{1, I}$. The threshold for stopping is set ε_1. The initial partitioning of the training set of input and output vectors into

$K \{(x_i, y_i)\}$ clusters, which is described by the clustering function ϕ, is set randomly.

2.2   Set iteration number $n = 1$. Specifies the maximum iterations quantity $N^{\max}$.

2.3   Calculation of the current sum of squares intra-cluster sums

$$z_1 = \begin{cases} 0, & n = 1 \\ WCSS^{(0)}(\phi) + WCSS^{(L)}(\phi), & n > 1 \end{cases}$$

2.4   Input vector cluster centroids calculation based on the k-means rule

$$m_{kj}^{(0)} = \frac{\sum_{i=1}^{I}[\phi(i) = k]x_{ij}}{\sum_{i=1}^{I}[\phi(i) = k]}, \ k \in \overline{1, K}, \ j \in \overline{1, N^{(0)}}.$$

2.5   Output vector cluster centroids calculation based on the k-means rule

$$m_{kj}^{(L)} = \frac{\sum_{i=1}^{I}[\phi(i) = k]y_{ij}}{\sum_{i=1}^{I}[\phi(i) = k]}, k \in \overline{1, K}, \ j \in \overline{1, N^{(L)}}.$$

2.6   Calculation of distance squares between vectors and cluster centroids

$$D_{ik} = \left\| \mathbf{x}_i - \mathbf{m}_k^{(0)} \right\|^2 + \left\| \mathbf{y}_i - \mathbf{m}_k^{(L)} \right\|^2, \ i \in \overline{1, I}, \ k \in \overline{1, K}.$$

2.7   Clustering function values calculation

$$\phi(i) = \arg\min_k D_{ik}, \ i \in \overline{1, I}, \ k \in \overline{1, K}.$$

2.8   Calculation of the intra-cluster squares sum in the case of input vectors

$$WCSS^{(0)}(\phi) = \sum_{i=1}^{I}\sum_{k=1}^{K}[\phi(i) = k]\left\| \mathbf{x}_i - \mathbf{m}_k^{(0)} \right\|^2.$$

2.9   Calculation of the intra-cluster squares sum in the case of output vectors

$$WCSS^{(L)}(\phi) = \sum_{i=1}^{I}\sum_{k=1}^{K}[\phi(i) = k]\left\| \mathbf{y}_i - \mathbf{m}_k^{(L)} \right\|^2.$$

2.10   Condition of completion.
       If $\left| WCSS^{(0)}(\phi) + WCSS^{(L)}(\phi) - z_1 \right| > \varepsilon_1$, and $n < N^{\max}$, then increase the iteration number, i.e. $n = n + 1$, go to step 2.3.

3   The centroids calculation of cluster weights of associative links between adjacent stages.

3.1   The training set of weights matrices of associative relations between transformations steps $\{\mathbf{V}_i^{(l-1,l)}\}$, $i \in \overline{1, I}$, $l \in \overline{1, L}$, is set a threshold for stopping ε_2.

3.2   The current sum calculation of intra-cluster squares sums

$$
z_2 = \begin{cases} \sum\limits_{l=1}^{L} \sum\limits_{i=1}^{I} \sum\limits_{c=1}^{C} \sum\limits_{u=1}^{U} \left(v_{icu}^{(l)} - \frac{1}{I} \sum\limits_{z=1}^{I} v_{zcu}^{(l)} \right)^2, & K = 2 \\ \sum\limits_{l=1}^{L} WCSS^{(l-1,l)}(\phi), & K > 2 \end{cases}
$$

3.3   Weights cluster centroids calculation of associative links between neighboring stages based on the k-means rule

$$
w_{kcu}^{(l-1,l)} = \frac{\sum_{i=1}^{I}[\phi(i) = k]v_{icu}^{(l)}}{\sum_{i=1}^{I}[\phi(i) = k]}, \ k \in \overline{1, K}, \ l \in \overline{1, L}, \ c \in \overline{1, N^{(l-1)}}, \ u \in \overline{1, N^{(l)}}.
$$

Consider that $w_{kcu}^{(l-1,s)} = w_{kcu}^{(l-1,l)}$, $s \in \overline{l+1, L}$.

3.4   Calculation of intra-cluster squares sums in the case of associative links weights between neighboring stages

$$
WCSS^{(l-1,l)}(\phi) = \sum_{i=1}^{I} \sum_{k=1}^{K} [\phi(i) = k] \left\| \mathbf{V}_i^{(l-1,l)} - \mathbf{W}_k^{(l-1,l)} \right\|^2.
$$

3.5   Completion Condition

If $\left| \sum_{L=1}^{L} WCSS^{(l-1,l)}(\phi) - z_1 \right| > \varepsilon_2$ and $K > K^{\max}$, then increase the clusters quantity, i.e. $K = K + 1$, go to step 2.

7 Structural and Parametric Identification Algorithm of a Neural Network Transformations Model in Batch Mode for GPU Implementation

For the proposed method of structural and parametric identification of the neural network transformations model on the example of an audit we consider an algorithm designed for implementation on GPU through the parallel processing technology CUDA (see Fig. 3).

Step 1   Set the maximum clusters quantity $K^{\max}$. Initialize the initial quantity of clusters K.

Step 2   Set the training set $\{(\mathbf{x}_i, \mathbf{y}_i)\}$, $i \in \overline{1, I}$, the threshold for stopping ε_1. Initialize the values of the clustering function φ.

Step 3   Initialize iteration number n. Initialize the maximum number of iterations $N^{\max}$.

Fig. 3 Block diagram of the structural and parametric identification algorithm of the neural network model of transformations in the batch mode
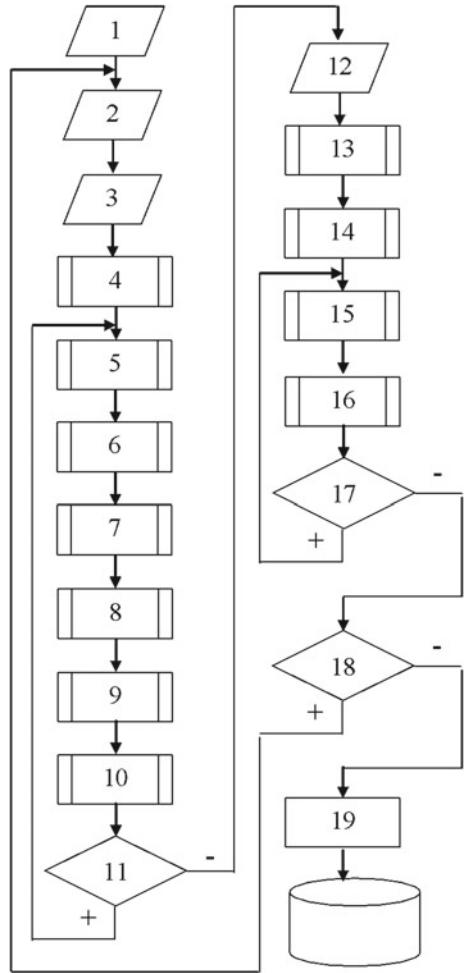
Step 4   Calculation of the current sum of intra-cluster squares sums z_1.

Step 5   Calculation of input vector clusters centroids based on reduction, using $K \cdot N^{(0)} \cdot I$ GPU threads that are grouped into $K \cdot N^{(0)}$ blocks. The result of each block is the centroid $m_{kj}^{(0)}$.

Step 6   Calculation of output vector cluster centroids based on the k-means rule, using $K \cdot N^{(L)} \cdot I$ GPU threads that are grouped into $K \cdot N^{(L)}$ blocks. The result of each block is the centroid $m_{kj}^{(L)}$.

Step 7   Calculate the squares of distances between vectors and centroids of clusters based on reduction, using $I \cdot K \cdot \left(N^{(0)} + N^{(L)}\right)$ GPU threads that are grouped into $I \cdot K$ blocks. The result of each block is the distance square D_{ik}.

Step 8   Calculate clustering function values based on reduction, using $I \cdot K$ GPU threads that are grouped into $I^{\phi(i)}$.

Step 9   Calculation of intra-cluster squares sum in the case of input vectors. First, on the basis of reduction, using $I \cdot K \cdot N^{(0)}$ GPU threads which are grouped in $I \cdot K$ blocks, in each block the product of $d_{ik} = [\phi(i) = k]\left\|\mathbf{x} - \mathbf{m}_k^{(0)}\right\|^2$. Then on the basis of reduction, using $I \cdot K$ GPU threads that are grouped into I blocks, the sum of $d_i = \sum_{k=1}^{K} d_{ik}$. Then on the basis of reduction, using GPU threads I that are grouped in 1 block, in a block the $WCSS^{(0)}(\phi) = \sum_{i=1}^{I} d_i$.

Step 10   Calculation of intra-cluster squares sum in the case of output vectors.

First, on the basis of reduction, using $I \cdot K \cdot N^{(L)}$ GPU threads that are grouped in $I \cdot K$ blocks, in each block the product of $d_{ik} = [\phi(i) = k]\left\|\mathbf{y}_i - \mathbf{m}_k^{(L)}\right\|^2$. Then on the basis of reduction, using $I \cdot K$ GPU threads that are grouped into I blocks, the sum of $d_i = \sum_{k=1}^{K} d_{ik}$. Then on the basis of reduction, using GPU threads I which are grouped in 1 block, the $WCSS^{(0)}(\phi) = \sum_{i=1}^{I} d_i$.

Step 11   If the completion condition is satisfied, i.e., $|WCSS^{(0)}(\phi) + WCSS^{(L)}(\phi) - z_1| > \varepsilon_1$, and $n < N^{\max}$, then increase the iteration number, i.e. $n = n + 1$, go to step 4.

Step 12   Operator input of the training set $\left\{V_i^{(l-1,l)}\right\}$, $i \in \overline{1, I}$, $l \in \overline{1, L}$. Initialization of the threshold for stopping ε_2.

Step 13   Calculating the current sum of intra-cluster squares sums z_2.

Step 14   Initialize stage number l.

Step 15   Compute centroids of cluster weights of associative links between adjacent steps based on reduction, using $K \cdot N^{(l-1)} \cdot N^{(l)} \cdot I$ GPU threads that are grouped into $K \cdot N^{(l-1)} \cdot N^{(l)}$ blocks. Each thread calculates the centroid $w_{kcu}^{(l-1,l)}$.

Step 16   Calculation of intra-cluster squares sums in the case of associative links weights between adjacent steps. First, on the basis of reduction, using $I \cdot K \cdot N^{(l-1)} \cdot N^{(l)}$ GPU threads which are grouped in $I \cdot K$ blocks, in each block the product of $d_{ik}^{(l-1,l)} = [\phi(i) = k]\left\|\mathbf{V}_i^{(l-1,l)} - \mathbf{W}_k^{(l-1,l)}\right\|^2$. Then on the basis of reduction, using $I \cdot K$ GPU threads that are grouped into I blocks, in each block the sum of $d_i^{(l-1,l)} = \sum_{i=1}^{I} d_{ik}^{(l-1,l)}$. Then on the basis of reduction, using GPU threads I which are grouped in 1 block, the $WCSS^{(l-1,l)}(\phi) = \sum_{i=1}^{I} d_i^{(l-1,l)}$.

Step 17   If $l < L$, then increase the step number, i.e. $l = l + 1$, go to step 15.

Step 18   If the completion condition is satisfied, i.e., $\sum_{l=1}^{L} WCSS^{(l-1,l)}(\phi) - z_2 > \varepsilon_2$, and $K > K^{\max}$, increase the number of clusters, i.e. $K = K + 1$, go to step 2.

Step 19   Write the obtained centroids $\mathbf{m}_k^{(0)}$, $\mathbf{m}_k^{(L)}$, $\mathbf{W}_k^{(l-1,l)}$ to the database.

8 Experiments and Results

The proposed method of audit data analysis based on an ensemble of neural network transformation models was implemented through the Matlab package. The following constraints were imposed on the dimensions of the data:

$$N^{(l-1)} \cdot N^{(2)} \leq 1024, \ K \leq 1024, \ I \leq 1024.$$

The results of comparison of the proposed method using GPU with the traditional method of k-means clustering for training of neural network models with associative memory (on the example of manufacturing enterprise data) are presented in Table 2.

Evaluation of computational complexity of the proposed method using GPU and traditional clustering method was based on the number of distance computations, the computation of which is the most costly place of the method. At that N^{max}—the maximum number of iterations, $N^{(0)}$—the number of neurons in the input layer, K—the number of clusters, I—the power of the training set.

The traditional clustering method for training neural network models with associative memory usually does not provide structural identification (automatic determination of the number of clusters), and joint clustering of training input and output data, which reduces the analysis accuracy (see Table 2).

Traditional clustering method for training neural network models with associative memory usually does not provide batch mode (allows parallel implementation), which increases computational complexity (see Table 2).

The traditional clustering method for training neural network models with associative memory usually does not provide more than one hidden layer, which narrows the scope of their applicability to the analysis of audit data (see Table 2).

The proposed method makes it possible to eliminate these drawbacks.

Table 2 Comparison of the proposed and traditional clustering method

Indicator	The method	
	Proposed	Traditional
Analysis accuracy	0.95	0.8
Computational complexity	$O\left(\log_2\left(I \cdot K \cdot N^{(0)}\right)N^{max}\right)$	$O\left(I \cdot K \cdot N^{(0)} \cdot N^{max}\right)$
Automatic detection of the number of clusters	+	−
Joint clustering of input and output data	+	−
Batch training mode	+	−
The presence of more than one hidden layer	+	−

9 Conclusions

The actual task of increasing the efficiency of audit automation at the stage of data analysis was solved by transforming and summarizing the audit indicators data by means of an ensemble of neural network models with associative memory.

1. A structural model of audit data transformation is proposed, based on which a neural network model for automatic formation of generalized patterns in checking the "Completeness" premise is formed.
2. To improve the training speed of the proposed ensemble of neural network transformation models, a method of structural and parametric identification in batch mode based on the k-means rule was developed.
3. The proposed method made it possible to achieve analysis accuracy of 0.95 due to the adaptive structure of neural network transformation models and joint clustering of training input and output data.
4. We propose to extend the capabilities of neural networks with associative memory by using a different number of hidden layers corresponding to the stages of transformations.
5. The algorithm of structural and parametric identification based on k-means, intended for implementation on GPU by means of CUDA technology, is created.
6. The proposed method of structural and parametric identification based on the rule of k-means can be used for the intellectualization of SPR audit.

Prospects for future research is to study the proposed method for a wide class of artificial intelligence tasks, as well as to create a method of data transformation and generalization to solve the problems of auditing.

References

1. Priyanka D, Priyanka JV, Papa Rao S (2020) Statistical analysis of various measures in auditing practices using optimization techniques. Sci Technol Dev 9(6):37–87. http://journalstd.com/gallery/6-june2020.pdf
2. Schultz M, Tropmann-Frick M (2020) Autoencoder neural networks versus external auditors: detecting unusual journal entries in financial statement audits. In: Proceedings of the 53-rd Hawaii international conference on system sciences, Hawaii, USA, 7–10 Jan, pp 5421–5430. https://doi.org/10.24251/HICSS.2020.666. http://hdl.handle.net/10125/64408
3. Nonnenmacher J, Kruse F, Schumann G, Gómez JM (2021) Using autoencoders for data-driven analysis in internal auditing. In: Proceedings of the 54th Hawaii international conference on system sciences (HICSS 2021), Hawaii, USA, 5–8 Jan, pp 5748–5757. https://doi.org/10.24251/HICSS.2021.697. http://hdl.handle.net/10125/71317
4. Bodyanskiy Y, Boiko O, Zaychenko Y, Hamidov G , Zelikman A (2020) The hybrid GMDH-Neo-fuzzy neural network in forecasting problems in financial sphere. In: Proceedings of 2nd international conference on system analysis and intelligent computing (SAIC), Kyiv, Ukraine. IEEE, pp 1–6. https://doi.org/10.1109/SAIC51296.2020.9239152
5. Crajaa P, Kima A, Lessmanna S (2020) Deep Learning application for fraud detection in financial statements. Discussion paper, No. 2020-007, Humboldt-Universität zu Berlin, IRTG 1792 "High dimensional nonstationary time series", Berlin

6. Dai J, Vasarhelyi MA (2020) Continuous audit intelligence as a service (CAIaaS) and intelligent app recommendations. J Emerg Technol Account 17(2):1–15. https://doi.org/10.2308/jeta-10751

7. Barmak AV, Krak YV, Manziuk EA, Kasianiuk VS (2019) Information technology separating hyperplanes synthesis for linear classifiers. J Autom Inf Sci 51(5):54–64. https://doi.org/10.1615/JAutomatInfScien.v51.i5.50

8. Neskorodieva T, Fedorov E, Izonin I (2020) Forecast method for audit data analysis by modified liquid state machine. In: Proceedings of the 1st international workshop on intelligent information technologies and systems of information security (IntelITSIS 2020), Khmelnytskyi, Ukraine, 10–12 June 2020: proceedings—CEUR-WS, vol 2623, pp 25–35. http://ceur-ws.org/Vol-2623/paper3.pdf

9. Neskorodieva T, Fedorov E (2020) Automatic analysis method of audit data based on neural networks mapping. In: VII International conference IT&I "Information technology and interaction", 2–4 Dec, Kyiv, pp 36–40. http://eportfolio.kubg.edu.ua/data/conference/6329/document.pdf

10. Neskorodieva T, Fedorov E (2021) Method of spectral clustering of payments and raw materials supply for the compliance audit planning. Radioelectron Inform Manage 1:127–135. https://doi.org/10.15588/1607-3274-2021-1-13

11. Singh UP, Jain S, Tiwari A, Singh RK (2019) Gradient evolution-based counter propagation network for approximation of noncanonical system. Soft Comput 23:4955–4967. https://doi.org/10.1007/s00500-018-3160-7

12. Cho JW, Park H-M (2016) Independent vector analysis followed by hmm-based feature enhancement for robust speech recognition. Sig Process:200–208

13. Bartlett MS, Movellan JR, Sejnowski TJ (2002) Face recognition by independent component analysis. IEEE Trans Neural Netw 13(6):1450–1464

14. Achmad MI, Adinugroho H, Susanto A (2014) Cerebellar model articulation controller (CMAC) for sequential images coding. In: The 1st international conference on information technology, computer, and electrical engineering. https://doi.org/10.1109/ICITACEE.2014.7065734

15. Neal RM (1992) Connectionist learning of belief networks. Artif Intell 56:71–113

16. Dayan P, Hinton GE, Neal RM, Zemel RS (1995) The Helmholtz machine. Neural Netw 7:889–904

17. Kohonen T (1995) Self-organizing maps. Springer-Verlag, Berlin

18. Haykin S (2006) Neural networks: a complete course. Williams Publishing House, Moscow

19. Chiueh TD, Goodman RM (1991) Recurrent correlation associative memories. IEEE Trans Neural Netw 2(2):275–284

20. Kobayashi M (2017) Quaternionic Hopfield neural networks with twin-multistate activation function. Neurocomputing 267:304–310. https://doi.org/10.1016/j.neucom.2017.06.013

21. Du KL, Swamy MNS (2014) Neural networks and statistical learning. Springer Verlag, London.https://doi.org/10.1007/978-1-4471-5571-3

22. Park Y (2010) Optimal and robust design of brain-state-in-a-box neural associative memories. Neural Netw 23(2):210–218. https://doi.org/10.1016/j.neunet.2009.10.008

23. Khristodulo OI, Makhmutova AA, Sazonova TV (2017) Use algorithm based at Hamming neural network method for natural objects classification. Procedia Comp Sci 103:388–395. https://doi.org/10.1016/j.procs.2017.01.126

24. Barszcz T, Bielecki A, Wójcik M (2020) ART-2 artificial neural networks applications for classification of vibration signals and operational states of wind turbines for intelligent monitoring. In: Advances in condition monitoring of machinery in non-stationary operations, pp 679–688. https://doi.org/10.1007/978-3-642-39348-8_58

25. Sivanandam SN, Deepa SN (2006) Introduction to neural networks using Matlab 6.0. The McGraw-Hill Comp. Inc., New Delhi

26. Javidmanesh E (2017) Global stability and bifurcation in delayed bidirectional associative memory neural networks with an arbitrary number of neurons. J Dyn Sys Meas Control 139(8). https://doi.org/10.1115/1.4036229

27. Fischer A, Igel C (2014) Training restricted Boltzmann machines: an introduction. Pattern Recogn 47:25–39
28. Brusco MJ, Shireman E, Steinley D (2017) A comparison of latent class, K-means, and K-median methods for clustering dichotomous data. Psychol Methods 22(3):563–580. https://doi.org/10.1037/met0000095
29. Fu Z, Wang L (2012) Color image segmentation using Gaussian mixture model and EM algorithm. Multi Signal Process:61–66. https://doi.org/10.1007/978-3-642-35286-7_9
30. Aggarwal CC, Reddy CK (2014) Data clustering: algorithms and applications. CRC Press, Boca Raton, FL

Method of Adaptive Error Control Coding of Structured Messages Using Parameterization of Reed-Solomon Codes

Roman Andrushchenko and **Andrii Rohovenko**

Abstract More and more business processes are going online, and special information services, APIs, SDKs are created that interact with people and other information systems. In the course of their operation, the TCP/IP network stack protocols are mostly used for communication. The wireless communication devices, mobile devices: tablets, smartphones, etc. are also becoming increasingly popular. All this leads to the need to pay more attention to the quality and security of data transmission over wireless networks, as well as to improve methods of detecting and correcting of errors that occur during data reception and transmission. The article describes the method of error correction in data packets based on the structure analysis of messages, analyses possible ways of implementation and analyses its effectiveness in combination with Reed-Solomon codes. The proposed method will improve data transmission processes over network protocols.

Keywords HTTP · JSON · Data serialization · Error control coding · Reed-Solomon codes · Computer network · Network software

1 Introduction

1.1 Target Settings

Information systems have been increasingly used in various activity spheres in recent decades, and in the last few years, the special attention has been paid to wireless communications and the Internet of Things [1], which require special attention to the reliability and security of data channels used by nodes of information systems for interaction and communication with each other and with other information systems. The number of devices connected to the Internet is growing at an incredible rate,

R. Andrushchenko (✉) · A. Rohovenko
Chernihiv Polytechnic National University, 95 Shevchenko Str., Chernihiv 14035, Ukraine
e-mail: arbamor@ukr.net

A. Rohovenko
e-mail: arogovenko@gmail.com

which is the main reason for the development of IPv6 protocol in communications due to the need to expand the address space and ensure growth. A lot of Internet of Things devices connected to the global network appeared in a last few years. The combination of these factors results in the need to research, develop and improve the data transmission methods in telecommunication systems [2, 3].

The error detection and correction process is typically implemented as a part of low-level protocols that consider transmitted data as a stream or blocks of bits/bytes, without taking into account the high-level structure and semantics of messages [4, 5].

So, some of the piece of information that could theoretically be used for more successful data decoding is ignored. This approach is generally proven and mostly correct, as the protocols are low-level and should not consider the data transmitted. They operate with a set of bits only. However, all these protocols do not always cope with the tasks assigned to them, especially in the case of interference, noise that occurs during the transmission process. Therefore, it is advisable, if possible, to use all appropriate ways to improve the process of detecting and correcting of errors, including the use of as much information available as possible for recovery. One way to improve the process is to use knowledge about the data and messages transmitted, which are considered not just as a set of bits, but as a structure with its own characteristics.

The TCP/IP protocol stack is the most common in computer networks. It consists of several layers, and at each level, each protocol of the same layer is built on a protocol of a lower layer and each protocol has its own specific task. The lower layer protocols (which belong to the physical/channel layers of the OSI model) deal with, among other things, the detection and correction of errors that may occur during data transmission. These protocols consider data as a stream of bits/bytes and do not take their semantics, structure and content into account. In this regard, some of the information that could be used to improve the reliability of data transmission is not utilized. On the one hand, this approach simplifies the implementation of communication channels, but on the other hand, in case of interference in data transmission, the speed and reliability of algorithms and technical solutions decreases [6, 7].

Therefore, it is necessary to improve the data transfer process, and in this article, it is proposed to use knowledge about the data and messages transmitted, which are considered as a certain structure with its own characteristics, and not just as an encoded binary data stream.

1.2 Scientific Researches and Issues Analysis

In order to better research the achievements in the field of error control coding, the articles and developments concerning the combination of different coding methods, adaptive coding and peculiarities of their implementation, as well as modelling of their operation were considered [8–10]. In particular, article [8] presents the usage of turbo codes in power-line communication channel networks (PLC). An example of adaptive codes usage to correct the errors is given in the article [9], which also emphasizes the effectiveness of adaptive algorithms utilizing depending on the current state of a channel. The article [10] presents the features of the implementation of error

control coding based on FPGA boards. It should be noted that such implementations are higher in complexity, but their advantage is the greater speed of information processing through the use of parallel computing, which is provided by FPGA.

In addition, the peculiarities of data processing with the use of protocols of the OSI model are considered, in order to utilize as much information about the transmitted data as possible to increase the probability of successful message transmission over network channels [11]. The article [11] states and substantiates the importance to apply algorithms for search and correction of errors during data transmission not only in large and enterprise networks, but also in personal and small user networks.

It is worth noting that this article considers the encoding of data without the loss of information (i.e. it is not about audio or video streams where the partial data loss is permissible).

1.3 The Goals

The goal is to propose and test a more integrated approach to data preparation before transmission over network channels, which includes analysis of the message structure and data contained in the message. The reliability of the transmission of structured data over network channels should be increased with the help of the revealed regularities. The research outlines the effectiveness of the content analysis of structural peculiarities and messages in order to use the results of analysis to improve the performance factors of data transmission. This article proposes to improve existing methods of data encoding by using of information about the content and nature of the messages themselves, which in theory can increase the reliability of the transmission of structured data.

2 Method of Adaptive Error Control Coding

2.1 The Typical Composition of Structured Data

Abstracting from the formats and implementations (JSON, XML, ProtoBuf, etc.), they are united by one: messages are divided into components, which are often called "tokens". These components have a certain semantic meaning (often even the corresponding token name is specified in the message itself), usually unique within the message type, and a set of valid values for each token (type). An information system can have any number of message types and any number of tokens in one message: their number depends on the complexity of the system. In different programming languages, such messages are mapped in different data structures, but in most cases, they are classes (such as in Java) and structures (such as in C). The messages are serialized before being sent over the network channels, and deserialized after receipt.

The previous article [12] discussed a way to use all these peculiarities of structured messages to improve the quality of data transmission over the network channels. In this article we consider the features and implementation details of the proposed method.

2.2   The Client–server Communication

Let's suppose that the information system is built on the principle of client–server architecture. The server responds to the client requests, and there are $k > 0$ types of requests in the system. Each request has its own request structure, and a corresponding unique response structure. All requests and responses to requests are the structured messages that consist of a set of tokens. The scheme of a request (ask) with the index k has the form of $Ak = \{a_{k1}, a_{k2}, ..., a_{kn}\}, n > 0$. The corresponding scheme of reply (response) is: $R_k = \{r_{k1}, r_{k2}, ..., r_{km}\}, m > 0$. It should be noted that the number of tokens and their type in the request and response may differ. In addition, in more or less complex systems, the tokens are hierarchical and may themselves contain embodied messages, i.e., for example, the response to request number 1 may look like this: $R_1 = \{r_{11}, r_{12}\}$, where token r_{11} is a numerical identifier and token r_{12} is composite: $r_{12} = S \{s_1, s_2, ...\}$.

2.3   Dynamic Subtypes

Let's use the above features of structured messages and requests in order to increase the reliability of data transmission over a network channel. The concept of a "dynamic subtype" is introduced. This is statistical information about the data transmitted by a specific token. It may change over time and be supplemented. In the simplest case, we consider the numerical type of token (int, 32 bits). If a certain number of messages containing a given numeric token and the token value were in the range (100–150) were transmitted for a certain period of time, it means that the current dynamic subtype of the token is a number encoded by $w = 6$ bits of information (0–50) with an offset equal to $s = 100$. Thus, such a simple dynamic subtype for a numeric data type can be encoded by two numbers: the number of bits and the offset. Let's consider another case, for example, for text strings. If a text token contained only a small set of constants (enum), the number of which does not exceed 256, then the dynamic subtype can encode all text values with a numeric single-byte value. In this case, such a dynamic subtype will be encoded with a dictionary, which will contain "key"–"value" pairs, where the key will be a 1-byte number, and the value— the corresponding text string. Another way to encode a text (or byte) string is to use three values: an array of valid admissible characters A and the length of a byte string $\{L_{min}, L_{max}\}$.

The current state of the dynamic subtype for each token can be controlled in three ways:

- for each client of the information system separately;
- for all clients in general;
- for each client separately with a timeout (TTL).

In the first case, during the encoding/decoding process, it is possible to take into account the peculiarities of each client device separately. The disadvantage is the greater complexity of the algorithm and greater requirements for computing resources and memory of the device, especially on the server side, because it must store information about dynamic subtypes for each client, and the clients must be identified by a unique identifier (e.g., UUID).

In the second case, the implementation is simpler, but the operational efficiency of the method will be lower. The third way is a compromise between the first two methods. This paper considers the simplest option, i.e., shared state of dynamic subtypes for all client devices.

2.4 Method of Encoding

So, let's have a message of a certain type, which consists of a finite set of tokens. We introduce the concept of message structure history, which contains the information about the current dynamic subtypes of tokens, as well as information about the old subtypes, because synchronization between hosts does not happen instantly. The history can be represented as a list or an array, where the index of an element is its identifier. The main requirements for such a list are:

- the ability to quickly make changes (add a new history element in case of a subtype change detection with one of the message tokens);
- the possibility of direct and fast access to each history element with constant complexity O(1).

The general linked lists are not suitable for the current task (because they do not meet the second requirement), so it is better to use the lists based on arrays, which are expanded by allocating a new array, if necessary, but with a certain growth rate calculated by the formula: $S_{new} = 3 \cdot S_{old}$.

The initial value is $S_{old} = 10$. It is also possible to use a hash table, which in most cases enables to add and search for elements with almost constant complexity $O(1)$, provided the correct hash function is selected. The hash function should be calculated quickly and has the form of $y = f(x)$, where x is an array of data of arbitrary size, and y is a number/array of fixed size. The number of collisions in the calculation should be as small as possible. That is, the distribution of the function values should ideally approach a uniform distribution [13].

Also, in order to verify the correctness of the data contained in the message, it is necessary to add a checksum to it, with which it will be possible to answer the question

with a high probability: whether the decoding process is completed successfully or not. There are many methods for calculating of checksums. The parity bit is the simplest option, which consists of adding one bit to the message, which is equal to "1", if the number of units is even, and "0"—otherwise. In terms of coding theory, the parity bit can be considered as a code $(n, n-1)$ with matrices:

$$H_{1,n} = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$$

$$G_{n-1,n} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ 0 & 1 & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 1 \end{pmatrix} \tag{1}$$

It is used, for example, in the UART protocol. But it is not reliable, and can only detect an error in one bit [14].

IP packets use a 16-bit complement code as a checksum. The header is divided into 16-bit blocks; their sum is calculated. If there are transfers to a new bit after the calculation, they are added to the amount. This method is more reliable and is already used in the IP protocol at the network level of the OSI model.

Then the typical structure of the encoded message will look like (Fig. 1).

The error control coding method performed by the $encode_N(\ldots)$ function can be any method with arguments:

- data to be encoded (d);
- number of additional bits for encoding (Δc).

Furthermore, if the encoding method does not provide an easy way to set and change the code parameters, another way to increase the probability of successful decoding can be used by engaging the callback function, which will be called by the encoding algorithm in case the impossible decoding of a part of the message is detected. The implementation of such a function in the simplest case involves forecasting based on the current scheme of dynamic subtypes, possible variants of the symbol to be decoded, checking the result with a checksum. The disadvantage of this method is that it is necessary to change the decoding algorithm. Therefore, the usage of the first variant is considered. The general algorithm is depicted in Fig. 2.
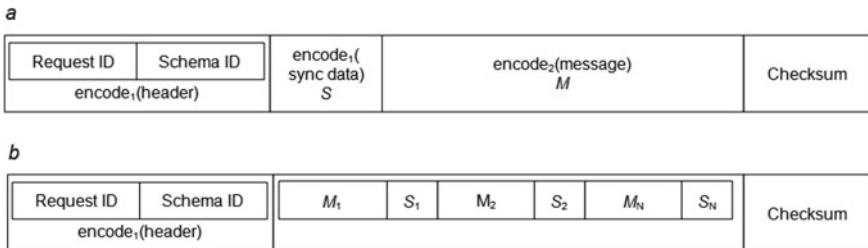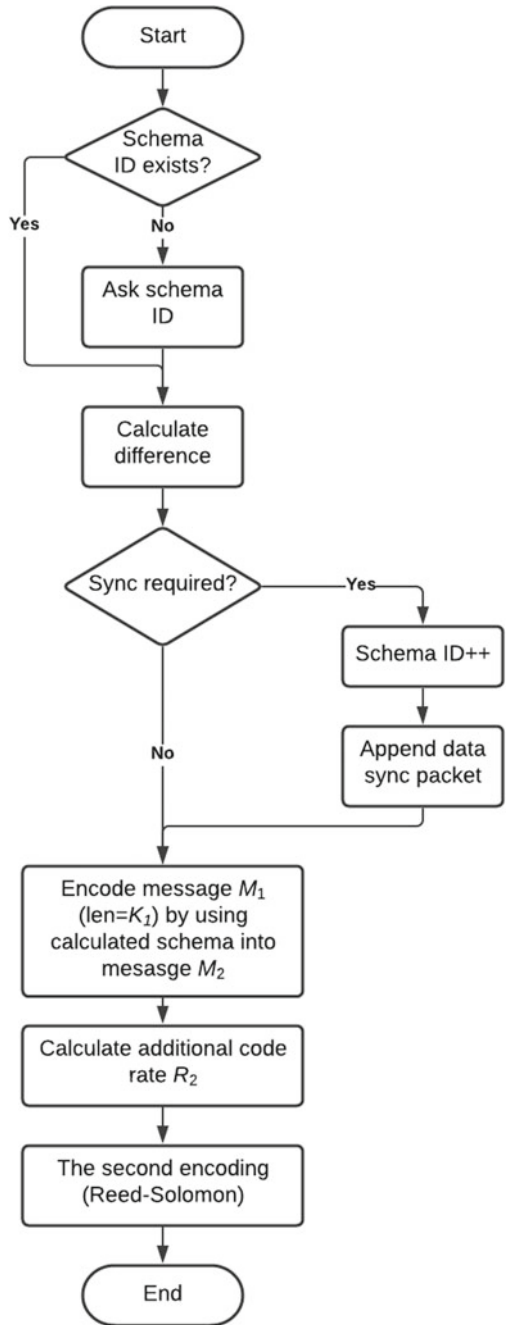

Fig. 1 Logical structure (**a**), physical structure (**b**)

We consider the results based on the Reed-Solomon code (implementation taken from the ZXing Open Source library) and the communication protocol of the ThingSpeak service for Internet of Things devices, developed by IoBridge company (now purchased by MathWorks).

The Reed-Solomon codes are a partial case of BCH codes. The BCH codes are the cyclic codes of some length n over the field $GF(q)$ with a distance d, under the condition that for some value of $b \geq 0$, the generator polynomial is:

$$g(x) = \text{lcm}\{M_i(x), \ i = b, b+1, \ldots, b+d-2\} \tag{2}$$

If we decrease the exponent of the min. polynomials that produce the generating polynomial of the cyclic code, the redundancy of the BCH code also decreases. The exponent cannot be less than one, and it is equal to one if x takes a value in the same field as the field $GF(2^m)$, which is used to build a check matrix of code. In such a field, the minimal polynomial of the element β is $x - \beta$. Since the exponents of the variable x correspond to the position of the codeword, the length of the code must not exceed the number of elements of the multiplicative group of the field.

The Reed-Solomon codes are often used as component codes in cascade constructions. In the method considered in this example, there is no such structure in the usual sense, but there is an add-on over the code in the form of a variable parameter of the code rate and pre-conversion of the message to a more efficient format, which increases the probability of the successful data transmission.

The advantage of the BCH codes and the Reed-Solomon codes in particular is the simplicity of their encoding and decoding algorithm in the channel with hard decision. The disadvantage is the lack of a reasonable decoding algorithm in channels with soft solutions. But this disadvantage is partially eliminated by usage of the method of "minimum generalized distance", which is reduced to multiple decoding in the channel with hard decision [15, 16].

If the division of the information message into parts and the alternation of encoded tokens with the encoded scheme is not taken into consideration, the procedure can be clearly depicted as follows (Fig. 3):

$$R_1 = K_1/N; \ R_2 = K_2 + K_S/K_N; \ \Delta C = K_1 - K_2 - K_S$$

That's why:

$$R_2 = (K_1 - \Delta C)/N$$

where

K_1 the number of data bits of message encoded by the Reed-Solomon encoder, bits;

N the total number of bits in the message, bits;
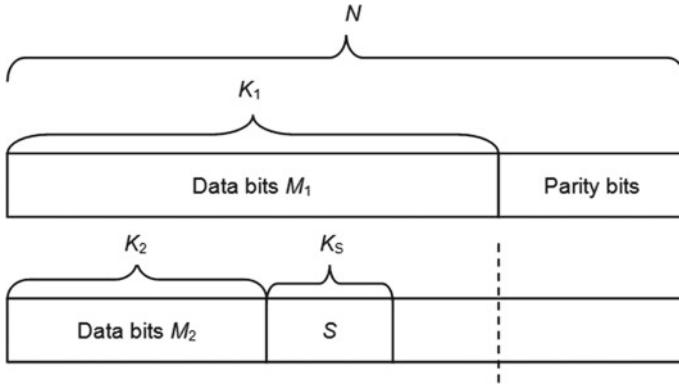
Fig. 3 The procedure of calculating of the amount of possible redundant bits of information

K_2 the number of data bits encoded by state encoder (message without state and schema), bits;

K_S the number of bits representing state and schema, bits;

ΔC the amount of space which can be used for adjusted error correction code;

R_1 the code rate of the origin Reed-Solomon code;

R_2 the max possible code rate after adjusting encoder parameters.

Let's take the following Reed-Solomon code: (15,13), GF(2^4) with the corresponding generator polynomials:

$$(15, 13) : g(x) = x^2 + 6x + 8 \tag{3}$$

The minimum polynomial is equal to: $\alpha = 2$. The code (15,13) can correct 1 error ($d = 3$). Since the code in fact operates not with bits, but with their sets of 4, one error is equated to an error in any bit in the set (or to an error in all bits together, i.e., the number of errors in the set can be from 0 to 4, but within one set their number is equal to one error), which improves the corrective capability of the code.

We take a set of the following messages of the same size:

```
1) 22 68 22 3a 31 2c 22 74 22 3a 31 37 35 2c 22 77 22 3a 34 35 7d
2) 22 68 22 3a 32 2c 22 74 22 3a 31 36 33 2c 22 77 22 3a 35 30 7d
3) 22 68 22 3a 32 2c 22 74 22 3a 31 32 32 2c 22 77 22 3a 34 34 7d
4) 22 68 22 3a 33 2c 22 74 22 3a 31 37 38 2c 22 77 22 3a 33 38 7d
5) 22 68 22 3a 33 2c 22 74 22 3a 32 30 31 2c 22 77 22 3a 33 39 7d
6) 22 68 22 3a 31 2c 22 74 22 3a 32 30 32 2c 22 77 22 3a 36 30 7d
7) 22 68 22 3a 31 2c 22 74 22 3a 31 37 36 2c 22 77 22 3a 34 38 7d
```

If the extra characters are removed, we obtain the token IDs and values:

```
1) 68 31 74 31 37 35 77 34 35 7d
2) 68 32 74 31 36 33 77 35 30 7d
3) 68 32 74 31 32 32 77 34 34 7d
4) 68 33 74 31 37 38 77 33 38 7d
5) 68 33 74 32 30 31 77 33 39 7d
6) 68 31 74 32 30 32 77 36 30 7d
7) 68 31 74 31 37 36 77 34 38 7d
```

The length of seven messages is 560 bits.

Let's construct the scheme of dynamic types of tokens for a set of messages (let the identifier be equal to 1). We have tokens with keys 0×68, 0×74 and 0×77, which can be matched to single-byte identifiers of the varint type (0×1, 0×2, 0×3). The token subtype 0×1 has a value in the range [0; 2] (0×2) with an offset of 0×31. The token subtype 0×74 has a value in the range [122; 202] (if we convert from ASCII to a decimal number) or if we specify an offset—then it is [0; 80] (0×60) with an offset of 122 ($0 \times 7A$). Performing a similar operation for the third token, we obtain the range [0; 22] (0×16) with an offset of 38 (0×26). So, the message scheme will look like:

```
Token ID    Origin Name    Type    Max     Offset
0x1         0x68           0x1     0x2     0x31
0x2         0x74           0x1     0x50    0x7A
0x3         0x77           0x1     0x16    0x26
```

If all numeric values are encoded with varint, and the text values are encoded with the text strings ending in zero-byte 0×0, and to indicate the number of tokens at the beginning of the scheme, the resulting scheme will look as follows:

```
01 | 03 | 68 00 01 02 31 | 74 00 01 50 7A | 77 00 01 16 26
```

The scheme is transmitted once, and the messages encoded using this scheme will look like this (the scheme number comes first, then it is the message):

```
01 00 35 07 01 29 0C 01 00 06 02 38 00 02 4F 01 00 50 16 00 36 0A
```

So we have: an original message size—560 bit. The size of the scheme—136 bit. The size of the resulting message is 168 bits without an identifier.

The first data transfer will contain a synchronization packet, so its size will be a maximum of $176 + 136 = 312$ bits of information. If the scheme does not change, then only the identifier is transmitted instead of the scheme, and therefore it is necessary to transfer $168 + 8 = 176$ bits of information. Hence, we have the values of the parameters: $K_1 = 560$, $K_2 = 176$, $K_S = 136$, $K_2 + K_S = 312$.

We encode the original message with the Reed-Solomon code, GF(24), $\alpha = 2$ (formula 4). Dividing the data array into blocks of 13 information bits, we obtain the following data symbols:

6	8	3	1	7	4	3	1	3	7	3	5	7
7	3	4	3	5	7	13	6	8	3	2	7	4
3	1	3	6	3	3	7	7	3	5	3	0	7
13	6	8	3	2	7	4	3	1	3	2	3	2
7	7	3	4	3	4	7	13	6	8	3	3	7
4	3	1	3	7	3	8	7	7	3	3	3	8
7	13	6	8	3	3	7	4	3	2	3	0	3
1	7	7	3	3	3	9	7	13	6	8	3	1
7	4	3	2	3	0	3	2	7	7	3	6	3
0	7	13	6	8	3	1	7	4	3	1	3	7
3	6	7	7	3	4	3	8	7	13	0	0	0

The result of encoding with the above code will look like:

6	8	3	1	7	4	3	1	3	7	3	5	7	6	15
7	3	4	3	5	7	13	6	8	3	2	7	4	15	11
3	1	3	6	3	3	7	7	3	5	3	0	7	2	9
13	6	8	3	2	7	4	3	1	3	2	3	2	7	14
7	7	3	4	3	4	7	13	6	8	3	3	7	13	10
4	3	1	3	7	3	8	7	7	3	3	3	8	11	12
7	13	6	8	3	3	7	4	3	2	3	0	3	6	1
1	7	7	3	3	3	9	7	13	6	8	3	1	6	1
7	4	3	2	3	0	3	2	7	7	3	6	3	12	1
0	7	13	6	8	3	1	7	4	3	1	3	7	2	13
3	6	7	7	3	4	3	8	7	13	0	0	0	9	3

The resulting size of the encoded message obviously increases by the number of redundant symbols, and for this code—by 2 symbols for each block. The resulting encoded message size is 165 symbols, or $165 \cdot 4 = 660$ bits. Where 80 bits are parity symbols. If we take the message with the scheme of subtypes, it contains 312 data bits. Therefore, in an equivalent size encoded message can include $660 - 312 = 348$ bits that can be used to improve the transmission reliability.

Code word length $N = 15$. $N_{total} = 660$—total number of bits of encoded message, $k_{total} = 312$—the total number of data bits, $V = N_{total}/N = 660/15 = 44$—the total number of code words (each code word $= 15$ bits).

So each codeword produced by the adjusted encoder should have minimum of 8 data bits:

$k = k_{total}/V = 312/44 = 8$ (upward rounding)—the minimal amount of data bits in each codeword.

So, we can use max. possible code (15,8) but such code does not satisfy the condition of $(n - k \bmod 2) = 0$. Therefore, it is necessary to round up, so the corresponding code that can be used is (15,9).

$$g(x) = x^6 + 7x^5 + 9x^4 + 3x^3 + 12x^2 + 10x + 12 \tag{4}$$

Table 1 Results for the example above

	encode$_1$[(15,13), message]	encode$_1$[(15,9), encode$_2$(message)]
Length, bits	660	540
Max acceptable percentage of errors, %	6.7	20

This code can correct 3 errors (d = 7) instead of 1. So, its corrective capability is three times greater. This means that the probability of successful transmission of a data packet encoded with 1) the use of information about the structure of the message content and 2) the Reed Solomon's code (15,9) is much higher than a data packet of the same size but encoded only by code (15,13):

```
0   1   0   3   6   8   0   0   0   1   0   2   3
1   7   4   0   0   0   1   5   0   7   10  7   7
0   0   0   1   1   6   2   6   0   1   0   0   3
5   0   7   0   1   2   9   0   12  0   1   0   0
0   6   0   2   3   8   0   0   0   2   4   15  0
1   0   0   5   0   1   6   0   0   3   6   0   10
0   0   0
```

The result of coding (Table 1).

```
0    1    0    3    6    8    0    0    0    14   9    9    5
12   0    1    0    2    3    1    7    4    0    0    9    4
10   1    3    4    0    1    5    0    7    10   7    7    0
6    15   3    6    1    14   0    0    1    1    6    2    6
0    1    3    3    15   11   15   15   0    0    3    5    0
7    0    1    2    8    8    15   0    9    12   9    0    12
0    1    0    0    0    6    10   12   8    0    14   1    0
2    3    8    0    0    0    2    4    10   12   0    1    2
12   15   0    1    0    0    5    0    1    6    11   2    1
0    5    11   0    0    3    6    0    10   0    0    0    0
0    4    4    10   10
```

The function of code parameters selecting in this case is obviously a step function. Therefore, the value of additional corrective capability (as a percentage) depends on the successful result of dividing the message into parts. In particular, if the original message was *N* bytes in size, and the internal encoder/decoder supports codes (15,*x*) in this case, the function is depicted in Fig. 4.

Let's explain the Fig. 4 depicted above. It represents an additional portion of errors in the message which can be safely detected and corrected after receiving the message by decoder, depending of amount of data which has been encoded by the
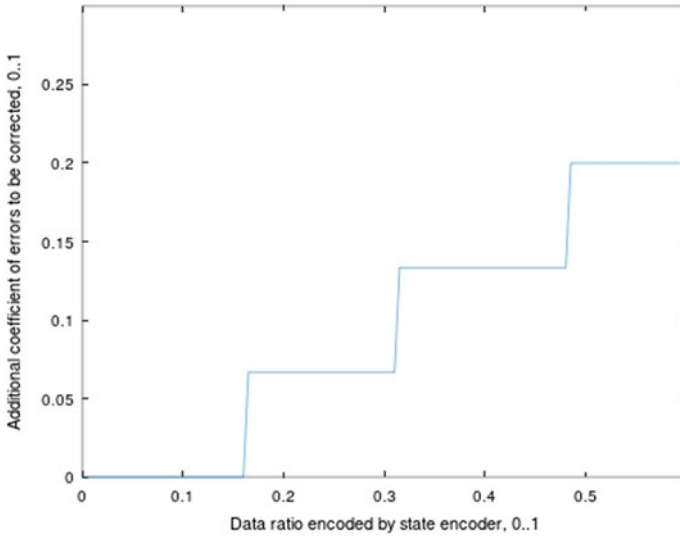
Fig. 4 Additional acceptable percentage of errors which can be corrected

encoder$_2$ (state encoder). For example, let's have an origin message containing 1001 bits. Reed-Solomon (15,13)-encoder produces encoded message with the length of 1155 bits which can be transmitted over network channel. State encoder can extend the amount of space within these 1155 bits for additional parity bits by reorganizing the message content. If state encoder reports to the Reed-Solomon encoder that there are additional $1001 \cdot 0.16 = 160$ bits which can be used for parity bits then the Reed-Solomon encoder may generate code with decreased code rate, e.g. (15,11) instead of (15,13), but without increasing the amount of data to be transmitted. So within the same 1155 bits the acceptable amount of errors which can be safely corrected by the decoder are increased by 6.6%. Please note that the Fig. 4 shows the theoretical limit, actually there are necessarily additional synchronization packets that increase the number of bits of information required for decoding during the transmission.

3   Conclusions and Suggestions

The proposed method of adaptive error control coding analyzes and prepares data for transmission over communication channels by combining 2 encoders: Reed-Solomon encoder (actually it can be replaced by any parameterized encoder) and state encoder. The reliability of the transmission of structured data over network channel is increased in case of using this combined approach as the corresponding decoder may accept and correct more errors than in case of using only Reed-Solomon encoder.

1. The article presents the possibility of effective use of the considered coding method by example: (a) the Reed-Solomon code and (b) encoding the message structure. The possibility to increase the corrective capability provided by the proposed encoding method was tested using a test dataset and the package "communications" in GNU Octave software [17].

2. This method can be combined with various encoding methods, in particular, any other code with the parameters which can be parameterized is possible to use instead of the Reed-Solomon code. In the simplest case, this can be an ordinary linear code (if the goal is to use as simple encoding/decoding algorithm as possible, which can be present in devices with limited resources) or other more complex codes, such as from the family of convolutional codes or turbo codes.

3. The further research can be carried out with the study of the coding efficiency for the data of different types and structures, as the choice of parameters depends on the results of the breakdown of the original messages into parts, which in turn may be different, depending on the internal structure of messages.

4. The algorithm for finding and selecting of tokens can be theoretically improved and optimized for faster operation. It can also be made more versatile by dividing it into two parts by analogy with the approach adopted in the Apache Thrift libraries: (a) token encoding/decoding; (b) classification of subtypes and generation of the range of acceptable values.

5. In addition, depending on the specific implementation of the method, other data structures (binary trees, arrays, etc.) can be used for implementation in different computing architectures. There is a theoretical possibility of parallel computing usage or even full-fledged parallelization based on FPGA.

References

1. Hemanth G (2020) A study on the Internet of Things. Int J Eng Trends Appl 7(4):5–9
2. Ansere JA, Han G, Liu L, Peng Y, Kamal M (2020) Optimal resource allocation in energy-efficient Internet-of-Things networks with imperfect CSI. IEEE Internet of Things J 7(6):5401–5411
3. Alberti AM, Scarpioni GD, Magalhaes VJ, Cerqueira A (2019) Advancing NovaGenesis architecture towards future Internet of Things. IEEE Internet of Things J 6(1):215–229
4. Koivo H, Elmusrati M (2009) Systems engineering in wireless communications. John Wiley & Sons, Chichester, U.K, pp 242–243
5. Plummer DC. RFC826: an ethernet address resolution protocol. In: Internet standard. https://datatracker.ietf.org/doc/html/rfc826
6. Lakshman TV, Madhow U, Suter B (2000) TCP/IP performance with random loss and bidirectional congestion. IEEE/ACM Trans Netw 8(5):541–555
7. Iliev TB, Hristov GV (2007) Simulation estimation of the forward error correction of turbo codes. In: 2007 8th International conference on telecommunications in modern satellite, cable and broadcasting services, pp 521–524
8. Abd-Alaziz W, Mei Z, Johnston M, Le Goff S (2017) Non-binary turbo-coded OFDM-PLC system in the presence of impulsive noise. In: 25th European signal processing conference (EUSIPCO), pp 2576–2580

9. Sunghwan K (2015) Adaptive FEC codes suitable for variable dimming values in visible light communication. IEEE Photonics Technol Lett 27(9):967–969
10. Zinchenko MY, Levadniy AM, Grebenko YA (2019) Concatenated error correction code implementation on FPGA. In: 2019 Systems of signal synchronization, generating and processing in telecommunications, pp 1–4
11. Pollard JK (2007) Packet error correction in personal area networks. In: First Asia international conference on modelling and simulation, pp 287–291
12. Andrushchenko R, Zaitsev S, Druzhynin O, Shelest M (2020) Method of encoding structured messages by using state vectors. In: Mathematical modeling and simulation of systems, pp 144–153
13. Mailund T (2019) The joys of hashing: hash table programming with C. Apress, Berkeley, California
14. Valvano JW (2014) Embedded systems: introduction to ARM® Cortex(TM)-M microcontrollers. Austin, Texas, pp 377–382
15. Kudriashov BD (2016) Foundations of coding theory, pp 70–187
16. Wesołowski K (2002) Mobile communication systems. J. Wiley, New York, pp 30–35
17. GNU Octave official web-site. https://www.gnu.org/software/octave/index

Usage of WBAN Systems and IoT Solutions in a Medical Environment

John N. Davies , **Mariya Verovko** , **Oleksandr Verovko,** **and Iryna Solomakha**

Abstract The COVID-19 pandemic has had a huge worldwide effect on everyone. A major issue that it has brought to light is the requirement for social distancing resulting in working from home with the ability to utilize video conferencing for meetings. This has become an essential part of everyday life, hence the requirement for reliable computer networks. Coping with social distancing in a medical environment is far more challenging however the use of Wireless Body Area Network (WBAN) and the ability to share the data has become enormous aid. A great deal of development in body sensors has taken place and these devices are becoming commercially available. WBAN devices can be used to measure many human conditions e.g. heart rate, temperature, sleep-patterns etc. to help with diagnosis. Most papers on WBAN concentrate on sensors but this paper investigates the suitability of the technology and the support network required for use in a medical environment. An integrated system using AWS IoT services to support WBAN devices in different environments is covered. The implementation of a smart hub for healthcare monitoring is proposed. The concept of an Intelligent Bed is introduced and a solution is provided consisting of a smart hub attached to a bed. Parameters associated with the wireless characteristics are calculated and use of the reliability of the AWS IoT services for use as a healthcare monitoring system is investigated.

J. N. Davies
CARDS, Glyndŵr University, Wrexham, UK
e-mail: johnndavies4@gmail.com

M. Verovko (✉) · O. Verovko
AgileVision sp. z o.o., Krakow, Poland
e-mail: miya.tevkun@gmail.com

O. Verovko
e-mail: averovko@gmail.com

I. Solomakha
Chernihiv Polytechnic National University, Chernihiv, Ukraine
e-mail: iveria60@gmail.com

1 Introduction

1.1 WBAN as an Emerging Solution in Healthcare Monitoring

A Wireless Body Area Network (WBAN) according to a medical dictionary is a wireless network typically used to monitor vital signs and other physiological parameters of the human body [1]. However WBAN covers a variety of consumer electronics applications as well as real-time health monitoring devices. The technology is based around the IEEE 802.15.6 standard aimed at short distance, extremely reliable wireless communication linked to the human body. WBAN devices may be embedded inside the body, surface mounted on the body or humans carrying in clothes pockets, by hand or in various bags [2]. This means that the sensor used to provide the data is not a fixed type and varies with the type of measurement required and manufacturer [3].

It is also necessary to consider how the data is to be collected, saved and shared to provide accurate and often live monitoring. There is little point in not using existing components where appropriate so the design should utilize Local Area Network (LAN), Metropolitan—Campus Network (MAN) and Wide Area Networks (WAN) and Cloud Applications. The security associated with this data has become very important due to the introduction of recent legislation to define who is able to access it [4]. Figure 1 shows WBAN in different medical environments.
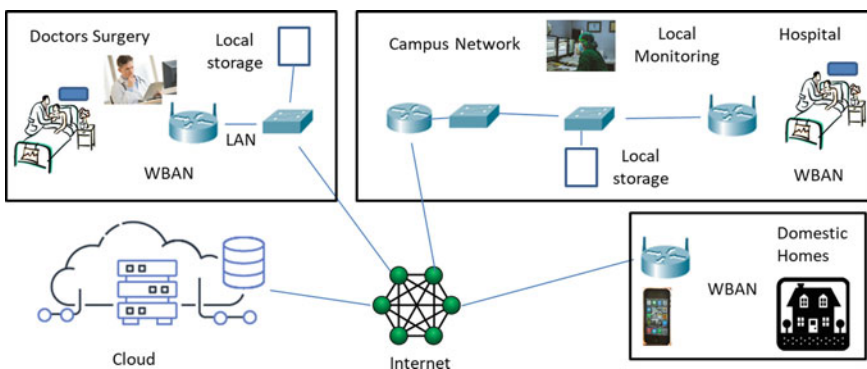

Fig. 1 WBAN usage in a range of environments

1.2 Features of WBAN Usage in Different Environments

Since there are a wide range of WBAN devices capable of measuring many different characteristics, often the environment where they are used dictates the network used. Hospitals have very complex environments that are always changing due to developments in ailments, treatments, equipment and working practices. Each department has its own specialized equipment for investigation and diagnosis of patient issues. Most departments have expensive equipment that provides electronic data that can be distributed around the hospital e.g. X-ray slides (images). Introduction of the WBAN in a standard hospital operation can optimize the usage of resources and reduce the load from medical personnel by improving data flows [5]. This paper makes a great simplification of the issues that are found in the various departments but by adopting a standards approach it is anticipated that future developments and changes can be made easier and provide benefits to all.

Doctors Surgery: Doctor's surgeries are now becoming a hub for handling patients before being funneled to the appropriate medical specialist. Present trends are to reduce the number of smaller surgeries and invest in larger sites with more facilities and increased capability. Introduction of WBAN (e.g. installation of basic health monitoring systems with automatic data exchange and synchronization) here would be an advantage since it provides consistency with the hospitals' consistency with the hospitals [6]. Doctor's surgeries would be treated like mini-hospitals without the overhead of overnight stays and provision of food etc. This functionality would enhance the normal running of the surgery. It would be advantageous if most of the appointments could take place by telephone or by video conferencing. Multi-way video conferencing is well established and by having specialists on hand it may reduce the number of separate appointments required.

Domestic Environment: The issue of health monitoring for people, who stay at home, is of great importance because people have to self-isolate even if they have mild symptoms. Most homes use Wi-Fi to connect to the internet router and are used for general purpose applications. While people are at home they have the ability to use a limited number of sensor devices due to, cost of equipment, knowledge of operation and limited data rates of their internet connection. But the high availability of Activity Trackers and other smart-home devices, which provide ability to measure many human conditions (heart rate, temperature, sleep-patterns etc.), makes the introduction of WBAN into the domestic environment possible [7]. Usage of WBAN based monitoring is a solution that can be used for remote health monitoring for the people, who can be treated at home [8]. Using this approach can also prevent incorrect diagnosis, since the data from sensors are more accurate than the description obtained from patients using audio or video conferencing.

2   Design of WBAN System and Its Components

2.1   General Approach for WBAN Development

Medical monitoring equipment is heavily regulated and rightly so. There are a number of International Standards bodies providing guidelines to ensure interoperability. From a computer networking, sensor point of view the IEEE the standard agreed for sensor networks is the IEEE15 range. Areas of interest include WBAN /Bluetooth, High/Low Rate WBAN, Mesh, Visible light, Peer Aware Communications, Key Management Protocol and Wireless Next Generation. IEEE15.6 covers Body Area Networks which is the appropriate standard for this area of research [9].

Sensors provided with Bluetooth interfaces can normally only store a limited amount of data and so there has to be a local Hub that can collect this data and either store it locally or use a network to transfer it on to the appropriate cloud application. Storing the data locally is not very convenient and so the use of a LAN to provide the data at a central point, e.g. a Nursing station, is seen as being essential. Sharing this data via a computer network provides great advantages to the well-being of the patient.

It is necessary to connect the WBAN to a LAN due to the short distance that WBANs are capable of operating over. Typically a LAN would provide a network that would operate over a ward in a hospital, a doctor's surgery or a patient's home. This can be provided using either a cabled network or a Wi-Fi wireless network. Utilizing a Campus network (MAN) in a hospital situation allows the monitoring of patients at a locally level or at a central point e.g. a Nursing station. To perform remote monitoring data exchange is performed using an Internet connection.

2.2   WBAN Smart Components

In hospitals patients are continuously monitored by equipment located at their bedside. Blood pressure, heart rate, pulse oximetry, respiratory rate etc. are measured by specialist equipment containing either wired, Wi-Fi or Bluetooth connections to allow remote monitoring.

There are a number of developments that are taking place in the area of monitoring. An example of this is the Vscan Extend™ handheld ultrasound that can be used for abdominal, cardiac, fetal/obstetrics and, more significantly in the present climate, thoracic/lung ultrasound. This is a portable device and has the ability to transfer images via a wireless network so can be used in emergency circumstances to upload important data before transporting the patient to a hospital. Another example is the Meditronic capsule endoscopy device PillCam™ [10]. It is a camera inside a capsule that is swallowed by the patient and captures images at a frame rate of 2–6 fps governed by the capsule speed as it travels through the body giving much more accurate images. There is a sensor belt worn around the patient's waist to receive

the images via a wireless network. A further example is the Dexcom CGM which is attached to the body just below the skin and every 5 min monitors the glucose content of the blood. Readings are then uploaded to a smart phone or similar device. This provides continual monitoring and alarms rather than the daily tests that are utilized at present resulting in the ability to detect changes at a much earlier stage hence less medical intervention required [11, 12].

Since the monitoring equipment is provided by many different companies it is necessary to attach some kind of computer to convert the data into a standard form. This simplifies the combination of different types of data for monitoring and storage and also helps integrate newly introduced sensors [13]. WBAN tend to operate over small distances and so by attaching a Smart Hub (a cheap small computer e.g. Raspberry Pi) to a network to collect data from sensors in a particular area allows creating a smart hub device. Such a smart hub can be used in a domestic environment to obtain data about a patient's health state or it can be attached to a bed in a hospital and make it an intelligent bed. Many sensor devices provide smart phone access the Raspberry Pi would control this which would help reduce security problems. A Raspberry Pi has various adapters which allow it to connect with sensors using Bluetooth etc. and obtain the data about patient health. However further data processing, secure transmission, storage and sharing are the tasks to be implemented. There are several approached and service that can be used for development of complete IoT solution.

2.3 Storage of Health Data in the Cloud

Storage of healthcare data has always been a complex task because of lack of standardization in data-collection. Medical personnel have typically used a mix of storages, e.g. paper notes, several electronic patient record systems (EPRS), which creates duplications and redundancy of stored data. Such approach leads to time wasting and increases the possibility of medical errors due possible data conflicts or data loss.

Described problem has been partially solved by the introduction of EHR systems. Usage of EHR systems significantly simplifies the process of data recording, data storage and data access. Already used in the most part of EU and US hospitals, EHR systems are moving on to the next stage of the development, which should increase data mobility by integration of the EHR system with portable electronic devices, which could be used by nurses, doctors or patients. Another huge milestone for healthcare data storage is the introduction of International Classification for Nursing Practice (ICNP), which provides an agreed set of terms to record observations and interventions of nurses. ICNP based framework is integrated with the EHR and significantly simplifies data sharing and data management [14, 15].

Standardization of healthcare data opens new possibilities for these data storage. One of the approaches that could be used is the usage of cloud computing, which is widely spread in many business areas. Existing restrictions in healthcare data storing don't allow EU healthcare to obtain full benefit from cloud data storage, since it

has limitations in scalability. However there is a set of cloud-based solutions that are already in use. They are AWS Cloud, Azure for Health, Apple Health Records, Cloud Healthcare API by Google, Watson Health by IBM etc. [16].

There are many benefits, which are provided by cloud-based solutions. Usage of the cloud significantly simplifies medical data access, which means that medical personnel can quickly and easily get or upload all necessary information and can be sure that the provided data is the latest and up to date. Other valuable aspect is the security and role-based availability of the information, which is already implemented by cloud provider. There is also a wide range of solutions for integration of cloud-based data with other services and applications, which makes it easier to use healthcare data in multiple systems. Cloud providers also perform real-time data backup, which significantly increases data reliability and prevents possible data losses. Considering this information, the conclusion can be made that strong healthcare data in a cloud is not just a trend, but also the way to improve healthcare data availability and operation of electronic medical systems and solutions.

3 Implementation of Smart Hub Using AWS

3.1 Development of IoT Solutions Using Amazon Services

AWS IoT is one of the leading IoT platforms that provide features and services to simplify implementation of IoT solutions. One of the AWS IoT Core features is the availability of the AWS IoT Device SDKs, which provide functionality to build IoT applications using various hardware platforms. These SDKs consist of open-source libraries, developer and porting guides and performs connection, authentication and data exchange between devices and AWS IoT Core using MQTT and WSS protocols. There are various AWS IoT Device SDK versions, implemented for C++, Java, Python and JavaScript [17, 18].

The smart hub concept proposed is based on the data exchange between Raspberry Pi and sensors. Further data processing, storage and sharing is the task that could be solved by extending AWS onto a Raspberry Pi (or similar hub) by using AWS IoT Greengrass. Current approach allows devices to process data locally while using cloud resources and services at the same time (storage, analytics, troubleshooting, management etc.). Processing of the local events and data is performed by AWS Lambdas, which operates on the local devices. The general architecture of the IoT solution, designed using AWS IoT Greengrass, can be seen in Fig. 2.

Considering operation issues and features, related to usage of AWS IoT Greengrass, the architecture of the intelligent bed solution is separated into two parts:

- AWS IoT Greengrass Core Device, which operates in the local environment (a hospital, at home etc.);
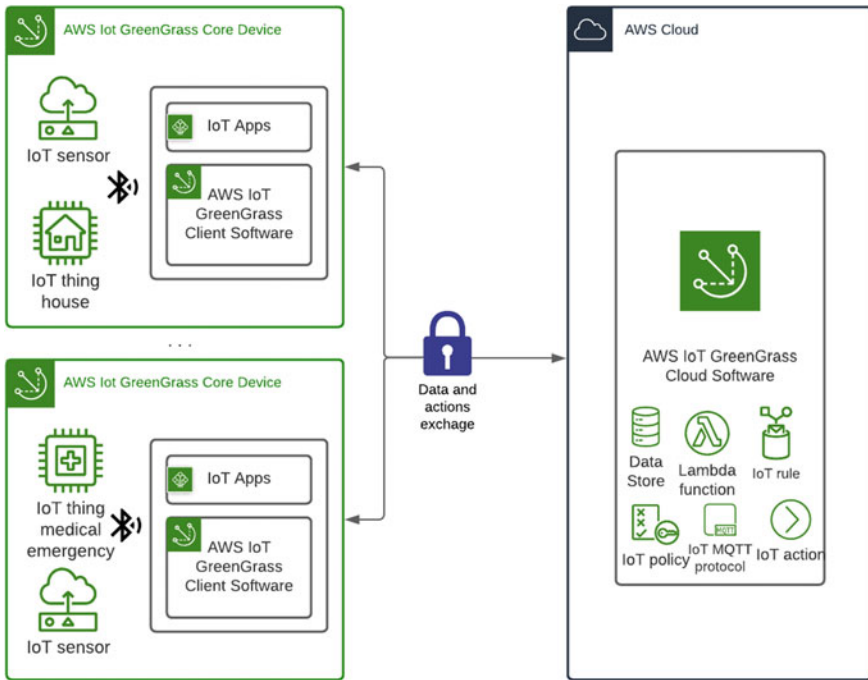- AWS IoT Greengrass Cloud Software, which operates in the cloud.

Fig. 2 Architecture of the IoT solution, performed using AWS IoT Greengrass

AWS IoT Greengrass Core Device includes sensors and other IoT things, which provides data for the Raspberry Pi. The Raspberry Pi needs to have AWS IoT Device SDK installed to perform cloud features. The data processing is performed by AWS Lambda, which operates directly on the device, but has a data exchange connection with the Amazon cloud.

AWS IoT Greengrass Cloud Software includes IoT Greengrass settings, such as IoT security policies, AWS Lambda settings and subscriptions, which identify data resources and direction, MQTT topics, which perform as data pipes and IoT rules, which declares how the data should be processed further (stored directly to DB, converted to SQS, SNS or other queue type).

The interaction between the two parts is performed using MQTT. Data is sent or received in the format of an IoT action and linked to a corresponding MQTT topic, which is identified in Lambda subscriptions. Any Lambda, placed in a domestic environment, could only read or write to defined topics. The security of data exchange is controlled by Amazon and is generally performed using SSL certificates, which should be stored on a local device. There are several ways to process data from MQTT topics. Each topic has its own unique arn and should be accessed using it. Direct data transferring from MQTT topic to Amazon SQS, Amazon SNS and AWS DynamoDB should be defined using IoT rules.

3.2 Design of the Smart Hub Using Amazon Cloud Services

Implementation of the IoT hub solution using AWS IoT Greengrass has been used for the development of a smart hub. AWS IoT Greengrass covers the tasks of data collection and transmission to the Amazon Cloud. Further analysis, processing and continuous data sharing to allow doctors to control the state of their patient remotely require usage of additional services behind AWS IoT Greengrass. The proposed architecture of the IoT based health monitoring system is present in Fig. 3.

Amazon services and units, used for intelligent bed implementation, are following:

- AWS IoT Greengrass that is generally used for data collection and data exchange between local device (Raspberry Pi) and cloud environment;
- AWS IoT Greengrass Lambda, which contains program code, that operates with data on a local device and sends obtained metrics to the cloud;
- IoT actions, data format for link between local and cloud environment;
- IoT MQTT protocol, used for secure data exchange;
- IoT policy performs security management and access control to cloud resources;
- IoT topic as a pipe for actions from Lambda, placed in a cloud;
- IoT rule, enables data transferring from IoT topic to Amazon SQS;
- Amazon SQS as a temporary data storage, which receives data from IoT Greengrass and send it next to other Amazon services;
- AWS Lambda function, which is subscribed to SQS topic and perform processing action on each data updates;
- Amazon DynamoDB stores collected metrics and data about patient health;
- AWS Lambda functions, obtain processed data from DynamoDB and then provide it to Web Application, which could be accessed by the doctor or other user;
- Amazon Cloudfront for web-application to make it available over the Internet.

Implementation of the proposed concept starts from the placement of the Raspberry Pi and healthcare sensors (pressure, pulse etc.) into patient domestic environment. Raspberry Pi requires Bluetooth and Wi-Fi connections, enabled AWS IoT Greengrass and working IoT Greengrass Lambda on it. At regular intervals the metrics data is read from sensors by Raspberry Pi Bluetooth adapter. The program
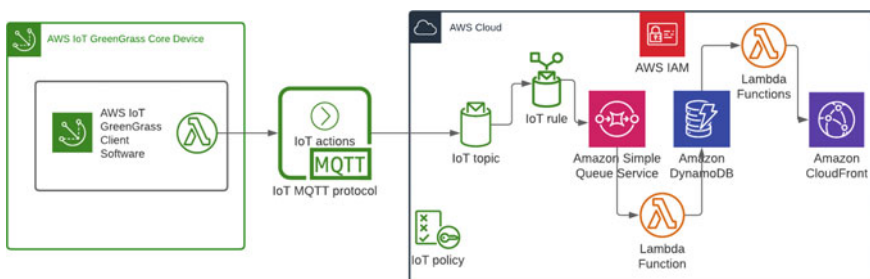

Fig. 3 Architecture of the smart hub, performed using AWS

code, which performs data requests is performed by IoT Greengrass Lambda. Metrics data are converted to IoT Action format by Lambda to be transmitted to AWS Cloud. IoT action is a data-structure in JSON format. Each IoT action is passed to AWS Cloud using MQTT and is placed into the specified MQTT topic. Each IoT Greengrass Lambda can push data to or read data from MQTT topic. The data flows are set directly in AWS IoTGreengrass settings.

Specified MQTT topic contains data from local device, but for further analysis and sharing should be placed in other Amazon storage or message service. There is a possibility to access MQTT directly from AWS Lambda using its' unique arn, but such an approach can lead to a security gap and significantly reduce flexibility. It requires hardcoded parameters in AWS Lambda and its' AWS IAM security policy, because AWS IAM security policy for IoT MQTT can't be attached dynamically. That is the reason why Amazon SQS is used as additional data converter and data storage. To transfer data from MQTT topic to Amazon SQS IoT rule is used, which is defined in IoT Greengrass settings using SQL syntax.

AWS Lambda is used as a SQS event listener. Current Lambda function is used to process obtained data and store calculated statistics, gathered metrics and average values for each patient. DynamoDB is used as data storage, where data is written from previous AWS Lambda.

Metrics from DynamoDB are used as a data source for AWS Lambdas, which provide information for Web-application, used by medical personal to control the state of each patient. AWS Lambdas are implemented as a REST API over HTTPS. The Web-application is placed in Amazon Cloudfront. Application code is stored in Amazon S3. Features of the design and security of the application for medical personal are not covered in current paper, since it doesn't relate to the topic of the investigation.

4 Analysis of Proposed Approach

4.1 Analysis of Data Flows and Data Loses in Cloud Architecture

The proposed approach contains complex data flow with a set of data changes, that can lead to data loose and misrepresentation, which is a critical point for healthcare data. The main reason of the possible data confusions are delays in data transferring between used Amazon services and units. An important task of current investigation is to find critical point, when the data correctness will be significantly impacted.

The overall time to deliver data from sensors to storage could be calculated using the function:

$$T = t_1 + t_2 + t_3 + t_4 \tag{1}$$

where t_1 is a time to deliver data from sensor to Raspberry Pi over Bluetooth and process it by local lambda function, t_2 is a time required to transfer data from local device to cloud over the Internet using MQTT, t_3 is AWS IoT rule operation time, t_4 is a time of the operation of second lambda function, which includes time on data obtained from SQS, data processing and its storing to Dynamo DB. Analyzing presented time units the conclusion can be made, that the longest time units are t_2 and t_4, since both of them include not only the time to transfer data over the network, but also running program code. Another key point, is that $t_4 > t_2$, because local lambda function only performs converting metrics to JSON format, while the second lambda function performs queries to DynamoDB to calculate statistics, which block its implementation till the data is not received. Considering the scenario described above there is a point when there could be data lose from SQS after obtaining data from MQTT before processing by AWS Lambda.

The overall data lose can be identified using the number of all packages passed to SQS (N_a), number of packages, that could be stored in SQS (N_s) and number of packages that could be received and processed by Lambda (N_r). In case of zero data lose $N_a = N_s = N_r$. However, this equation is not a constant, since both N_s and N_r are dependent on the AWS setting and affects each other. N_r is identified by a function:

$$N_r = \frac{1}{(t_L + dt)} * t_{op} \tag{2}$$

where t_L is a Lambda operation time, dt—time between Lambda operation cycles, t_{op}—overall time of system operation.

N_s can be defined as a function, which depends on t_s—time of message storage in SQS, t_{op}—overall time of system operation and N_a—total number of packages, that arrived to SQS:

$$N_s = \frac{t_s}{t_{op}} * N_a \tag{3}$$

According to the AWS specification, the time of Lambda function operation varies from 1.5 s (in cold start mode) to 15 m (maximum timeout) and could be identified as a set $T_r = [1.2, 900]$. Time the data could be stored in Amazon SQS is from 1 min to 30 days (with 4 days as default), so it is a set $T_s = [60, 2592000]$ [7]. So, the union of two sets when the data lose could be identified is $T_u = T_s \cup T_r = [60, 900]$, which means that $t_L \in [60, 900]$ and $t_S \in [60, 900]$, so 60 s and 900 s are boundary values of this parameters for both functions, which identifies N_r and N_s. For further calculations both t_s and t_L will be replaced by t_w.

Considering the previous calculation the conclusion that can be drawn is that data lose could be identified as the intersection for data unions obtained from the two functions $f(t_{op})$ for calculation N_s and N_r.

$$N_r = \frac{1}{(t_w+dt)} * t_{op}$$
$$\left\{ N_s = \frac{t_s}{t_{op}} * N_a \right\}$$
$$t_w \in [60, 900]$$
$$\delta N = N_s \cap N_r \tag{4}$$

To illustrate the results the graphical results were obtained for $N_a = 1000$ and average $d_t = 10$ ms. The result of each particular function N_r and N_s is the union of values between functions for boundary values 60 s and 900 s. So, the result was identified by the system of functions. The obtained result is shown in Fig. 4.

$$\frac{1}{60.2} * t_{op} > N_r > \frac{1}{900.2} * t_{op}$$
$$\left\{ \frac{60}{t_{op}} * 1000 < N_s < \frac{900}{t_{op}} * 1000 \right\}$$
$$\delta N = N_s \cap N_r \tag{5}$$

As can be seen from Fig. 4 there is an interception of the function. The data lose starts at the point when $60/t_{op} * 1000 = 1/60.2 * t_{op}$, $t_{op} = 1901$ s and the top value is obtained when $900/t_{op} * 1000 = 1/60.2 * t_{op}$, $t_{op} = 7363$ s.

This analysis shows that there is a possibility of data loss in a system implemented with proposed architecture. However, current numbers could be reduced to 0 by using the following AWS settings—set Amazon SQS storage time more than 15 min or leave it in default state (4 days). The second key point is AWS Lambda scalability—in the case of the Lambda operation taking too long AWS automatically creates one more to force the operation and prevent data loss.
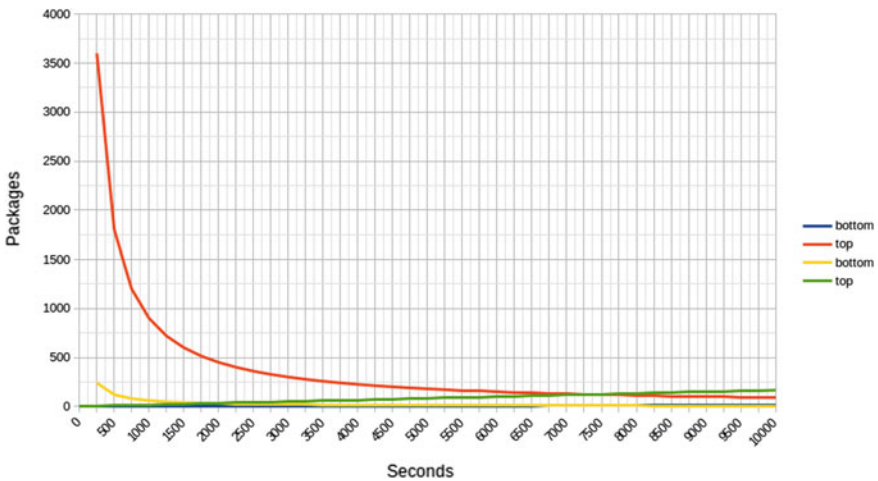

Fig. 4 Graphs of data loss

4.2   Wireless Distance Simulation

The proposed smart hub device could be used for the intelligent bed concept implementation in a hospital [19]. To define how many smart beds could be connected to a standard Wi-Fi point a simulation has been taken.

All hospitals have different layouts for the wards depending on their age, size, built material and management philosophy. This doesn't make it very easy to design and implement computer networks. The use of international standards and utilizing a standard design philosophy can simplify this task. There are certain basic physical attributes that can be used, a hospital bed is a maximum size of 0.96 m × 2.2 m, and space around a bed is 2 m × 2 m. These dimensions provide the basis for deciding which network technologies to use and are used as the parameters for a simulation.

Bluetooth and Wi-Fi are based on wireless technology operating in the gigahertz part of the spectrum. They utilize antennae and so follow the Maxwell–Faraday equation. There are some assumptions made in the simulator that simplify the calculations.

Let the transmission power $= P_t$. An important assumption made is that the medium is "free-space". Also the devices have a "hypothetical" isotropic antenna i.e. it radiates equally in all directions. If r is the distance, the power transmitted in a particular direction would be

$$\frac{P_t}{4\pi r^2} \tag{6}$$

It is assumed that the antenna on the receive side it identical to the transmit side and so the power received P_r given that S_r is the power hitting the antenna

$$P_t = \frac{G_r S_r}{4\pi r^2} \tag{7}$$

Let the gain on the receive antenna be G_r. The transmit side will also have a gain say G_t which would be multiplied by P_t. Additionally the power received is a function of the wavelength (λ) of the signal transmitted. Since the relationship between the wavelength and frequency (f) is defined as $\lambda = \frac{c}{f}$ where c is the speed of light, then the higher the frequency the lower the received power. So received power is

$$P_r = \frac{G_r \times G_t \times P_t \times c^2}{16\pi^2 r^2 f^2} \tag{8}$$

This is Friis equation [20]. To undertake calculations using this formula it is essential to use the correct units since power and gain are in (decibel-milliwatts) dBm. To put values into this equation it is simpler to take logs because the values for power and gain are quoted in decibel-milliwatts (dBm) which is a log value. This gives the equation

$$P_r = G_r + G_t + P_t + 20 \times \log_{10}\left(\frac{c}{4\pi rf}\right) \tag{9}$$

For Wi-Fi operating at 5 GHz at distance of 5 m the following values were used to calculate P_r the received power in dBm: $P_t = 20$ dBm, $G_r = G_t = 0$. The received power P_r was -37.9.

As a check the recommended power for reliable operation is greater than -65.

The simulation was run for a series of scenarios including: 2.4 GHz, 5 GHz at distances of 5–80 m in steps of 5 m.

Since the Wi-Fi signal is affected by materials in line of sight, a curtain was inserted 15 m and then replaced with a wall at 15 m these are plotted in Fig. 5. Additionally the value of minimum recommended P_r was plotted for reference (cutoff).

These findings were used to calculate the number of beds that a 5 GHz Wi-Fi would support from a distance point of view see Fig. 6.

By placing a Wi-Fi router in the center of the ward then the distance from the edges would $= 2.20 + 2 = 4.2$ m. Assume the height of the ward is 4 m then using Pythagoras, the max coverage $= 60$ m, so $X^2 = (60^2 - 4^2)$, $X^2 = (3600 - 16)$, $X = 59$ m in each direction. Each bed takes up 2.96 m and allowing for ends then it is possible to fit 18 bed stations per side.
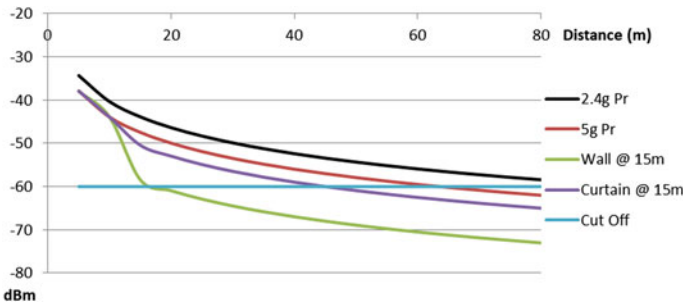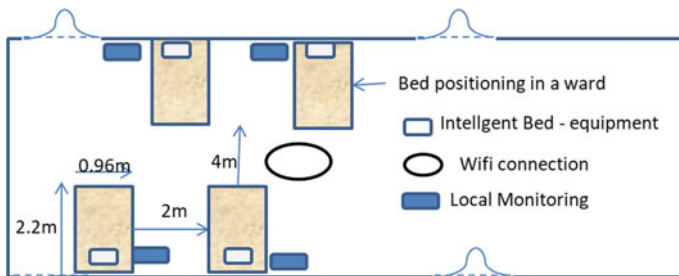

Fig. 5 Wi-Fi distance results

Fig. 6 Wi-Fi in hospital ward

5 Conclusions

The introduction of standard WBAN systems into the medical environment is one of the ways to improve healthcare monitoring in different environments. Usage of WBAN is a complex system to implement due to the separate technologies involved. However, the use of intelligent sensors and monitoring hubs supported by AWS cloud services can provide more accurate and up-to-date information about a patient's heath. The technology is now in a mature state and relatively low cost components can be used for the items like the sensors and Intelligent Bed.

The proposed intelligent monitoring hub has been constructed using Amazon services. Amazon IoT Greengrass is used to processes data from sensors and transmits it to the cloud. This provides an architecture that is appropriate for data collection period greater than 5 min, which is completely sufficient for patients, who are treated at home or visit doctor surgeries to check their health state. Such period is a function of the AWS Lambda operation features.

This solution is not recommended for more frequent data exchanges due to the possible data loss and low cost effectiveness. Continuous data streams, used for a patients in hospitals, requires updated solution, but still could be based on AWS IoT Greengrass. Using standard network protocols allow data from the sensors to be secured through to data storage. From a privacy point of view the data, once it is stored, is the sane as patient records.

References

1. Developing wireless body area networks standard. In: IEEE 802.15 WPAN™. Task Group 6 (TG6). Body area networks. http://www.ieee802.org/15/pub/TG6.html
2. Kurunathan JH (2015) Study and overview on WBAN under IEEE 802.15.6. U.Porto J Eng
3. Evans D (2011) The internet of things: how the next evolution of the internet is changing everything. CISCO White Pap 1:1–11
4. Islam SR, Kwak D, Kabir MH, Hossain M (2015) The internet of things for health care: a comprehensive survey. IEEE Access 3:678–708
5. Baker SB, Xiang W, Atkinson I (2017) Internet of things for smart healthcare: technologies challenges, and opportunities. IEEE Access 5:26521–26544
6. Hassan MK, El Desouky AI, Elghamrawy SM, Sarhan AM (2018) Intelligent hybrid remote patient-monitoring model with cloud-based framework for knowledge discovery. Comput Electr Eng 70:1034–1048
7. Sareen S, Sood SK, Gupta SK (2018) IoT-based cloud framework to control Ebola virus outbreak. J Ambient Intell Humaniz Comput 9:459–476
8. Catarinucci L et al (2015) An IoT-aware architecture for smart healthcare systems. IEEE Internet Things J 2:515–526
9. Wireless medical devices. In: U.S. Department of Health and Human Services (FDA) US Food & Drug Administration. https://www.fda.gov/MedicalDevices/DigitalHealth
10. Meditronic capsule endoscopy device PillCam™. https://www.medtronic.com
11. Keoh SL, Lupu E, Sloman M (2009) Securing body sensor networks: sensor association and key management. In: Pervasive computing and communications
12. Hamidi H, Fazeli K (2018) Using Internet of Things and biosensors technology for health applications. IET Wirel Sens Syst 8:260–267

13. Aizat Rahmat MA, Su ELM, Mohd Addi M, Yeong GluQo CF (2017) IoT-based non-invasive blood glucose monitoring. J Telecommun Electron Comput Eng:71–75
14. Plastic chips monitor body functions, research suggests. In: ScienceDaily. www.sciencedaily.com/releases
15. Hardiker N. Harmonising ICNP and SNOMED CT: a model for effective collaboration. In: Nursing informatics studies in health technology and informatics, vol 225, pp 744–745
16. De Raeve P. The world of cloud-based services: storing health data in the cloud. In: Heath Europa. https://www.healtheuropa.eu/cloud-based-services-storing-health-data-in-the-cloud/93053
17. AWS documentation. https://docs.aws.amazon.com
18. AWS official guide. https://aws.amazon.com
19. Verovko MV, Verovko OV, Kazymyr VV, Davies JN, Bhalla A. Body area networks in a medical environment. In: 2017 Internet technologies and applications (ITA), pp 92–97
20. Friis HT. A note on a simple transmission formula. IRE Proc 34(5):254–256

Architecture of Distributed Blockchain Based Intrusion Detecting System for SOHO Networks

Ivan Burmaka [ID]**, Mariia Dorosh** [ID]**, Igor Skiter** [ID]**, and Svitlana Lytvyn** [ID]

Abstract Intrusion detecting systems are widely used in corporate networks to protect them from intrusions, but SOHO networks often left being unprotected. One of the main reasons of this is that a corporate IDS are too expensive and complex for SOHO users. Opensource systems require good skills in setting up and configuring IDS and they work standalone without any collaboration. In this paper we propose the architecture of an intrusion detecting system which will be suitable for SOHO networks and will be easy to setup and configure. This architecture will also allow a collaborative intrusion detection by using information collected by nodes from other networks. Proposed architecture is based on a blockchain technology because it allows to setup trusted and secure data exchange between IDS nodes.

Keywords Intrusion detecting system · IDS architecture · IDS for SOHO networks · Decentralized IDS · Blockchain · Consensus protocol

1 Introduction

Cybersecurity specialists are constantly registering new attacks on computer networks. These attacks are becoming more advanced and complex every year, to resist them big companies are using an Intrusion detecting systems (IDS), which allows to decrease a negative impact of attack on a corporate network.

I. Burmaka (✉) · M. Dorosh · S. Lytvyn
Chernihiv Polytechnic National University, 95 Shevchenko str., 14035 Chernihiv, Ukraine
e-mail: Ivan.Bourmaka@stu.cn.ua

M. Dorosh
e-mail: mariyaya5536@gmail.com

S. Lytvyn
e-mail: chdtu.fld@gmail.com

I. Skiter
Institute for Safety Problems of Nuclear Power Plants National Academy of Science of Ukraine, Kyiv, Ukraine
e-mail: skiteris@ukr.net

313

But SOHO networks are more vulnerable for network attacks and intrusions, because in most of cases popular effective intrusion detecting systems are too expensive for such kind of networks. Very often there is no network security administrator, who can setup firewall and other security tools correctly. So, SOHO networks need a cheap, easy to setup and secure solution for intrusion detection and protection. And of course it must be compatible with most of standard hardware, so that it can be installed on a x64_86 compatible server or for a example a single board computer. Cosar et al. propose to use a Raspberry PI as an IDS node hardware, and it can be a good solution because the latest models of Raspberry Pi have a powerful ARM CPU and a big amount of RAM [1].

So our main goal is to built an architecture of IDS for a SOHO segment. The intrusion detection describes the process of monitoring network or system events for any sign of possible incidents [2]. An IDS is an application to realize the process of intrusion detection. Basically, an IDS can provide two main functions:

- Information recording. An IDS can monitor the target objects and record information locally, then the collected data can be sent to other facilities for analysis like a central event management system.
- Alert generation. The main task of an IDS is to generate alerts (alarms) to inform security administrators of important identified anomalies.

IDS errors can be divided into two kinds: false positive and false negative. The first type means that IDs can react on a normal network traffic as on the intrusion. False negative means passing intrusion like a normal traffic. Rates of this errors are an important measurement to decide whether an IDS is effective or not. So good IDS must allow a user to choose whether he wants to have a lower security rate and decrease number of false positive alerts or make a system as much secure as possible but with possible false positive alerts.

As mentioned, an IDS can be generally classified into HIDS and NIDS, whereas such classification can be more specific according to the deployed locations like a wireless-based IDS, which identifies malicious activities through monitoring wireless network packets and protocols. In practice, an IDS product often combines these two types of detection, as they can complement each other and provide a more thorough protection [3].

Of course it is possible to use some opensource intrusion detecting system like Snort, but such systems are a little bit hard to configure and adopt for new threads. So the solution of the IDS problem for the SOHO will be a distributed community based intrusion detecting system. The main idea of it is to use information collected by each participant to protect another participants by distributing information about attacks and a way to protect from them. But here we face a problem of setting a trust between the IDS nodes and the easiest way to solve a trust problem will be a blockchain technology.

The blockchain technology gives us a few advantages. First of all, from a blockchain definition we see that blockchain is a growing only, log like, data structure with cryptographically linked blocks, which means that data, which is stored in a blockchain cannot be modified by a malicious side so that it will be unnoticed. This

means that for example rules which are stored in a blockchain will be protected from a malicious modification. Another advantage is a high cryptographic protection, all records are signed with creator's key, so it is almost impossible to make a fake record for some blockchain participant.

2 Related Work

2.1 Intrusion Detection

The main problem of a standalone intrusion detecting modules is that such kind of IDS does not have any information about a protecting network. This means that intruders can use it to bypass the protection. It makes a great need for collaborative IDS or IDS network.

One of the earliest prototypes for a collaborative intrusion detection was a distributed intrusion detecting system (DIDS) [4]. It could utilize a distributed monitoring with a centralized data analysis to analyse the heterogeneous network. The next step in DIDS evolution was DOMINO architecture [5] which enables DNIDS deployed at diverse locations to securely share intrusion information between the heterogeneous nodes. But this solutions does not have full decentralization, data processing in this systems was centralized. And these solutions also don't solve the insider threat problem. But there were also a few tries to design a trust mechanism for IDS nodes with P2P technologies [6, 7].

Based on this observation, Fung et al. [8] proposed a challenge-based CIDN, where the trustworthiness of an IDS node depends on the received answers to the challenges. They first introduced a Host-based IDS framework that enables each HIDS to evaluate the trustworthiness of others based on its own experience and uses a forgetting factor to give more emphasis on the recent experience of each peer. In addition to this, Li et al. [9] identified that different IDS nodes may have different levels of sensitivity in detecting different types of intrusions. They proposed a notion of intrusion sensitivity (IS) that measures the detection sensitivity of the IDS in detecting different kinds of intrusions. Accordingly, they proposed an intrusion sensitivity-based trust management model [10] that could allocate the values of IS by means of a machine learning classifiers. Meng and Kwok proposed the idea of improvement of this idea by decreasing a false alarm level with using a method of voted ensemble selection [11].

But Li et al. [12, 13] found that there are few types of attacks which allow to compromise the challenge mechanism and send malicious feedbacks in some situations or to some specific nodes (Passive message fingerprint attack and special On-Off attack) [12].

2.2 Blockchain

A blockchain can be described as a distributed data structure, which is shared between the all participants of a network with a peer-to-peer architecture. The data structure itself is built of a back-linked list of blocks, where each block is identified by its cryptographic hash and also contains the hash of the previous block. This property allows to establish a cryptographic link between blocks, creating a so-called "blockchain" [14] that all participants can examine.

Regarding the control of these permissions, current blockchain implementations fall into three categories: public, consortium and private [15].

The creation of new blocks in a blockchain is regulated by a consensus protocol which orders how nodes are collaborating when creating blocks. This protocol may vary a lot and depends on both the type of the blockchain implementation and the threat model. To offer guaranteed security properties, public blockchains design the consensus part to be either computationally hard (Proof-of-Work) or based on the possession of a scarce resource within the system (Proof-Of-Stake). On the other hand, consortium and private blockchains apply some kind of Byzantine or benign fault tolerant algorithms, such as PBFT or SIEVE [16] to cope with malicious nodes.

Many studies have started researching the area of using a blockchain technology in distributed intrusion detecting systems. Alexopoulos et al. [17] described a framework of a blockchain-based CIDS, where they considered a set of raw alarms produced by each IDS as transactions in a blockchain, then all collaborating nodes employed a consensus protocol to ensure the validity of the transactions before delivering them in a block. This can guarantee that stored alerts are tamper resistant in the blockchain, but they did not implement and evaluate their method in practice.

Menget et al. say that the blockchain technology will help to improve the distributed intrusion detecting mostly in aspects of data sharing, alarm exchange and trust computation [3]. Kolokotronis et al. proposes the use of a trust-based blockchain in IDSs, referred to as trust-chain to protect the integrity of the information shared among the IDS peers, enhance their accountability and secure their collaboration by thwarting insider attacks [18]. Sharma et al. proposed DistBlockNet, a distributed secure software defined a networking architecture for the internet of things by integrating the blockchain technology, allowing a node to interact with others without the need of a trusted central controller [19]. The architecture of a signature-based IDS with blockchain for IoT also was proposed by Li et al. The idea of this solution is to use a blockchain to sets up the data exchange between the signature based intrusion detecting nodes [20]. But IoT is a very specific use case for IDS systems because IoT devices mostly have low computational power and specific software, so IoT experience cannot be fully ported to SOHO networks.

As we can see, the main role of a blockchain in distributed IDS is providing of a secure data exchange. So we need to take a look at blockchain data exchange solutions, because IDS can operate with a huge amounts of data in case of using a lot of monitoring nodes. Li et al. propose a distributed data storage scheme for IoT environment employing blockchain and cetrificateless cryptography with distributed hash tables[13], which allows to organize a flexible access right management. Another blockchain based data storage solution was proposed by Ren et al. where they use Merkle trees to securely link data blocks to a blockchain [21]. Distributed hash tables were also used here for an access control. Mao et al. also showed that a Merkle tree is a good way for data integrity verification for cloud storages.

3   Architecture of a Blockchain Based IDS for SOHOs

The system being designed by us must have the following characteristics to be suitable for SOHO networks:

- Fully distributed. It is important characteristic in our case, because SOHO networks mostly provide cheap solutions with average reliability which means that any time any node can stop working, but the whole system must continue working. The system also must not have hard centralization, because central nodes are a single point of a failure for the whole system. The distribution will make the system more robust and reliable.
- Effective intrusion detection is also an important parameter, because if the system cannot protect from most known attacks it is useless. So good IDS needs periodical signatures updates which can help to protect from the newest threats and intrusions. Also good IDS need to have heuristics modules which can help to find a new and unknown thread and create signatures for them.
- Collaborative intrusion detection related to the effective detection and means that each node uses information from other nodes for detecting and preventing intrusions. For example information about heuristically detected intrusion can be distributed between other nodes and then all nodes will detect such kind of intrusions without using a lot of computational resources.
- Easy configuration is very important for SOHO, where people need a system which will be easy to setup and configure, because in SOHO networks administrators mostly do not know how to correctly and securely configure the advanced IDS. So the best solution for such kind of networks will be the IDS which is installed with the optimal default settings and the administrator can choose which services are used in this network and the level of detection sensitivity. Of course the system must have possibility for an advanced configuration for those who can tweak settings manually.
- Platform independence that is the system must work on the general purpose hardware and be compatible with as much hardware as possible, because a specific hardware can be too expensive for SOHO.
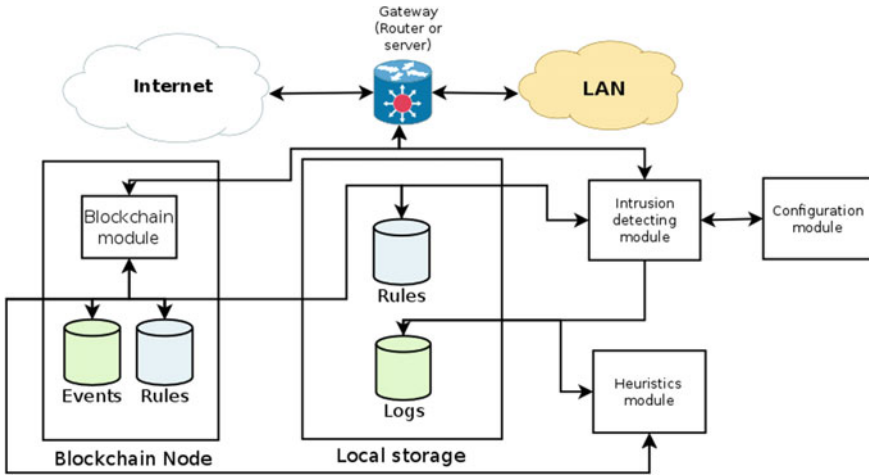
Fig. 1 Architecture of distributed blockchain based intrusion detecting system

The most important components of every intrusion detecting system are monitoring module, data processing module and alarm module. Our system, of course, will also have these components, but it will be extended with a blockchain related and collaborative modules (Fig. 1).

So a blockchain part of our IDS is responsive for trusted and secure communication between IDS nodes via the global (or local) network. As in the most of blockchains, in our IDS each node is connected with some limited number of other nodes and can send and receive messages to all connected nodes. The connections are chosen so that all nodes form an enclosed network where any node can send a message and it will be delivered to all other nodes.

The blockchain module is setting up connections with other IDS nodes and syncing chain of blocks with IDS rules and important events which are used for the intrusion detection and heuristics. The rules which are distributed between IDS nodes with a blockchain are then stored in a local rules database then the set of rules is processed with a configuration module component which selects from our set of rules only the rules which are related to the services which are used on a current node and general rules which are used for all nodes.

The intrusion detecting module reads the rules form a local database and uses them for the intrusion detection and prevention, and writes events to a log. Important events are distributed with a blockchain module, because such events can help to determine some unknown attacks with a heuristics module.

Then information from IDS module is processed by a heuristics module, and if it founds some suspicious activity, it creates a rule for it and sends it to a blockchain module for distribution between other nodes. Heuristically created rules will be marked as the rules for a high sensitivity (before a manual rules check) detection to decrease a level of false positive alerts.

4   Blockchain Structure

As we can see, IDS needs a lot of information to work with. First of all it is a set of rules which the intrusion detecting module needs for quick traffic processing. This set of data is only increasing most of the time, so it looks good to use a blockchain for saving and distributing a set of rules.

Another kind of information for a collaborative distributed IDS is an event log. The event log is also increasing all the time, so it also can be distributed via the blockchain. But if IDS rules have an extremely small size (for example the set of SNORT community rules is less than 500 kilobytes) and they can be easily placed to a blockchain as block data, when we are trying to store the events log, we face a problem that a size of event logs is increasing very quickly, which will not be good for a blockchain storage and in most of cases events the log information is temporary and some nodes may not want to keep it saved forever. For example single-board nodes possibly will not have such a lot of storage space. This means that we need somehow to store the temporary information, so that the node can choose it to store the information for a long time on not. The simplest way to do this is to move the information outside the blockchain structure and use a provable data possession (PDP) for checking data integrity [21]. PDP mechanism is mostly used in P2P networks for data integrity verification. First PDP was based on a probabilistic strategy to complete the integrity verification, using the homomorphic properties of the RSA signature mechanism to aggregate the evidence into a small value, greatly reducing the communication overhead of the protocol. But now the most common PDP use the Merkle hash tree [21]. In our case this tree also helps to link a data with a blockchain block. Figure 2 illustrates a blockchain block with Merkle tree structure. Merkle hash tree is a popular technique for data integrity checking. A Merkle hash tree is a tree in which every non-leaf node is labeled as the hash value of its children nodes, and every leaf node is labeled as the hash value of a data block. There is a root on the top of the Merkle hash tree. One of the most common hash algorithm for Merkle tree is a Tiger hash, which is well optimized for 64 bit processors and do not have such vulnerabilities as MD5 [22]. Using a Merkle tree will also help to search for data between the nodes quickly, because any node can quickly announce just available roots and other nodes will see if this node has a data they need. Of course such idea of a temporary event log storage will cause in future such situation when old event logs will be lost forever after some period of time, but in most of cases the older event log becomes less useless, because it really uses a lot of computational resources to analyze a huge amounts of logs for a long time and log information will become irrelevant after some period of time.
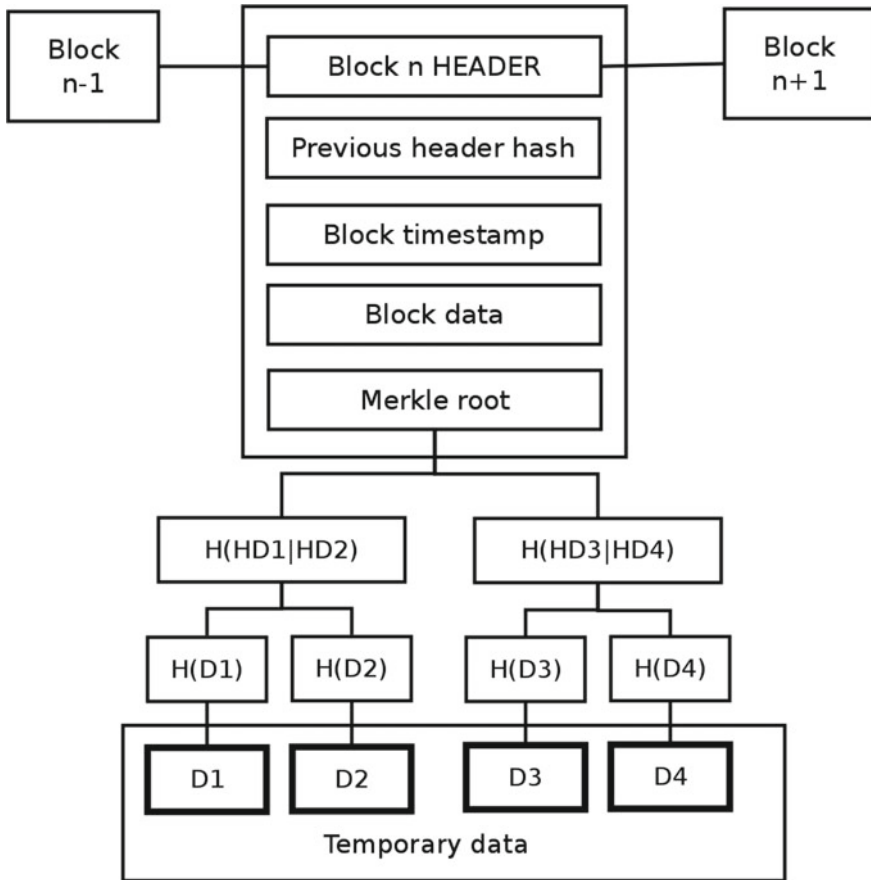
Fig. 2 Blockchain block structure with linked Merkle tree

5 Consensus Protocol

The early version of a blockchain, such as Bitcoin, is implemented by a clever mechanism called Proof of Work (PoW). In such a system, each block contains a list of transactions, a random salt and the hash of the previous block. A new blockchain be created only if a random salt is found. Finding a valid salt is a crypto puzzle that needs lots of computational power and the computing process is called Mining. Once a crypto puzzle is solved, a new block is created and the transactions can be written into this new block. The smart idea behind PoW is that the transactions in the system are performed and audited by a large group of miners who take great efforts to run this system. Therefore, it is reasonable to assume that the system is secure as long as the majority of miners are honest. However, this scheme does not fit our structure because of high energy consumption and a low transaction speed.

For example, the most famous cryptocurrency system using PoW is the Bitcoin, it is only able to process seven transactions per second [23].

To tackle the power-consumption issues of the Bitcoin, other mechanisms have been explored to replace PoW such as Proof of Stake (PoS), Proof of Space (PoSpace) and Rem (Intel SGX), etc. [24–26].

Li et al. assert that the Rem algorithm is the most suitable for systems with low-end hardware. The idea of "Rem" is to replace the power-consuming "Proof of Work (PoW)" with "Proof of Useful Work (PoUW)" where the miners provide trustworthy reports on CPU cycles devoted to inherently useful workloads [27]. But the main problem here is that this algorithm is based on Intel's SoftwareGuard extensions "SGX", which permits trustworthy code to be executed in an isolated and tamper-free environment and SGX can prove remotely the result of such executions. This means that this algorithm is platform dependent and can work only on Intel processors. So, we think that the best consensus algorithm here will be PoS like algorithm. 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 sytem 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 woks 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 consensus 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 the 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 is based on a current blocks queue in a memory pool.

A maximum staked value is limited with a modular arithmetic, so a staked value will be reset periodically for nodes which do not use their stake. This will help to prevent the situation when the malicious node will stake a high value and will have almost 100% probability to generate a block. The simulation model shows that this consensus protocol allows to achieve a stable block creation [28]. Such kind of a consensus protocol does not protect from malicious nodes appearing (if we have a permissionless blockchain), but there are a few methods which allow to ban malicious nodes so that they cannot take part in the block creation and lose all stake. This will make attack tries inconvenient for an attacker.
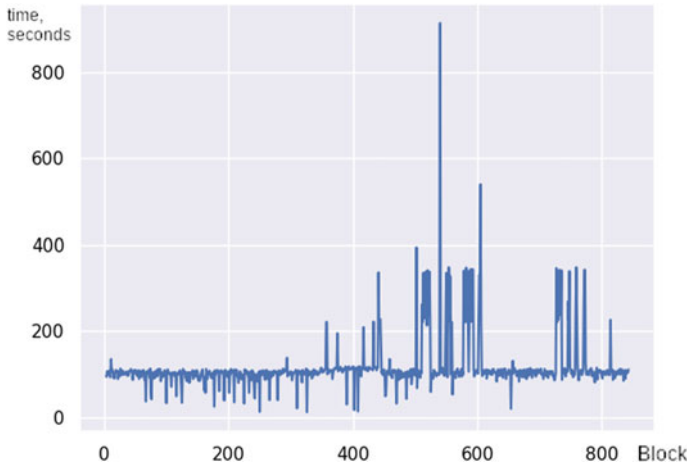
Fig. 3 Block generating time for blocks
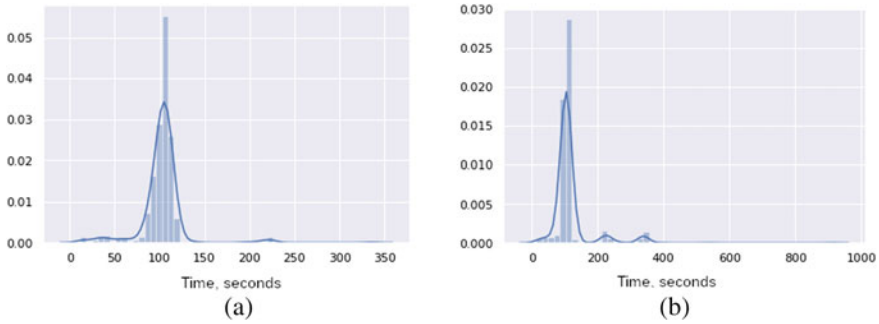
Fig. 4 Block generating time distribution for **a** two nodes, **b** three nodes

Test implementation of such kind of the consensus protocol shows that such protocol can provide a stable block generating process. Two and three nodes configuration were tested with a time module of 120 s. Such testing shows that the time used to generate one block in most cases is less than 120 s. Testing results are shown on Fig. 3.

The distribution shows that the increasing number of nodes can increase the time used to generate a block because of the conflicting blocks in a chain, as we can se on the plots at Fig. 4.

So, such consensus protocol will be convenient for a distributed intrusion detecting system because of optimal CPU usage and a stable block generating process.

6 Rules Management and Configuration

Rules database with the intrusion signatures are one of the most important part of the most intrusion detecting systems, because the signature based detection is the fastest method of the intrusion detection and it does not require a lot of computational resources. This means that the signature based detection can be used even on a low-end hardware. Signature based methods also have a low false positive error rate in compare to the anomaly based detection. But rules cannot work standalone, because a network traffic needs preprocessing before it can be compared to the signatures. Most of protocols have their specific preprocessing algorithms. For example Snort IDS has for such purpose dynamic processing modules and these modules are much more complex than just a set of text rules.

But to be effective the signature based detection system needs to be proper configured. We propose to use a simplified configuration method to make the system easy to configure for almost everyone. First of all we propose to sort rules database by a protocol, so that a user can choose only the protocols that his system uses. This can help to decrease the used rules database size. It also will help to disable unused dynamic traffic processors so it also will decrease the CPU load for traffic processing.

The rules can also be divided into high sensitivity rules, for users who need a maximum security level, but the other side of these rules is that such kind of rules cause a higher false positive error rate, by marking some normal activity as suspicious. The rules with a lower sensitivity rate are more suitable for most of users but these kind of rules can miss some suspicious activity.

If we just disable rules for unused services system they will have vulnerabilities. So to make system safe, we need to close all ports which can be related to disabled services. This will decrease the probability tat someone will use some vulnerabilities of services which a user forgets to disable. Another step to make the system secure is to use a set of rules which will detect the requests to the disabled services even if their ports are closed. So, for example, if someone tries to access HTTP service on a protected server but HTTP was disabled in IDS configuration, then ports will be blocked and IDS will report the suspicious event about a probable attack. Such IDS behavior can help to detect those attackers who are scanning for specific services and their vulnerabilities. For example, if distributed IDS finds that one IP address is trying to access one service on many protected servers it can at least blacklist such address and even report the appropriate ISP about a malicious activity.

And here we come to a collaborative part in rules management. Firstly it is related to heuristically created rules, based on the list of suspicious events. This means that the list of suspicious events which is distributed with a blockchain can be processed by all nodes for public events (or just by a small list of trusted nodes for private events). So, the node which collects enough information to automatically create the detection rule creates it and makes a transaction to a blockchain with a new rule. Such automatically created rules are marked as high sensitivity rules before the manual check by cybersecurity specialists to prevent the increasing of a high false-positive error rate.

7   Conclusion and Future Work

This paper has proposed the architecture of a collaborative distributed intrusion detecting system for SOHO networks. Using blockchain in this architecture for the data exchange allows setting trust between the nodes of a distributed IDS without using a central trusted node, which makes the system easy to scale and deploy. Using a blockchain technology also allows to protect data exchange from spoofing and attacks like MITM. A blockchain here also is used for data verification with Merkle tree technology, which also allows making system more secure and protects from using damaged data. Such architecture will be effective for the collaborative intrusion detection in SOHO because it allows to join multiple networks to help each other with the intrusion detection.

There are a few directions for future work. First and the main direction for future work is an implementation of the proposed IDS architecture and testing its performance and security in the artificial environment. If testing shows that the system is suitable for protecting SOHO networks it can be tested on a real network and future work will be related to improving the performance and security of such kind of IDS.

References

1. Coşar M, Karasartova S (2017) A firewall application on SOHO networks with Raspberry Pi and snort. In: 2017 international conference on computer science and engineering (UBMK), pp 1000–1003
2. Scarfone K, Mell P (2012) Guide to intrusion detection and prevention systems (IDPS). Tech. Rep. NIST Special Publication (SP) 800-94 Rev. 1 (Draft), National Institute of Standards and Technology. https://csrc.nist.gov/publications/detail/sp/800-94/rev-1/draft
3. Meng W, Tischhauser EW, Wang Q, Wang Y, Han J (2018) When intrusion detection meets blockchain technology: a review. IEEE Access 6:10179–10188, conference Name: IEEE Access
4. Snapp SR, Brentano J, Dias GV, Goan TL, Heberlein LT, Ho Cl, Levitt KN, Mukherjee B, Smaha SE, Grance T, Teal DM, Mansur D (1991) DIDS (Distributed Intrusion Detection System)—motivation, architecture, and an early prototype. In: In Proceedings of the 14th national computer security conference, pp 167–176
5. Yegneswaran V, Barford P, Jha S Global intrusion detection in the DOMINO overlay system. Technical Report, University of Wisconsin-Madison Department of Computer Sciences. https://minds.wisconsin.edu/handle/1793/60340. Accepted 2012-03-15T17:16:57Z
6. Duma C, Karresand M, Shahmehri N, Caronni G (2006) A trust-aware, P2P-based overlay for intrusion detection. In: 17th international workshop on database and expert systems applications (DEXA'06), pp 692–697. ISSN: 2378-3915
7. Tuan TA (2006) A game-theoretic analysis of trust management in p2p systems. In: 2006 first international conference on communications and electronics, pp 130–134. IEEE
8. Fung CJ, Baysal O, Zhang J, Aib I, Boutaba R (2008) Trust management for host-based collaborative intrusion detection. In: De Turck F, Kellerer W, Kormentzas G (eds) Managing large-scale service deployment. Lecture notes in computer science. Springer, Berlin, Heidelberg, pp 109–122

9. Li W, Meng Y, Kwok LF (2013) Enhancing trust evaluation using intrusion sensitivity in collaborative intrusion detection networks: feasibility and challenges. In: 2013 ninth international conference on computational intelligence and security, pp 518–522

10. Li W, Meng W, Kwok LF (2014) Design of Intrusion Sensitivity-Based Trust Management Model for Collaborative Intrusion Detection Networks. In: Zhou J, Gal-Oz N, Zhang J, Gudes E (eds) Trust Management VIII. IFIP Advances in Information and Communication Technology, Springer, Berlin, Heidelberg, pp 61–76

11. Enhancing false alarm reduction using voted ensemble selection in intrusion detection. Int J Comput Intell Syst 6(4). https://www.tandfonline.com/doi/abs/10.1080/18756891.2013.802114

12. Li W, Meng W, Kwok LF (2018) Investigating the influence of special on-off attacks on challenge-based collaborative intrusion detection networks. Future Internet 10(1):6. https://www.mdpi.com/1999-5903/10/1/6, number: 1 Publisher: Multidisciplinary Digital Publishing Institute

13. Li W, Meng W, Kwok LF, Ip HHS (2016) PMFA: toward passive message fingerprint attacks on challenge-based collaborative intrusion detection networks. In: Chen J, Piuri V, Su C, Yung M (eds) Network and system security. Lecture notes in computer science. Springer, Cham, pp 433–449

14. Antonopoulos AM (2014) Mastering bitcoin: unlocking digital cryptocurrencies. O'Reilly Media, Inc. google-Books-ID: IXmrBQAAQBAJ

15. Okada H, Yamasaki S, Bracamonte V (2017) Proposed classification of blockchains based on authority and incentive dimensions. In: 2017 19th international conference on advanced communication technology (ICACT), pp 593–597

16. Cachin C, Schubert S, Vukolić M (2016) Non-determinism in Byzantine fault-tolerant replication. arXiv:1603.07351 [cs]

17. Alexopoulos N, Vasilomanolakis E, Ivánkó NR, Mühlhäuser M (2018) Towards blockchain-based collaborative intrusion detection systems. In: D'Agostino G, Scala A (eds) Critical information infrastructures security. Lecture notes in computer science. Springer, Cham, pp 107–118

18. Kolokotronis N, Brotsis S, Germanos G, Vassilakis C, Shiaeles S (2019) On blockchain architectures for trust-based collaborative intrusion detection. In: 2019 IEEE world congress on services (SERVICES), pp 21–28. ISSN: 2642-939X

19. Sharma PK, Singh S, Jeong YS, Park JH (2017) DistBlockNet: a distributed blockchains-based secure SDN architecture for IoT networks. IEEE Commun Mag 55(9):78–85. Conference Name: IEEE Communications Magazine

20. Li W, Tug S, Meng W, Wang Y (2019) Designing collaborative blockchained signature-based intrusion detection in IoT environments. Future Gener Comput Syst 96:481–489. http://www.sciencedirect.com/science/article/pii/S0167739X18327237

21. Ren Y, Liu Y, Ji S, Sangaiah AK, Wang J (2018) Incentive mechanism of data storage based on blockchain for wireless sensor networks. https://www.hindawi.com/journals/misy/2018/6874158/. ISSN: 1574-017X. Library Catalog: www.hindawi.com. Pages: e6874158. Publisher: Hindawi Volume: 2018

22. Mao J, Zhang Y, Li P, Li T, Wu Q, Liu J (2017) A position-aware Merkle tree for dynamic cloud data integrity verification. Soft Comput 21(8):2151–2164. https://doi.org/10.1007/s00500-015-1918-8

23. Scalability—Bitcoin Wiki. https://en.bitcoin.it/wiki/Scalability

24. Dziembowski S, Faust S, Kolmogorov V, Pietrzak K (2015) Proofs of space. In: Gennaro R, Robshaw M (eds) Advances in cryptology–CRYPTO 2015. Lecture notes in computer science, Springer, Berlin, Heidelberg, pp 585–605

25. Kiayias A, Russell A, David B, Oliynykov R (2017) Ouroboros: a provably secure proof-of-stake blockchain protocol. In: Katz J, Shacham H (eds) Advances in cryptology—CRYPTO 2017. Lecture notes in computer science. Springer, Cham, pp 357–388

26. Zhang F, Eyal I, Escriva R, Juels A, Renesse Rv (2017) REM: resource-efficient mining for blockchains, pp 1427–1444. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/zhang

27. Li R, Song T, Mei B, Li H, Cheng X, Sun L (2019) Blockchain for large-scale internet of things data storage and protection. IEEE Trans Serv Comput 12(5):762–771. Conference Name: IEEE Transactions on Services Computing
28. Burmaka I, Stoianov N, Lytvynov V, Dorosh M, Lytvyn S (2020) Proof of stake for blockchain based distributed intrusion detecting system. In: International scientific-practical conference. Springer, pp 237–247

Models of Social Behavior of Learning Material Acceptance in Science-Centered Approach in Education

Volodymyr Shevchenko [ID]**, Denys Berestov** [ID]**, Igor Sinitsyn** [ID]**, and Maksym Brazhenenko** [ID]

Abstract The paper considers methods of mathematical modeling based on the use of the theory of cellular automata as the main tool for modeling the processes of public opinion formation. As the object of modeling was chosen social behavior in the context of the acceptance of educational material in the science-centered approach in education. The theory of cellular automata was chosen as a modeling tool because it is a relatively simple and at the same time effective method for modeling the interaction of single-type objects. The developed basic cellular automata of the model has extended rules of determination of cell state and determination of cell vicinity. It has also been derived the dependence of the information perception coefficient on the difference between the state of the cell and its surroundings. These improvements allow the model to be used to predict changes in social group preferences. At the same time, the model was implemented in two programming languages Python and MatLab, which allows to choose the simulation environment when the peculiarities of the problem statement change.

Keywords Cellular automata · Mathematical modeling · Social behavior

1   Introduction

1.1   *Actuality of the Work*

In an ever-changing world, each person needs to ensure that he or she is in demand as a specialist in his or her field of work [1, 2]. To ensure this it is necessary to have fast learning skills and an objective picture of the world. Higher education allows you to get the above things. A quality university education consists of many aspects: the study of the educational program, social activity, and scientific activity

V. Shevchenko (✉) · D. Berestov · M. Brazhenenko
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
e-mail: vladimir_337@ukr.net

I. Sinitsyn
Institute of Software Systems, National Academy of Sciences of Ukraine, Kyiv, Ukraine

[3]. In this paper, we will refer to the science-centered approach to education as one in which science serves as a driving force, encouraging the student to pay more attention to the subjects of the program and to go beyond the program to deepen their knowledge in their major. Properly instilled in the student scientific approach allows further development of non-standard innovative solutions. But forming exactly the scientific approach in the student is a difficult process because many factors that prevent its formation. It is useless to force to do science. Therefore, students are attracted to science through the belief that it is useful for their professional growth. That is, to a certain extent there is a place of agitation.

Therefore, it is relevant to ensure the effective formation of the student's understanding of the need to participate in scientific activities because of the influence of teachers. In accordance with this, it is also relevant to develop a model of the learning process in a science-centered approach to predict the consequences of management decisions about the agitation work with students.

With the help of models based on cellular automata, it is relatively easy to simulate processes that involve the same-type objects of the real world. Therefore, it was proposed to use the theory of cellular automata to model the dissemination of information in society.

1.2 Purpose of Work

To develop methods for predicting changes in public opinion through the development of a model of social behavior on the acceptance of educational material in a science-centered approach to education.

2 Analysis of Existing Developments and Problem Statement

2.1 Analysis of Existing Developments

Research on the theory of cellular automata (CA) has been conducted since the end of the 1940s when the idea of CA was "on the surface" [4]. Therefore, the creation of this theory is attributed to four scientists, who worked on its ideas independently and simultaneously. John von Neumann was working on the theory of self-replicating machines, and his laboratory colleague S. Ulam was developing a mathematical model of crystal growth. The exchange of ideas led to the creation of a CA model of systems evolution. Around the same time, N. Wiener and A. Rosenbluth developed a CA model of the excitable medium to describe the propagation of impulses in neural networks. Further, the CA theory gained popularity in the 1970s, when ordinary amateurs became interested in it. In particular, John Conway made an important

contribution when he developed the CA "Game of Life". Conway was interested in the problem proposed by Neumann, a model of a self-reproducing machine [5]. Conway simplified Neumann's ideas and created the "Game of Life" rules. This CA and its modifications have influenced several branches of mathematics, computer science, and physics.

In the context of behavioral models, it is necessary to mention the main modeling approaches: dynamic models based on systems of differential equations, hybrid models based on differential and algebraic equations. Algebraic equations are usually sufficient to predict the economic consequences of management decisions [6]. But the considered problem formulations rely on the averaged characteristics of the systems, which do not take into account the topology of the systems under study [7]. Models based on CA can provide new opportunities, especially when it is necessary to take into account both statistics and topology of the network. The peculiarity of the theory of CA is that it can be also used to model and describe a large number of phenomena [8, 9], for example, to simulate the spread of forest fires [10], car traffic on one lane [11], quantum world [12] and, importantly in the context of our study, in epidemics in society [6, 13, 14]. However, there are few studies directly related to the use of CA for societal behavior in the context of educational processes, in particular at the university. This generates **contradictions** between the peculiarities of modeling problems with a large number of parameters and the narrow set of control parameters of existing models like Conway's "Game of Life".

2.2 General Concepts of the Theory of Cellular Automata

All CA which will be considered further are guided by the following rules:

1. Each CA consists of a cellular field containing cells that can be in predefined discrete (e.g., 0 or 1) or non-discrete states (e.g., a value between 0 and 1).
2. The cell field can be represented as an oriented or undirected graph, or it can take the form of a checkerboard, which can also be closed according to some rules.
3. Changing the states of the cells of a CA is guided by a pre-defined list of rules.
4. The vicinity of a cell is the set of cells, on which the state of the sought cell at the next iteration of the automata depends.
5. The state of each cell of the automata is redefined at each next iteration of the CA in accordance with the rules for changing the state of cells.

The considered CA without further improvements is a rather abstract model, for which the life situation is quite difficult to find, so it is relevant to improve it in the direction of introducing the necessary set of rules for CA development.

Fig. 1 Cell field with an improved rule of possible cell states
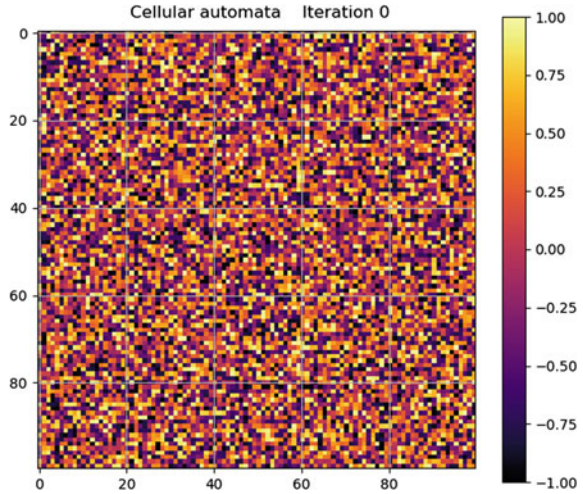
3 Improvement of the General Rules of the Cellular Automata

3.1 Improvement of the Rule of Possible Cell States

In most CA, cells can accept only well-specified discrete values, which is not appropriate in the situation of thought dissemination in society. In this case, the use of non-discrete values is more appropriate. Consider further that a cell's opinion can take values between -1 and 1, where the extreme limits correspond to diametrically opposite points of view on some issue. In this case, the cell field will be represented as a colored cell field (Fig. 1). The lighter the cell on the field, the closer its value is to 1.

3.2 Improvement of the Rule for Determining the State of a Cell

Since the object of our research (society) is somehow a closed system and it is difficult (in some cases even impossible) to distinguish objects of the society that are at the line, we will consider the cell field as closed, which topologically forms a torus form.

To calculate the number of neighbors we use the rule, which was deduced by Moore [11], that is, each cell will have 8 neighbors in its vicinity. Since a person's opinion is to some extent influenced by the opinions of the people communicating with him, the state of the cell at iteration $t + 1$ can be represented as a function of 9 variables, where the first variable corresponds to the state of the underlying cell, and
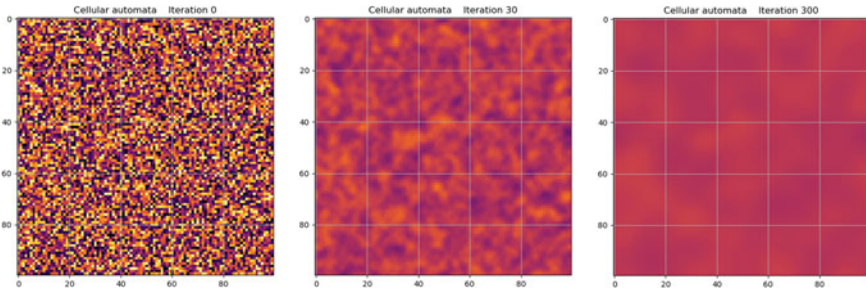
Fig. 2 CA with an improved cell state determination rule

the other 8 variables are the states of the neighbors. We propose this kind of function:

$$s_{t+1} = s_t + \left(\frac{\sum_{i=1}^{n} n_i}{n} - s_t \right) * k \tag{1}$$

where n—the number of neighboring cells, n_i—the state of the i-th cell from the environment of cell s_t, s_t—the state of the cell at iteration t, s_{t+1}—the state of the cell at iteration $t + 1$, k—coefficient of susceptibility to another's opinion.

The susceptibility coefficient can be preset by the user or depend on the preset parameters. The implementation of CA with the implemented rule to determine the state of the cell is shown in Fig. 2. In general, CA tends to the state of equilibrium (society's coming to common opinion). It is connected with the absence of external factors, which could influence the process of formation of a thought.

3.3 Introduction of the Basic Rule for Determining the Vicinity of a Cell

In classical CAs [5], the state of a cell depends on the states of the cells in its vicinity, namely 8 neighboring cells. This makes it impossible to model systems where objects can contact different number of objects. Therefore, it was proposed to introduce a basic rule for determining the neighborhood of a cell, which allows the simulation to set the necessary neighborhood for the cells of the CA (interaction radius). For example, radius 2 allows having 24 neighboring cells. The number of cells in the vicinity at an arbitrary radius is determined by the formula:

$$n = (2R + 1)^2 - 1$$

where R—radius of the intrapopulation interaction.

3.4 The Complicated Rule for Determining the Vicinity of a Cell

The basic rule of determining the cell neighborhood can be complicated by introducing a "terrain map", which is a set of interaction radius values for each cell of the automata field. The values for each cell can be set randomly or by the user. This will make it possible to simulate a social situation in which different objects interact with different numbers of objects.

To test the workability of the complicated rule for determining the neighborhood of the cell field, the cells were divided into four sectors, in each of which the cells share the same interaction radius. The division of the field is depicted in Fig. 3. An example of work of the CA with the complicated rule of the definition of a cell vicinity is shown in Fig. 4.

At iteration 30 of the CA (Fig. 4) we can see clear boundaries of established sectors (Fig. 3). Depending on the radius of interaction, cells of sectors will change their values with different speeds, with sectors with $R = 3$ and $R = 4$ heading to a state of equilibrium much faster than sectors with radius $R = 2$ and $R = 1$. From this, we can draw the intermediate conclusion that the breadth of the social circle strongly influences the formation of opinion in society.

Fig. 3 Dividing the field into sectors

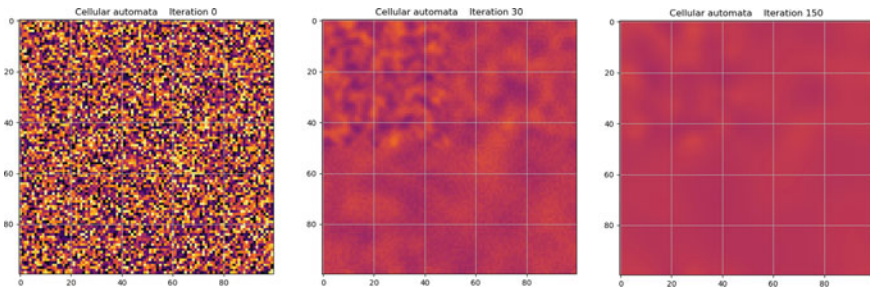| $R = 1$	$R = 2$
$R = 3$	$R = 4$

Fig. 4 CA with the complicated rule of determining the vicinity of a cell

4 Improvement of the Cellular Automata in the Direction of Modeling of Social Behavior

4.1 The Rule of Dependence of the Susceptibility Coefficient on the Difference Between the State of a Cell and Its Environment

Human life, for the most part, consists of choices about different developments: to go or not to go somewhere; to buy or not to buy something; to take part or not to take part in some process. When choices are not life-threatening, a person tends to choose what most of her environment accepts. If the opinion one holds does not coincide with the opinion of others, one is reluctant to change one's opinion under the influence of others. At the same time, when the opinion of the environment is similar to his own, he strengthens the existing opinion even more. Thus a clear dependence of sensitivity to another's opinion on the difference with one's own opinion emerges.

This dependence can be taken into account and implemented through the introduction of the rule of dependence of the susceptibility coefficient on the difference in the state and in the cell and its environment. The corresponding value of the coefficient k can be calculated using a modified logistic function:

$$k(\Delta) = \frac{L}{1 + e^{-h*(|\Delta|-x_0)}} \tag{2}$$

where Δ is the difference between the value of the cell state and the environment, L—the maximal possible value of the coefficient k, x_0—point of the middle of the logistic curve, h is the rate of growth of the logistic curve.

The value of the difference between the value of the cell state and the environment is calculated by the formula:

$$\Delta = \frac{\sum_{i=1}^{n} n_i}{n} - s_t \tag{3}$$

where n is the number of neighboring cells, n_i is the state of the i-th cell from the environment of cell s_t, s_t is the state of the cell at iteration t.

The dependence plot of the susceptibility coefficient based on the logistic function is shown in Fig. 5. The parameters of the proposed function (Fig. 5) take values $L = 1$, $x_0 = 0.5$, $h = -15$. In any case, when changing the formulation of the problem, the coefficients can be changed to fit the peculiarities of the modeling object.

Thus, formula (1) in the rule of determination of the cell state can be combined with the proposed rule of dependence of susceptibility coefficient on the cell state in formula (2) and (3):
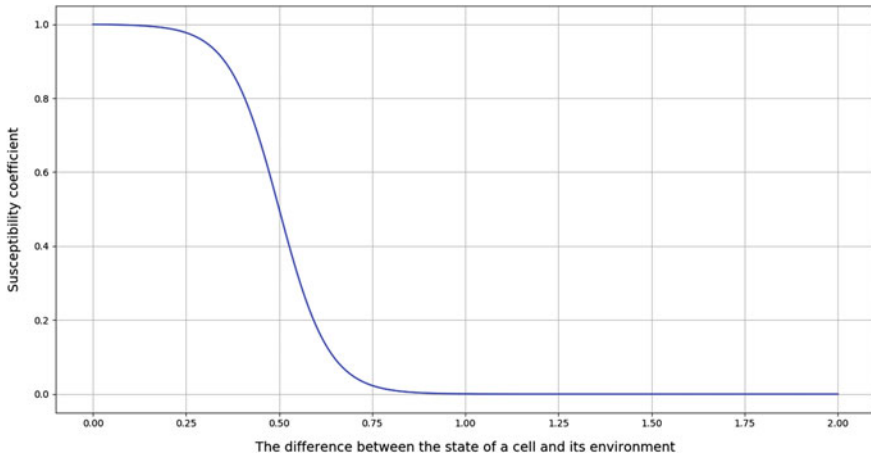
Fig. 5 The graph of susceptibility coefficient dependence on the state of a cell and its environment

$$s_{t+1} = s_t + \Delta * \frac{L}{1 + e^{-h*(|\Delta|-x_0)}}$$

An example of CA operation with the introduced rule of dependence of the susceptibility coefficient on the difference between the states of the cell and its environment is shown in Fig. 6. It is necessary to note, that in contrast to previous models (Fig. 4) in the given one (Fig. 6) one can observe the formation of separate paired epicenters of cells with large values (from 0.75 to 1) and opposite values (from −1 to −0.75). This is related to the appearance of the cells' ability to maintain their own opinion for some time despite the "hostile" environment. This manifestation corresponds to concrete cases of real life, so the new rule is appropriate when modeling the behavior of society.
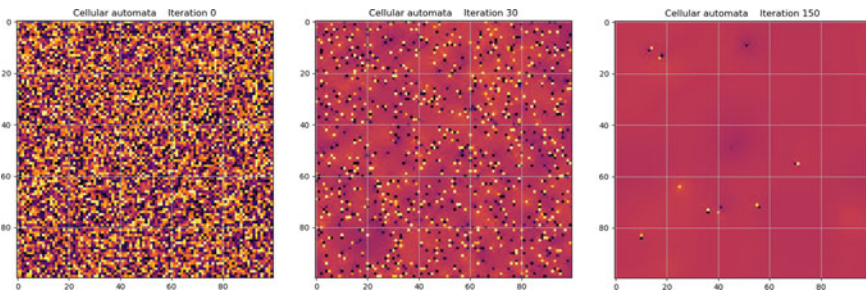

Fig. 6 CA with the introduced rule of dependence of the susceptibility coefficient on the difference between the states of the cell and its environment

4.2 The Rule of Cyclic Variation of the Radiuses of Field Cell Interactions

In real life, people are constantly on the move, every day a person is in at least two different circles of communication: home and work. Similar processes take place during the educational process: there are general lectures, separate seminars, and practical classes. Thus, students are constantly changing their social circles. At the same time, the difference in the performance of individual groups is often due to the difference in the initial preparation of students for learning, students' activity, and the overall favorable environment in the group. The first two factors in the model are taken into account in the formation of the initial cellular field, while the latter was not taken into account before.

Therefore, it was proposed to introduce a rule of cyclic change of cell field interaction radiuses, which allows setting several "terrain maps" of interaction radiuses for its subsequent cyclic change during the automata operation. To check the rule work let's set cyclic variation of interaction radiuses with shift by 2 values, that is the cell with radius $R = 4$ of interaction changes it to $R = 2$. One cycle will last for 2 iterations of the automata.

An example of CA work with the introduced rule of cyclic change of interaction radiuses of cells of the field is shown in Fig. 7. After the rule was introduced, the model changed its behavior: more individual paired cell epicenters with large (0.75 to 1) and (-1 to -0.75) small values are formed on the field. This is due precisely to the fact that cells communicate with different numbers of surrounding cells when the interaction radii change cyclically. Under individual group conditions, this means that another group's view of an issue has less impact on the group under study and is less likely to have people with an opinion that is cardinally different from the others.
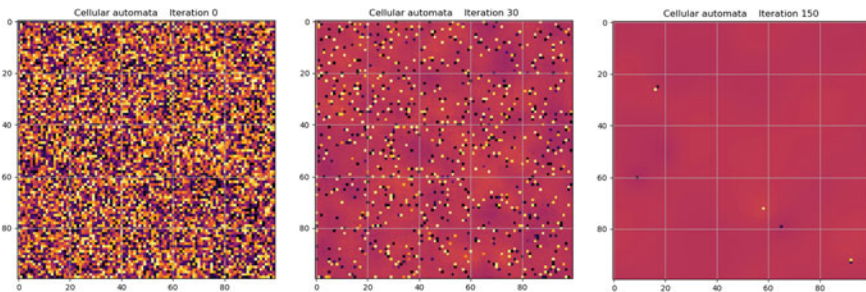

Fig. 7 Implementation of the rule of cyclic variation of the radiuses of field cell interactions

5 Approbation of Modified Cellular Automata in Modeling the Learning Process in the Science-Centered Approach

For modeling of processes of thought formation in a society the mathematical and algorithmic formalization which has been perfected above on simpler models with changing conditions of change of a condition of cells, definition of periphery of cells, dependence of a factor of susceptibility from a difference of states of a cell and its environment and cyclic change of radiuses of the interaction of field cells has been used in work. Thus, the abstract CA was adapted to the tasks of modeling the processes of thought formation in society. For further approbation, let us clarify the problem statement, terminology, and content of states and processes.

5.1 Formulation of the Task of Modeling the Lifecycle of the Learning Process Based on a Modified CA

Cell states. Each cell of the field is an individual, whose consciousness can perceive a certain opinion about some issue. An individual's opinion can vary within a non-discrete scale from -1 to 1, where the left boundary (1) is one of attitude to the question, the middle (0) is a neutral position to the question, and the right boundary (1) is opposite to the left boundary of opinion. In what follows, we will assume that the ideal-right (from the point of view of the researchers who use the model) opinion is the opinion that corresponds to the value 1.

An individual's opinion changes under the influence of communication with the individuals around him. The dynamics of the society (cell field) as a whole are very much influenced by the initial values of the field cells' thoughts and the radius of the interaction of each of the cells. At that, without external interference, the average opinion of the society tends to a neutral value in the range from -0.1 to 0.1.

The radius of interaction in the information field is always different and depends on the type of communication channel: personal, public events, mass media, and the like. Therefore, the definition of the interaction radius for each cell of the field during the real simulation is an important element to ensure the success of the simulation. In the learning process model, the interaction radius of each cell cannot be the same for everyone. And this makes the complicated rule of determining the cell vicinity a necessary part of the learning process model.

5.2 Testing of the Model on the Example of the Learning Process with the Science-Centered Approach

The considered statement of the problem was tested on the model of the learning process in science-centered approach [15] in education, which focuses on teaching students through encouraging them to scientific activity and creativity during their studies.

For the **positive opinion** in the model is considered the opinion that "it is useful for the student to be engaged in science for learning and self-development as a specialist". The opposite **negative opinion** is that "it makes no sense for the student to engage in science because it is not needed in the future".

It is worth noting that the results of modeling depend not only on the initial state of the cells in the society but also on the controlling influences of the external environment.

5.3 Problem Statement

The cells are students of the same stream. For the approbation, we will assume that one thread consists of 100 students so that the field has a size of 10 by 10 cells.

For simplicity, **the field** is represented in two-dimensional form, and to increase contacts, it is convolved into a torus. The initial values of cells are set randomly within the specified numbers. In the case of modeling a real flow of some specialty, to determine the states of the cells of the field it is sufficient to conduct a sociological survey on the topic of the student's attitude to scientific activity and activity. This will also make it possible to establish the value of student interaction radii. The same initial cell field (Fig. 8a) and interaction radius field (Fig. 8b) are used for the model.
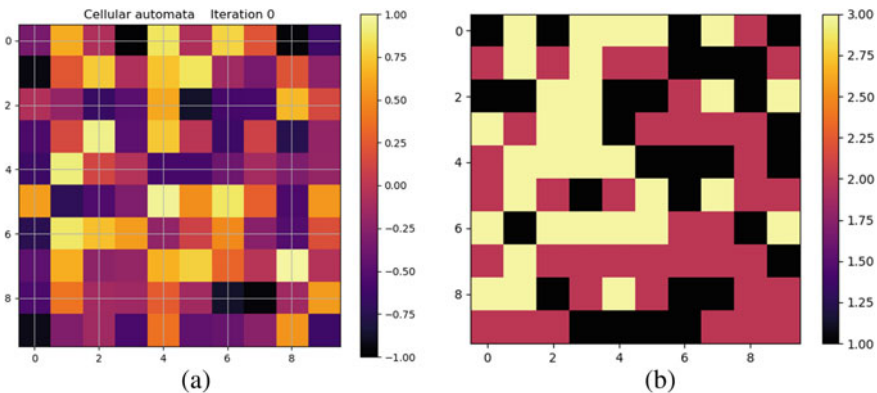

Fig. 8 Initial fields of the automata. **a** The initial cell field, **b** the field of interaction radiuses

When conditions change, the cellular field will change as well.

Let us consider 6 variants of the development of events:

- With a randomly distributed (by opinion and ability) stream of students, whose thoughts are not influenced from outside.
- With the flow divided into groups and without influence from the outside.
- With group-divided flow and with influence on a group of 4 flow students.
- With group-divided flow and with influence on a group of 16 students in the flow.
- With group-divided flow and influence on a group of 36 flow students.
- With grouped flow and impact on a group of 36 students in the flow.

5.4 Variant 1: Random Distribution with No External Influence

To analyze changes in total flux opinion, we will use a graph of flux opinion value dependence on automata iteration (time). In the case when the cell field is set randomly without division into subgroups and without influence, the flow opinion predictably comes (Fig. 9) to the value close to zero (0.011), which corresponds to the neutral attitude of students to participation in scientific activity. Of course, for a while, the general opinion fluctuates because each student initially has his or her own independent opinion, but eventually, everyone's opinion will come to a neutral general opinion.
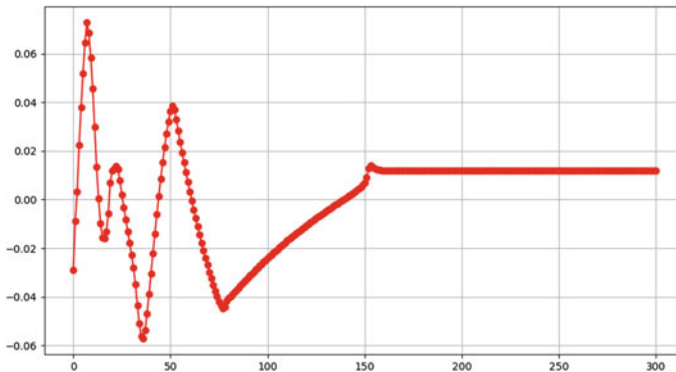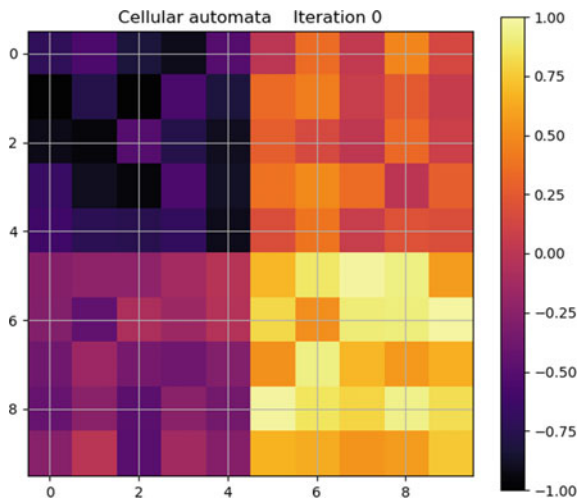

Fig. 9 The graph of the time dependence of the total opinion of the students of the stream

5.5 Variant 2: Division into Groups in the Absence of External Influence

Since usually groups of streams are formed at the beginning of the first based on the rating score or scores of entrance exams, it is appropriate to use the model with the division into subgroups. For this purpose let's form a new flow cell field, which will consist of 4 groups, in each of which students will have attitude to scientific activity in the range from -1 to -0.5 (the upper-left edge of the field—students are nihilists, that they are afraid to do science and do not see the point in it), from -0.5 to 0 (left lower edge—students who are afraid to do science and are inactive in the learning process), 0 to 0.5 (right upper edge—students who are not afraid to do science but lack general motivation), and 0.5 to 1 (right lower edge—science-oriented students who can do science with a little outside intervention). The cellular field created (Fig. 10) will also be used in subsequent models.

In the case of splitting the flow into groups and not interfering with the process of opinion formation, the average opinion on the flow (Fig. 11) comes to a neutral state 4.4 times faster than in variant 1 (Fig. 9). This is due to the fact that in a divided stream, like-minded students are gathered in one group, which leads to less discussion activity between students, as was the case in the previous model. The final average opinion is still almost neutral -0.16. Therefore, in order to form a correct attitude of students to scientific activity, professors, departmental and faculty leaders need to be actively involved in the formation of students' opinions.

Fig. 10 Cellular field of the stream with separation into groups

Fig. 11 The graph of the time dependence of the total opinion of the students of the stream

5.6 Variant 3: Division into Groups and Presence of Influence on 4 Students

In variant 1 and variant 2 the value of average opinion on the flow always tends to zero because the natural state for most students is neutral. But this can be influenced externally by conducting student agitation. To influence students' opinions, it was decided to choose four students with whom active agitation (encouragement to write papers for scientific competitions, to write theses, and conduct research) would be conducted. Students from the best group (with an opinion value of 0.5 to 1) were selected for greater effectiveness. As a result of the intervention, the average ratio of flow (Fig. 12) to research activity increased by 0.45 relatives to the value without the intervention, which is a fairly significant difference.

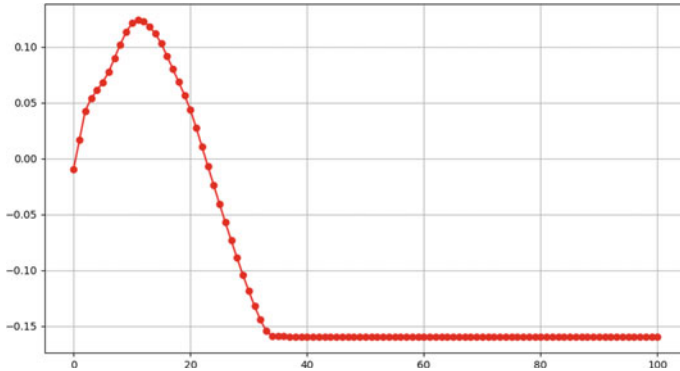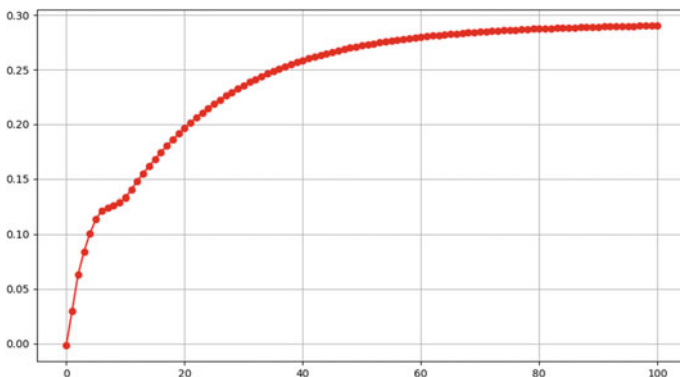**Fig. 12** The graph of the time dependence of the total opinion of the students of the stream
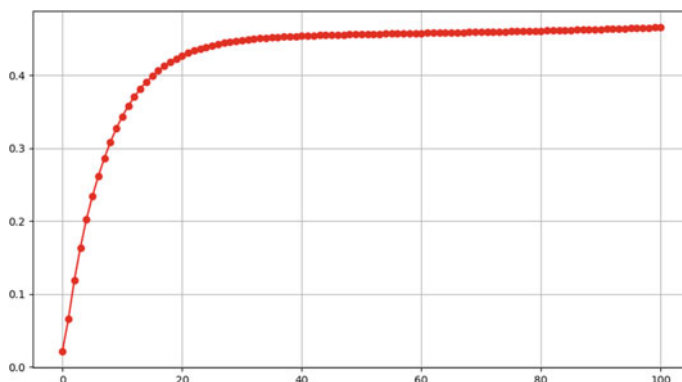
Fig. 13 The graph of the time dependence of the total opinion of the students of the stream

5.7 Variant 4: Division into Groups and Presence of Influence on 16 Students

In this case, the number of students with whom agitation is conducted has increased to 16. As a result of increasing agitation to 16, the average ratio of flow (Fig. 13) to academic activity increased by 0.63 (i.e., is 0.47) relative to the value without the intervention. At the same time saturation, when the rate of change of the flow opinion decreases, comes earlier by 40 iterations of the automata, which is also an important indicator because in this case the resources used on agitation can be saved.

5.8 Variant 5: Division into Groups and Presence of Influence on 36 Students

In this case, the number of students with whom agitation is conducted increased to 36, due to which the average ratio of the stream (Fig. 14) to academic activity increased by 0.79 (i.e., is 0.63) relative to the value without the intervention. The value achieved is within the range that was set when the best group was formed (0.5 to 1). This makes student learning and collaboration with faculty easier and more effective.

5.9 Variant 6: Division into Groups and the Impact on 36 Students from Different Groups

It is also reasonable to simulate the case when the students with whom the agitation is to be conducted are selected evenly in all groups of the flow. The number of students
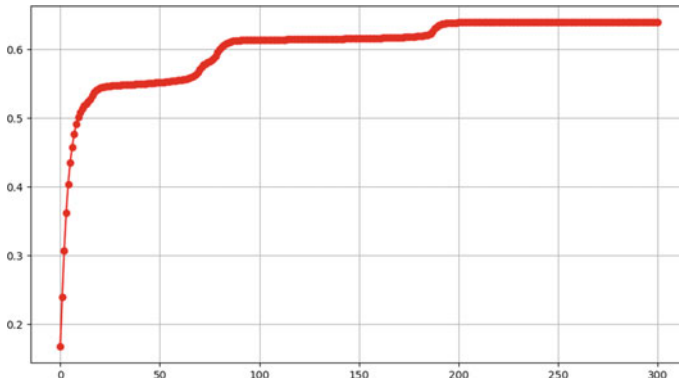
Fig. 14 The graph of the time dependence of the total opinion of the students of the stream

with whom agitation is carried out let us set 36. As a result, the average flow ratio (Fig. 15) of the academic activity increased by 0.77 (i.e., it is 0.61) relative to the value without the intervention. Campaigning took place quite effectively, but at the same time not as much as in variant 5. At the same time in the case of variant 5 campaigning took place systematically without sharp jumps in values.

Based on the created variants of development of the model of the educational process at the science-centered approach the comparative table with an initial number of students prepared for scientific activity, a final number of students prepared for scientific activity, and the difference between these values was created (Table 1).

It is worth noting that existing approaches to predicting the spread of ideas in the learning environment do not take into account the involvement of students with the help of other students, but only through teachers. This leads to an error ranging from 20 to 44%, which is a significant deviation. The developed model gives a gain in estimating the dynamics of change in student preferences at the level of
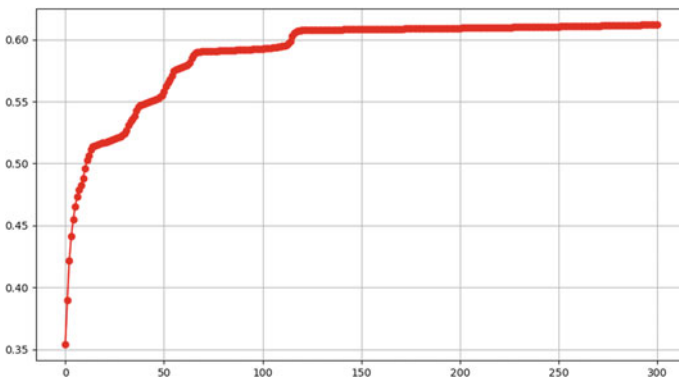

Fig. 15 The graph of the time dependence of the total opinion of the students of the stream

Table 1 Comparison of options for changing students' opinions

Variant	Initial number	Final number	Difference between values
1	25	0	−25
2	25	0	−25
3	25	5	−20
4	25	66	41
5	36	77	41
6	36	80	44

41–44%. Based on the results of the study (Tab. 1), the most effective is the 6th variant of development, which makes 80% of the students of the stream active in scientific activities. Depending on the results of the survey of students and further modeling of various options for the development of events can be selected as the most optimal. Non-interference of the faculty administration, department, and teachers in the process of formation of the attitude to scientific activity leads to almost complete absence of any activity of students, which in turn worsens the quality of student learning and formation of highly qualified specialists.

6 Conclusions

- The paper developed methods for predicting changes in public opinion through the development of a model of social behavior on the acceptance of educational material in a science-centered approach to education.
- The work has for the first time created a model of the educational process in the science-centered approach to education based on the model of CA, taking into account the additional parameters of the influence of cell state definition parameters, the basic and modified rules for determining the vicinity of cells, the rule of dependence of the susceptibility coefficient on the state of the cell and its environment.
- The developed model allows us to predict the dynamics of individual thought propagation in the learning process. The improved CA can be used to model both the objects of society and the process of dissemination of ideas contributing to the implementation of a science-centered approach.
- The paper developed software tools to analyze the dynamics of CA using the model of the learning process in the science-centered approach. The developed model gives a gain in assessing the dynamics of changes in student preferences at the level of 41% to 44%, in contrast to existing approaches, which give an inaccuracy of 20–44% when predicting.
- CA model is implemented in two algorithmic languages Python and MatLab, which allows you to choose the simulation environment when changing the features of the problem statement.

- The direction of further research is to expand the rules of CA, in particular, the rule to take into account unforeseen influences from outside the simulated society.

References

1. Bartosh OP, Hvozdetska BG, Varha NI, Zoska YV, Nikon NA (2020) Professional socialization of youth in Central and Eastern Europe. Perspektivy nauki i obrazovania—Persp Sci Educ 48(6):103–116. https://doi.org/10.32744/pse.2020.6.9

2. Chasmer LE, Ryerson RA, Coburn CA (2021) Educating the next generation of remote sensing specialists: skills and industry needs in a changing world. Can J Remote Sens. https://doi.org/10.1080/07038992.2021.1925531

3. Conger D, Kennedy AI, Long MC, McGhee R (2021) The effect of advanced placement science on students' skills, confidence, and stress. J Human Resour 56(1):93–124. https://doi.org/10.3368/jhr.56.1.0118-9298R3

4. Zhukov AE (2017) Cellular automata in cryptography Part 1. Cybersecurity Issues 3(21):70–76. https://doi.org/10.21681/2311-3456-2017-3-70-76

5. Conway's game of life [electronic resource] Website LifeWiki. www.conwaylife.com/wiki/Conway's_Game_of_Life

6. Shevchenko V, Shevchenko A, Fedorenko R, Shmorhun Y, Hrebennikov A (2019) Designing of functionally stable information systems optimal for a minimum of losses. In: CADSM 2019, 15th international conference on the experience of designing and application of CAD systems (CADSM), 26 Feb–2 March, 2019, Polyana-Svalyava (Zakarpattya), UKRAINE, IEEE Ukraine section, IEEE Ukraine Section (West), MTT/ED/AP/EP/SSC societies joint chapter part number: CFP19508-USB, pp 36–40. ISBN: 978-1-7281-0053-1

7. Petrov P, Dimitrov G, Ivanov S (2018) A comparative study on web security technologies used in Irish and Finnish banks. In: 18 International multidisciplinary scientific geoconference SGEM 2018: conference proceedings, 2–8 July 2018, Albena, Bulgaria, vol 18; Inf Geoinformatics Remote Sens 18(2.1):3–10

8. Schneckenreither G, Popper N, Breitenecker F (2015) Methods for cellular automata and evolution systems in modelling and simulation. In: 8th Vienna international conference on mathematical modelling MATHMOD 2015, 18–20 Feb, 2015, Vienna, Austria, pp 141–146. https://doi.org/10.1016/j.ifacol.2015.05.151

9. Shevchenko V, Berestov D, Sinitcyn I (2020) Built-in processor for sharing passwords through the open information space. In: 2020 16th International conference perspective technologies and methods in MEMS design (MEMSTECH), proceeding, Lviv, 22–26 April 2020, pp 40–44. https://ieeexplore.ieee.org/document/9109523

10. Freire J, DaCamara C (2019) Using cellular automata to simulate wildfire propagation and to assist in fire management. In: Natural hazards and earth system sciences (NHESS), 22 Jan, 2019, Lisbon, Portugal, pp 169–179

11. Salcido A (2011) Equilibrium properties of the cellular automata models for traffic flow in a single lane. In: Salcido A (ed) Cellular automata—simplicity behind complexity, pp 159–192

12. Piroli L, Cirac I (2020) Quantum cellular automata, tensor networks, and area laws. Phys Rev Lett:5. https://doi.org/10.1103/PhysRevLett.125.190402

13. Zhang Y, Gong C, Li D (2021) A prognostic dynamic model applicable to infectious diseases providing easily visualized guides: a case study of COVID-19 in the UK. Sci Rep 11:8412. https://doi.org/10.1038/s41598-021-87882-9

14. Xiao M, Zhan Q, Li Y (2021) Research on combating epidemics based on differential equations and cellular automata. In: 2021 International conference on advances in optics and computational sciences (ICAOCS) 2021, 21–23 Jan, 2021, Ottawa, Canada, p 7. https://doi.org/10.1088/1742-6596/1865/4/042143

15. Kiray SA, Kaptan F (2012) The effectiveness of an integrated science and mathematics programme: science-centred mathematics-assisted integration. Energy Educ Sci Technol Part B: Social Educ Stud 4(2):943–956

Research of Topological Properties of Network Reflections Obtained Using Different Algorithms for Scanning Initial Networks

Dmytro Lande and **Oleh Dmytrenko**

Abstract Many modern types of researches of networks use mechanisms for their monitoring, after which conclusions about the topology of such networks are drawn. This paper shows that this approach is wrong. The reflections of the initial networks obtained as a result of monitoring and partially reflect the properties of these initial networks is often significantly different. The properties of these network reflections significantly depend on the algorithms used for scanning the initial network. To demonstrate this statement, the paper researches the properties of network reflections obtained using three scanning algorithms, which implement the following principles: (1) the transition to the node according to the PageRank algorithm; (2) the transition to the node that has the highest PageRank; (3) the transition to the node that has the highest degree. The networks based on the Erdös-Renyi and Barabási-Albert models are considered as basic. The paper shows that the values of the characteristics of network reflections, which are close to the topology and characteristics of the initial networks, are achieved only during approaching the scanning steps to the number of nodes in these initial networks. The obtained results are important in the methodical plan and can be considered as a statement of a problem of finding the algorithm of scanning of a complex network using of which gives most adequately parameters of the initial network (first of all, the degree distribution).

Keywords Monitoring · Network scanning · Network topology · Barabási-Albert model · Erdös-Renyi model · PageRank algorithm · Network density · Degree · Degree distribution · Power law

D. Lande · O. Dmytrenko (✉)
Institute for Information Recording, National Academy of Sciences of Ukraine, Kyiv, Ukraine

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

1   Introduction

Recently, more and more scientific papers have appeared on the study of huge dimension information networks. Not only the content of such networks but also the structure (nodes and connections) can be methodically attributed to the category of Big Data. In practice, it turns out that it is impossible to cover some network in full by common methods. In such cases, special algorithms are used. Due to the past development of so-called peer-to-peer networks, algorithms such as the Breadth-first Search (BFS) method [1], the Random Breadth-first Search (RBFS) method, the Intelligent Search Mechanism (ISM) method, the Depth-first Search (DFS) [2, 3], the Dijkstra's algorithm [4], the Floyd-Worshell algorithm [5, 6], the Bellman-Ford algorithm [7, 8], the finding connection points and bridges in a graph [9], etc. [10, 11].

All these algorithms provide the ability for traversing networks and searching tree or graph data structures (primarily, it was assumed that they will be used to find or integrate targeted content). Using these algorithms, some researchers draw conclusions about the topology of networks, in particular co-authorship networks [12]. The use of probabilistic models of the networks [13–16] makes it possible to analyze the structure of dependencies between the corresponding nodes in this network, to find the probabilities of the existence of certain connections and to obtain estimates of various network characteristics. But it was found that mentioned above models do not always display the real network topology. Therefore, this paper considers and researches how the topology of the network reflection, i.e. the network built using a limited number of scanning steps, depends on the initial network and the scanning algorithm.

2   Random Network Model

For modelling, as an example, three artefact networks namely, Barabási-Albert [17], Erdös-Renyi [18] and Watts-Strogatz [19] networks are investigated. These random networks models can be considered as prototypes of many real networks.

2.1   Barabási-Albert Network: Model of Preferential Connection

Most real artefact networks have a power-law distribution. It turned out that this distribution is due to an effect called the cumulative advantage or preferential attachment. The power-law networks include Barabási-Albert networks (Fig. 1) [17].

To build these networks, a special procedure is used, which consists of the fact that new nodes are gradually added to the initially small number of nodes, links from which are more likely to connect to those nodes that have more links. That is, in
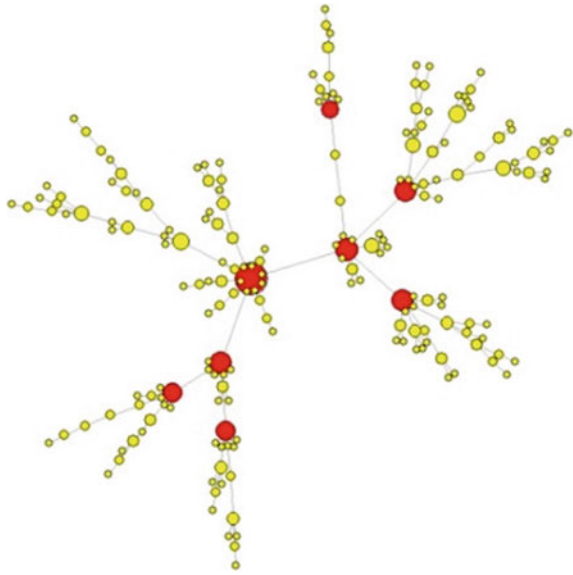
the process of network growth, new nodes are more likely to form connections with those nodes that are already characterized by a large number of connections.

It has been proven that it is the "rich getting richer" (phenomenon that leads to the emergence of power laws in networks). Obviously, when a new node joins the network, only one link is used, i.e. the number of edges in the network is comparable to the number of nodes, and the network is quasi-hierarchical (the hierarchy can be violated only in the initial composition of nodes).

The degree distribution of the Barabási-Albert model is scale-free, i.e. it obeys the power law

$$P(k) \sim k^{-3} \tag{1}$$

2.2 Erdös-Renyi Network

The object that has been intensively reseached in graph theory and is directly related to complex networks is the classical random graph, or Erdös-Renyi random graph. It was proposed and researched in the late 1950s by Paul Erdős and Alfred Rényi [20].

The Erdös-Renyi network can be constructed by randomly distributing *m* connections between *n* nodes. This model is equivalent to a model in which the value of the number of edges *m* is replaced by the corresponding probability *p* of a new edge appearing in the graph (Fig. 2).

Fig. 2 Erdös-Renyi network
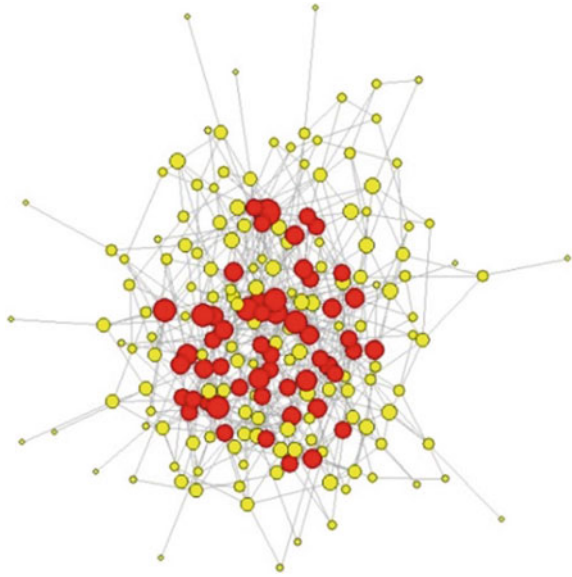
It is sometimes called the random graph model or sometimes just the Poisson random graph model because of the Poisson degree distribution for $n \to \infty$ and np = const

$$P(k) \sim \frac{(np)^k e^{-np}}{k!} \qquad (2)$$

3 The Proposed Network Scanning Algorithms

To build reflections of networks and further research the dependence of the network characteristics of the obtained reflections of networks on the network characteristics of the initial networks besides the different number of scanning steps, three proposed scanning algorithms, which implement the following principles are also used: (1) the transition to the node according to the PageRank algorithm; (2) the transition to the node with the largest value of PageRank; (3) the transition to the node with the largest value of degree.

3.1 Building the Network Reflection Using the PageRank Algorithm

The PageRank algorithm [21, 22] was proposed in 1996 at Stanford University by Larry Page and Sergey Brin as part of a research project on a new type of information retrieval system. The system uses the PageRank algorithm to measure the importance of website pages and rank them in their search engine results. Taking into account the structure and text of hyperlinks, the PageRank algorithm simulated a random walk of an Internet user starting from a random page. The more random visits to a page, the higher its ranking.

The PageRank for a page A is calculated according to the following rules. Let T_1, …, T_n be the pages that link the page A. The algorithm also uses a damping factor d, the values of which are between 0 and 1, and usually equals 0.85. The function $C(T)$ is equal to the number of links outbound from page T. Then the PageRank of page A, $PR(A)$, is equal to:

$$PR(A) = (1 - d) \sum_{i=1}^{n} \frac{PR(T_i)}{C(T_i)} \tag{3}$$

In other words, the PageRank are random values, the sum of which for all pages will be equal to 1.

To calculate the PageRank, the Internet space is represented as an oriented graph, the vertices of which correspond to the pages, and the edges correspond to the hyperlink between them. Let n pages be included in the search index. Then a matrix of transitions M of size $n \times n$ is created to model a random walking. The element of this matrix m_{ij}, which is in row i and column j has a value of $1/k$ if the page with number j has k outgoing links, among which there is one that link page with number i. If there is no such outgoing link, the element m_{ij} equals 0.

The probability distribution of finding a random traveler can be described by a column vector whose row j will be equal to the probability of being on the page j [23]. This vector corresponds to the simplest and idealized variant of the PageRank.

To build a reflection of the initial network according to the described above PageRank algorithm, you need to make a traversing along the path, which is determined by a random walk across the initial network. The number of network scanning steps is limited. During the scan, all nodes and connections are fixed. The set of these nodes and connections and determines the reflection of the initial network.

3.2 Building the Network Reflection Using the Algorithm of Maximum PageRank

At the initial stage, the PageRank characteristic is assigned for each node of the initial network. Next, the network is scanned as follows: from the current node via the outgoing link the transition to the node that has the highest PageRank is made. The network scanning process is continued during a limited number of steps. If during traversing the network there is a return to the node from which the transition has already been made, then the next transition via the not yet passed outgoing link that leads to the node with the highest PageRank is made. If there is no such outgoing link, the scanning process is continued from any random node of the initial network.

3.3 Building the Network Reflection Using the Algorithm of Maximum Degree

Node degree is a characteristic of the total number of both incoming and outgoing links [11, 24].

Building the network reflection using the algorithm of maximum degree is made on the same principle as in the previous case, however, the defining characteristic of the node is not PageRank, but its other characteristic the degree of the node is used.

Since all the proposed algorithms to some extent have a probabilistic nature, then for a particular initial network many different reflections of the network can be obtained. Therefore, the resulting values of the characteristics of the network reflections are averaged over many implementations.

4 Results

The network obtained as a result of using the Barabási-Albert model was used for the research. The graph corresponding to the generated network consists of 200 vertices randomly connected by 398 edges (each new node attached to the network was connected to the existing one only using the 1 edge). The density of the obtained network is 0.01, and the average degree is 1.99.

The network scanning was performed using the three algorithms proposed above. The maximum number of scanning steps was 50, 100, 150, and 200.

Figures 3, 4, and 5 presents different reflections of the initial Barabási-Albert network for respectively different network scanning algorithms.

Figure 3 shows the reflection (highlighted in bold) of the initial Barabási-Albert network, which was obtained using 50 scanning steps, which in turn were made on the principle of random walking, i.e. on the principle that uses the common PageRank algorithm. The obtained reflection of the initial network consists of 18 nodes and 33
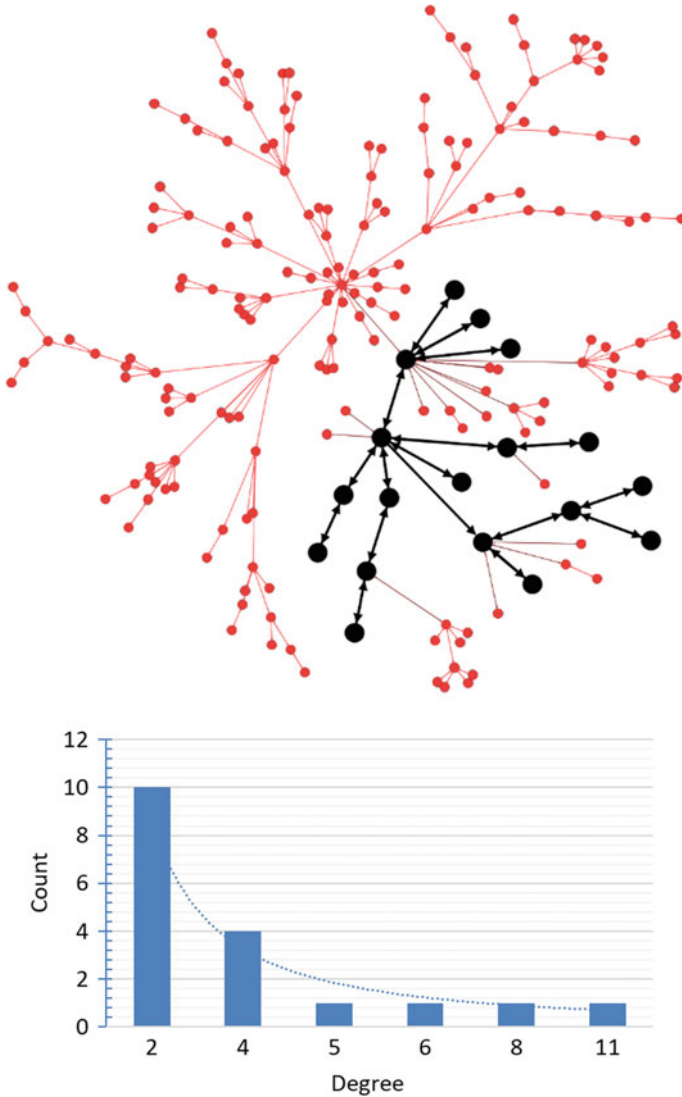
Fig. 3 The reflection of the initial network built using the PageRank algorithm, and the degree distribution of the obtained network reflection

edges (Fig. 3). The average node degree is 1.833, and the network density is 0.108. The degree distribution of the nodes of the obtained reflection, as well as the degree distribution of the initial Barabási-Albert network, at least asymptotically follows a power law.

The research was also done using the algorithm that implements the principle of transition to the node that has the highest degree. After 100 scanning steps of the

Fig. 4 The reflection of the initial network built using the algorithm that implements the principle of transition to the node that has the highest degree, and the degree distribution of the obtained network reflection

initial Barabási-Albert network using the above mentioned algorithm of maximum degree, the reflection of the network, which consist of 37 nodes and 69 edges, was obtained (Fig. 4). The average node degree is 1.865, and the network density is 0.052.

After 200 steps scanning of the initial Barabási-Albert network using the algorithm of transition to the node that has the highest PageRank, the reflection of the network that consist of 65 nodes and 126 edges was obtained (Fig. 5). The average node degree is 1.938, and the network density is 0.03.

Analyzing the results presented in Figs. 3, 4, and 5, it can be seen that networks that are the reflections of the initial Barabási-Albert network and built using a limited

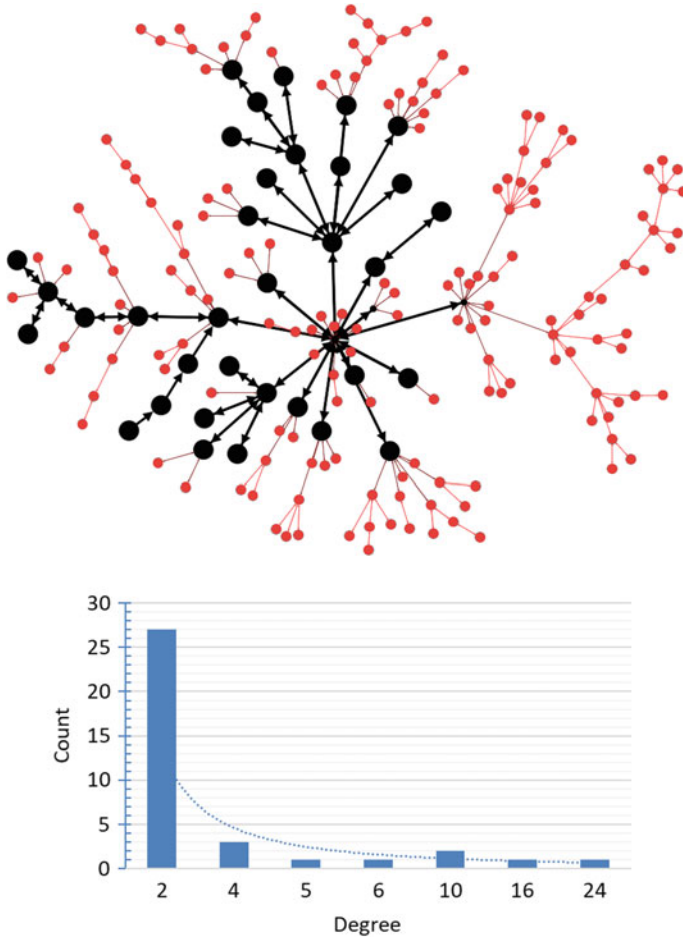**Fig. 5** The reflection of the initial network built using the algorithm that implements the principle of transition to the node that has the highest PageRank, and the degree distribution of the obtained network reflection

number of scanning steps, have the same power-law degree distribution as the initial network.

In general, it can be observed that during approaching the number of scanning steps to the number of nodes in the initial network for each of the considered algorithms for scanning this network, the degree distribution of the obtained network reflections approximates a power-law degree distribution, by which the Barabási-Albert network is characterized a priori.

Also, to research the dependence of the network characteristics of the reflections on the number of scanning steps as an initial network was used a network built using the Erdös-Renyi model. The graph corresponding to the generated network consists of 200 vertices randomly connected by 428 edges (the probability for edge creation

is 0.01). The average node degree the generated network is 2.14, and the network density is 0.011.

To build the reflection of the initial network (Fig. 6), 200 scanning steps using an algorithm that implements the principle of transition to the node that has the highest degree is made. As result, the reflection of the initial Erdös-Renyi network, which
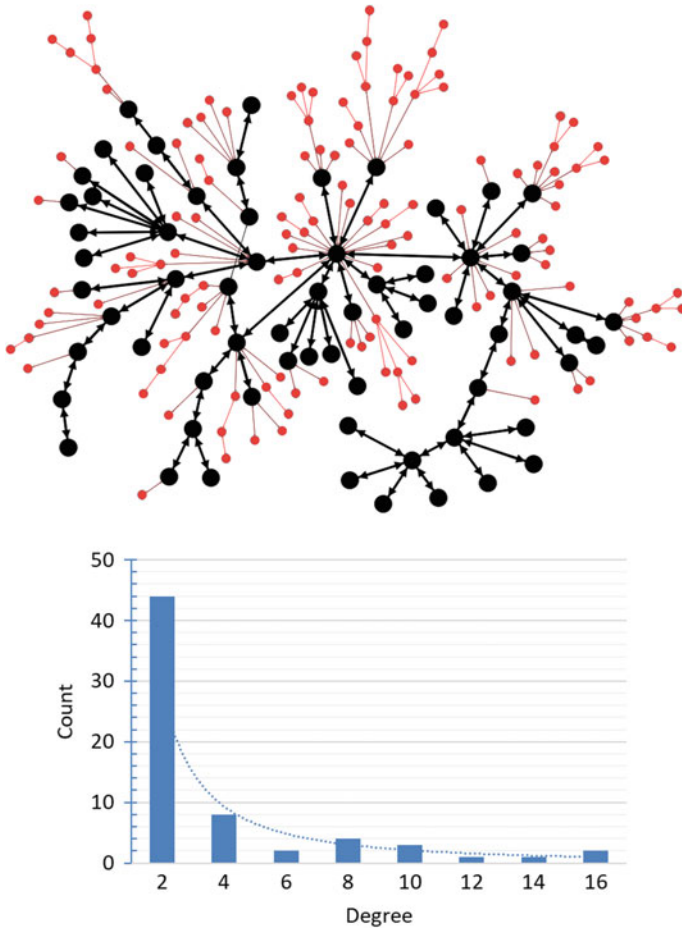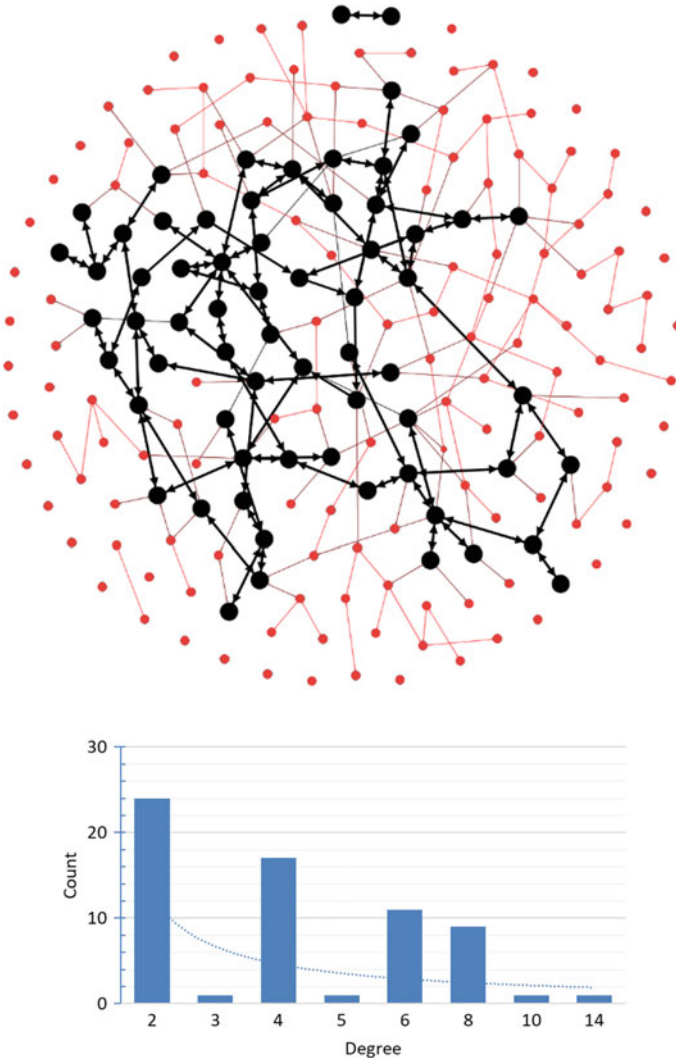

Fig. 6 The reflection of the initial network built using the algorithm that implements the principle of transition to the node that has the highest degree, and the degree distribution for the obtained network reflection
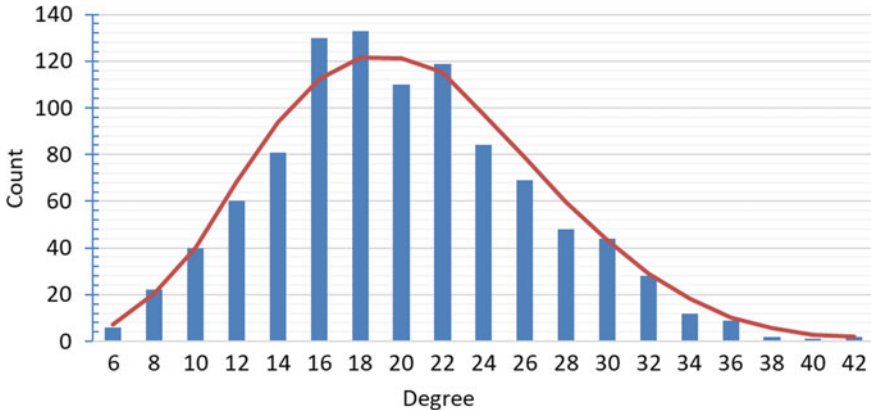
Fig. 7 The Poisson degree distribution of the initial Erdös-Renyi network

consists of 65 nodes and 143 edges, was obtained. The average node degree is 2.2, and the network density is 0.034.

Figure 6 also presents the degree distribution of the nodes of the obtained network reflection. As can be seen, using a limited number of scanning steps, the reflection of the initial Erdös-Renyi network, which has distinct from a Poisson degree distribution was obtained. While the initial Erdös-Renyi network has a Poisson degree distribution.

After more detailed research on the example of networks built using the Erdös-Renyi network model, and applying different scanning algorithms, the reflection of the original networks, which have close to a power-law degree distributions were obtained. For example, the Erdös-Renyi network, which has 1000 nodes and 10,030 connections (the probability for edge creation is 0.01) and which has a Poisson degree distribution (Fig. 7) was researched. 500 scanning steps using the algorithms that implement the principle of transition to the node that has the highest PageRank and degree, respectively. As a result, the close to a power-law degree distribution of the corresponding network reflections were obtained (Figs. 8 and 9).

The experimental data presented in Table 1 illustrate the dependence of the characteristics (the average node degree and the network density) of the networks reflections on the algorithms used for scanning the initial networks (Barabási-Albert and Erdös-Renyi networks) and on the number of scanning steps.

In general, for a different limited number of scanning steps, depending on the chosen scanning algorithm, the different characteristics of the reflections of the Barabási-Albert and Erdös-Renyi networks were obtained.

It was found that increasing scanning steps leads to an approximation of the average degree and density of the obtained reflections of the Barabási-Albert and Erdös-Renyi networks to the average degree and density of these initial networks. That is, when approaching the scanning steps to the number of nodes in the initial networks for each of the proposed scanning algorithms, it is possible to achieve such
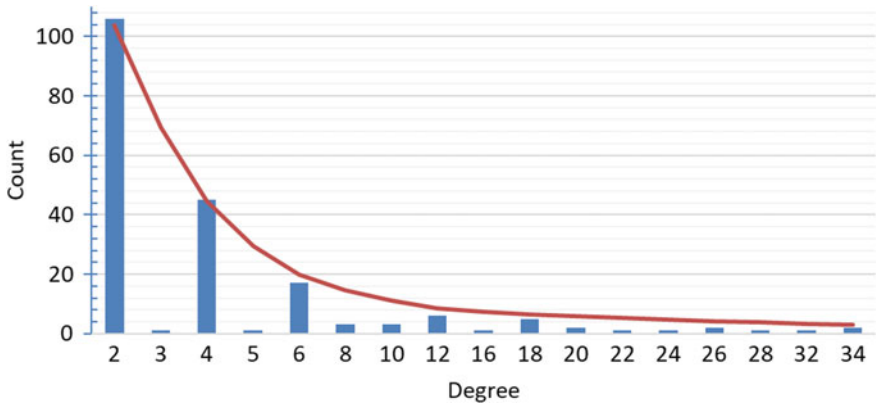
Fig. 8 The degree distribution of the Erdös-Renyi network reflection that built using the algorithm of maximum PageRank
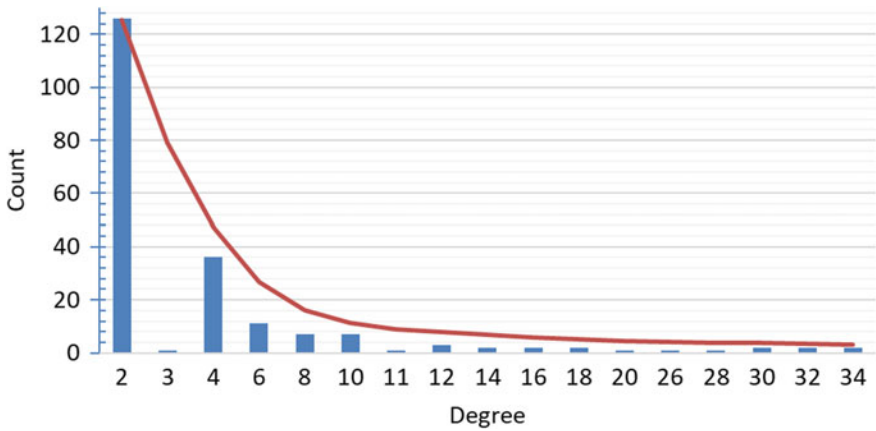
Fig. 9 The degree distribution of the Erdös-Renyi network reflection that built using the algorithm of maximum degree

characteristics of the networks reflections that are close to the characteristics of the initial networks. Also based on the data obtained by computational experiments, it is shown that almost all reflections of the initial networks that built using the model Barabási-Albert and Erdös-Renyi, and also using a limited number of scanning steps and the different scanning algorithms have a power-law degree distribution.

Table 1 Dependence of the network characteristics of the reflections on the number of scanning steps and the chosen scanning algorithm

Initial network	Scanning algorithm	Number of scanning steps	Number of nodes	Number of edges	Average degree	Density
Barabási-Albert network (Number of nodes is 200, Number of edges to attach from a new node to existing nodes is 1)	PageRank	50	18	33	1.833	0.108
		100	33	61	1.848	0.058
		150	43	78	1.814	0.043
		200	60	115	1.917	0.032
	Max PageRank	50	25	44	1.76	0.073
		100	36	70	1.944	0.056
		150	50	95	1.9	0.039
		200	65	126	1.938	0.03
	Max degree	50	25	44	1.76	0.073
		100	37	69	1.865	0.052
		150	53	102	1.925	0.037
		200	63	122	1.937	0.031
Erdös-Renyi network (Number of nodes is 200, Probability for edge creation is 0.01)	PageRank	50	20	34	1.7	0.089
		100	42	70	1.667	0.041
		150	58	103	1.776	0.031
		200	80	132	1.65	0.021
	Max PageRank	50	23	41	1.783	0.081
		100	38	79	2.079	0.056
		150	56	108	1.929	0.035
		200	73	147	2.014	0.028
	Max degree	50	21	38	1.81	0.09
		100	48	95	1.979	0.042
		150	57	118	2.07	0.037
		200	65	143	2.2	0.034

5　Conclusion

The paper shows that the characteristics of network reflections, which are close to the topology and characteristics of the initial networks, are achieved only during approaching the scanning steps to the number of nodes in these initial networks.

For the first time, the experimentally obtained characteristics of the reflections of networks depending on the characteristics of the initial networks, the scanning algorithms and the number of scanning steps are presented.

Also, based on the data obtained by computational experiments, it was shown that almost all reflections of the initial networks that built using the model Barabási-Albert and Erdös-Renyi, and also using a limited number of scanning steps and the different scanning algorithms have a power-law degree distribution.

The obtained results are important in the methodical plan and can be considered as a statement of a problem of finding the algorithm of scanning of a complex network using of which gives most adequately parameters of the initial network (first of all, the degree distribution).

References

1. Cormen TH, Leiserson CE, Rivest RL, Stein C (2009) Introduction to algorithms. MIT Press
2. Tarjan R (1972) Depth-first search and linear graph algorithms. SIAM J Comput 1(2):146–160
3. Korf RE (1985) Depth-first iterative-deepening: an optimal admissible tree search. Artif Intell 27(1):97–109
4. Dijkstra EW (1959) A note on two problems in connexion with graphs. Numer Math 1(1):269–271
5. Floyd RW (1962) Algorithm 97: shortest path. Commun ACM 5(6):345
6. Warshall S (1962) A theorem on Boolean matrices. JACM 9(1):11–12
7. Bellman R (1958) On a routing problem. Q Appl Math 16(1):87–90
8. Ford Jr LR, Fulkerson DR (2015) Flows in networks. Princeton University Press
9. Tarjan RE, RE T (1974) A note on finding the bridges of a graph. Inf Process Lett 2(6):160–161
10. Snarskii AA, Lande DV (2015) Modeling of complex networks: tutorial. In: K.: Engineering. ISBN 978-966-2344-44-8
11. Lande D, Snarskii A, Dmytrenko O, Subach I (2020) Relaxation time in complex network. In: Proceedings of the 15th international conference on availability, reliability and security (ARES 2020), Ireland, Europe
12. Newman ME (2006) Finding community structure in networks using the eigenvectors of matrices. Physical Rev E 74(3):036104
13. Jensen D, Neville J (2002) Data mining in social networks. In: Proceedings of the national academy of sciences symposium on dynamic social network analysis, pp 289–302
14. Bonchi F, Castillo C, Gionis A, Jaimes A (2011) Social network analysis and mining for business applications. ACM Trans Intell Syst Technol (TIST) 2(3):1–37
15. Leskovec J, Kleinberg J, Faloutsos C (2005) Graphs over time: densification laws, shrinking diameters and possible explanations. In: Proceedings of the eleventh ACM SIGKDD international conference on knowledge discovery in data mining, pp 177–187
16. Leskovec J, Backstrom L, Kumar R, Tomkins A (2008) Microscopic evolution of social networks. In: Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining, pp 462–470
17. Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–512
18. Erdős P, Rényi A (1959) On random graphs I. Publicationes Mathematicae 6:290–297
19. Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393(6684):440–442
20. Erdős P, Rényi A (1960) On the evolution of random graphs. Publ Math Inst Hung Acad Sci 5(1):17–60
21. Page L, Brin S, Motwani R, Winograd T (1999) The PageRank citation ranking: bringing order to the web. Stanford InfoLab
22. Langville AN, Meyer CD.: Google's PageRank and beyond: the science of search engine rankings. Princeton University Press

23. Rajaraman A, Ullman JD (2011) Mining of massive datasets. Cambridge University Press
24. Albert R, Barabási AL (2002) Statistical mechanics of complex networks. Rev Mod Phys 74(1):47

video analytics. In general, the structure of motion is a very important component that can be used for various kinds of forecasting, strategic planning, decision-making on the coach's substitutions, etc. [2].

In order to solve the problem of representation and, as a consequence, understanding of the structure of motion, we will solve the problem of obtaining a mask of a moving object.

Today, there are many ways to determine the mask of a moving object. One of the most common approaches to this problem is background subtraction. Background subtraction is a technique in the fields of image processing and computer vision where the image foreground is extracted for further processing (object recognition, etc.) [3]. The essence of the method is to select the background. That is, those pixels that do not change from frame to frame are the background, and those that change are, accordingly, a moving object.

Let us consider this approach in more detail. Let there be a certain set of frames. The value of a separate pixel on a separate frame can be described by the following formula:

$$V_i(x, y, t),$$

where V_i — value of intensity of the pixel; x — coordinate of the pixel by the abscissa axis; y — coordinate of the pixel by the ordinate axis; t — time.

Then the background subtraction can be described as follows:

$$|V_i(x, y, t) - V_i(x, y, t + 1)| > Th,$$

where Th — threshold value.

That is, if the difference between the intensity of the pixel value of a particular frame and the next one exceeds a certain threshold value, then this pixel belongs to a moving object.

This approach has both advantages and disadvantages. The advantages include ease of implementation and the speed of the algorithm. But the disadvantages include the instability of this approach to changes in illumination, the presence of noise, etc.

There are many modifications of this method, which consist in a more complex calculation of the background. For example, the calculation of the background relative to several frames, taking into account the average value of the background pixels [4]:

$$V_{i+1} = \alpha I_i + (1 - \alpha) V_i,$$

video analytics. Working with video files in terms of data analytics is a step-by-step review of each frame. The result is created based on the data of current and previous frames. So, the main task is to process each image of the video. It is possible to divide sports into two groups. The video files of the first group have events with complex spatiotemporal relationships. The second group has simple actions based on sequences of short movements.

In paper [2], was proposed an approach to detecting moving objects in a video sequence by removing the background. The authors suggest first to create a background model before removing it from each sequence image and then segment the moving objects. Another method, the principle of which is presented in [3], is to detect motion using optical flow. But the optical flow cannot be calculated locally, because only one independent measurement is available from a sequence of images at a given point, while the flow velocity has two components.

Usage concepts of models, methods, and algorithms for understanding the structure of motion are presented in [4–12]. After analysis of mentioned works, it is possible to confirm that the decision of a problem can be based on classical algorithms as well as deep learning approaches.

Mathematical models for finding the change in motion should be used in simplified environments of sports analytics [4, 5]. In paper [6], was proposed an approach based on the estimation of the spatiotemporal boundaries of the detected events by maximizing the normalized Laplacian spatiotemporal operator in space and time scales. Continuation of this work can be a combination of any mathematical classifiers [7, 8]; semi-hidden models [9]; generative models [10, 11]. The disadvantage of these works is the lack of a comprehensive description of the ordering and placing of the features consistently over time.

Papers [12, 13] aims to use temporal structures to understand the motion based on dynamic Bayesian networks and Markov models. The disadvantage of these works is the complicated design [14], detailed training data, which can be expensive to collect [15]. The solution to this problem is presented in [16], but the paper does not care about the time of object motion.

From the literature review, it is established that each approach to understanding the structure of motion in sports analytics has its advantages and disadvantages. Therefore, the development of models for understanding the movement in sports analytics, changes their parameters, combining algorithms to support each other's strengths, and highlighting their weaknesses in understanding the structure of motion has a scientific and practical interest.

2 Presentation of the Main Material

2.1 Model

The chosen concept involves the usage of image segmentation. "Mask-RCNN" architecture [14] was chosen for this task. This is a popular and accurate model based on neural networks of two-stage detection with a prediction of the mask (Fig. 1).

First look at the generation of a mask, the main element for which, as for any network that works with images, is the convolution operation (1).

$$O(i, j) = \sum_{k=1}^{m} \sum_{k=1}^{n} I(i + k - 1, j + l - 1) * K(k, l) \tag{1}$$

where

i runs from 1 to $M - m + 1$, and j runs from 1 to $N - n + 1$,
$M * N$ is the dimension of the input image.

The convolution layer of the network allows to combine values of adjacent pixels and find more common features of the image. To do this, a small square matrix (in Mask-R CNN it is 3×3, 1×1 pixels), which is called the kernel, is consistently applied to the image. Each element of the kernel has its weight, which is multiplied by the corresponding values of the pixels of the image. The sum of these products is the result of the convolution operation. It should be noted that the coefficients of the convolution kernel are finding during the training of the model, while the number, size, and position of the kernel are determined before training.

The parts of the input image that may contain an object pass through a set of successive convolutional layers paired with the ReLu activation layer. Activation
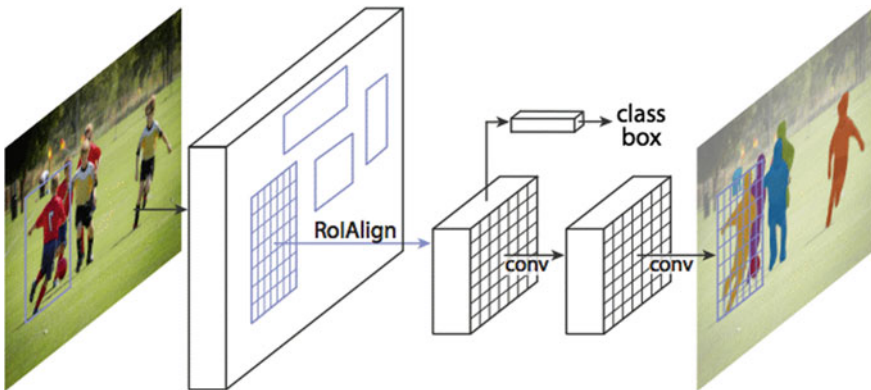

Fig. 1 Mask-RCNN architecture [14]

in the neural network is an attempt to emulate the work of neurons in the brain of living beings. In nature, the signal from one neuron goes to another under certain conditions, in ReLu such a condition is to compare the value with zero. If the result after the convolution is greater than zero, then the unchanged number is returned, otherwise, it will be set to zero.

As a result, the mask and the model's confidence in the prediction are returned. The described procedure is performed for each class the model can predict and one mask with the highest level of confidence is selected. Since there will be only one class—for humans, the mask for the human position will be compared with a mask consisting of only zeros and means the absence of a human.

2.2 Dataset

The "COCO" dataset [15] was used in this article. It is designed for object detection, segmentation, human keypoint detection, material segmentation, and signature generation tasks. For now, its authors have collected more than 330 K images. Not all of them are annotated with masks, so a smaller amount of data is used for the segmentation task. Usually, a training subset is used for the training process, but for this work, another part was used. First, 79 categories that not people, were ignored. Second, a validation dataset was used. This choice was made since the model must learn on new images. Nowadays learning neural networks do not require the training weights from scratch. To facilitate future research, deep learning libraries contain pre-trained models. New models begin to learn not with random coefficients, but with already known ones. Since weights can already determine some patterns in the data, it is much faster to get an acceptable result for new classes and images. This approach is called transfer learning. Therefore, the data for the new model was selected from the validation subset, and the weights were pre-trained using training data. Thus, the model obtained during this work is more focused on the images of people from the test set.

In total, 2693 annotated images were used. Additionally, the image was divided into 3 groups in the ratio of 7.5: 2: 0.5, for training, validation, and testing subsets. The training set contains 2019 photos, and the validation—539. The model saw them both during the training process, but the weights were updated only using training samples. The final learning results were tested on a smaller set, which contains 135 unknown to the neural network images.

The process of learning masks requires a file, where the true coordinates to the people in the images will be. A single data description format is accepted for working with the "COCO" dataset. Because the data was filtered and can no longer be described by the source file provided by the "COCO" authors, a new annotation file was created. Each image is described by a separate dictionary (data structure). Annotation example:

{ "segmentation": [[480.0, 96.71, ... , 477.58, 243.13],
[479.64, 297.78, ..., 479.64, 333.87]],
"iscrowd": 0, "image_id": 0, "bbox": [359.0, 59.58, 121.0, 342.35],
"area": 15618.319299999996, "category_id": 1, "id": 1713044 },

"category_id" represents the class number of the object (will always be equal to 1), "bbox" is the bounding box of the object, "image_id" is image number, "iscrowd" is a label indicating whether there is a cluster of objects in the image, "area" is the area occupied by the object. The most important field is "segmentation". It must be an array of arrays, each of them contains consecutive pairs of numbers denoting the x and y coordinates. By connecting these points, the contour of the object can be obtained. Although the purpose of the model is to predict the mask, providing binary mask images for all instances is inefficient in computational terms, so simplified notation is used. Nevertheless, the result of applying the network to an image is still a binary matrix.

For working with the "COCO" dataset was used a cocoapi [16]. It helps to download, analyze, and visualize annotations. The API supports several annotation formats, so this software product was used to make it easier to work with large datasets. The source code is freely available on "GitHub", so anyone can use it.

2.3 Training Parameters and Metrics

Detectron2 is a powerful tool for working with neural networks [17]. This is a popular library of computer vision models which allows users to integrate the latest computer vision technologies into the workflow. The advantage of "Detectron2" is that it is distributed with an "Apache License" [18], which permits to use of the software for any purpose, freely distribute, and modify it. This library is based on "PyTorch", a framework for working with neural networks. The advantage of using such tools is that they allow you to speed up processes by computing on GPUs. "Google Collaboratory" provides a limited number of hours of GPU free usage, so the model was trained in this environment.

This work is based on the "Mask R-CNN", and the "Detectron2" includes its implementation. In addition to previously created models, new ones can be created thanks to the Detector2. To do this, several steps should be done: dataset registration, model selecting, and training configurations description.

The dataset can be registered by adding the path to the appropriate folders in the DATASETS.TRAIN and DATASETS.TEST variables. MERGE_FROM_FILE and MODEL.WEIGHTS indicate the architecture of the model and the weights by which it should be initialized, these fields should reference the actual files in the Detectron2 library.

Some others important hyperparameters: IMS_PER_BATCH is the number of images in one batch (5 was used); MAX_ITER parameter that characterizes the

number of iterations (500 was used); BATCH_SIZE_PER_IMAGE is responsible for the number of different objects that can be found by the model in one photo (128 was used); BASE_LR determines the initial learning rate (0.002 was used).

The authors of the "Mask R-CNN" model proposed the following loss function (2):

$$L = L_{cls} + L_{box} + L_{mask} \tag{2}$$

where

L_{cls}—loss function for classification,
L_{box}—loss function for bounding box detecting,
L_{mask}—loss function for mask prediction.

The L_{mask} metrics its most important for one class image segmentation. It is the average binary cross-entropy for masks (3):

$$L_{mask} = -\frac{1}{m^2} \sum_{i,j=1}^{m} \left[y_{ij} \log \hat{y}_{ij}^k + (1 - y_{ij}) \log(1 - \hat{y}_{ij}^k) \right] \tag{3}$$

where

y_{ij}—ground truth class label for the pixel at (i, j);
$\hat{y}_{ij}^k$—is the value that the model predicted for the same pixel for a particular class k.

Another important value for evaluating our model is a metric called COCO mAP [19]. To describe it, some other concepts need to be clarified.

Intersection Over Union (IOU):

$$IOU = \frac{area(M_p \cap M_{gt})}{area(M_p \cup M_{gt})} \tag{4}$$

where

M_p—mask predicted by model,
M_{gt}—ground truth mask.

True Positive (TP): object detected correctly.
False Positive (FP): object detected incorrectly.
False Negative (FN): object not detected, but it is present.
True Negative (TN): object not detected, and it is absent.

These metrics depend on the threshold value. Each time IOUs are compared to a threshold and belong to one of these categories.

Precision is the fraction of relevant instances among the retrieved instances (5), and recall is the fraction of relevant instances that were retrieved (6).

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{all\ detections} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{all\ ground\ truths} \tag{6}$$

The Precision $\times$ Recall curve is a graphic way of estimating a model. For each threshold value from 0 to 1, Precision (5) and Recall (6) are calculated and represented in pairs on the graph as axes. Average Precision (AP) is defined as the area under the Precision x Recall curve. It is a number from 0 to 1. Mean Average Precision (mAP) is the average AP for all classes. For the COCO dataset, count 10 of these values for thresholds from 0.5 to 0.95 with steps of 0.05. Thus obtain a single metric (7).

$$mAP_{coco} = \frac{mAP_{0.50} + mAP_{0.55} + \ldots + mAP_{0.95}}{10} \tag{7}$$

To better understand the model performance, the values for the threshold of 0.5 and 0.75, as well as mAP_{coco} (7) for objects of different sizes are highlighted. For predictions of less than 32^2 pixels, more than 96^2 pixels, and in between.

Segmentation accuracy can also be useful:

$$segmentation\ accuracy = \frac{TP}{PT + FP + FN} \tag{8}$$

2.4 Evaluating the Model

Figure 2 shows the learning results for several metrics: total loss function; loss function for masks; learning rate; FN; FP; accuracy. These data were obtained for a training dataset with 8215 people instances in 2019 photos. There is a fast increase or decrease in the first 200 iterations, then change insignificantly. This shows the correctness of the selected number of steps of the learning algorithm. A smaller number would not lead to the best network, and a larger one would take a long time without model improvement.

The best value for the total loss function 0.6776 and masks only loss function—0.2234. Although the overall value is quite large, it is better to focus on the segmentation metric. The pixel-wise accuracy is 0.9023, and the losing of objects level is 0.08572.

The mAP_{coco} metrics for test and validation datasets described in Table 1.

For each value, the dataset, which was not seen by the model, gives better results. It indicates that the network has studied the general tendency to determine the human positions. For the IOU 0.5 threshold, the model makes the best prediction, so this threshold will be used in the future.
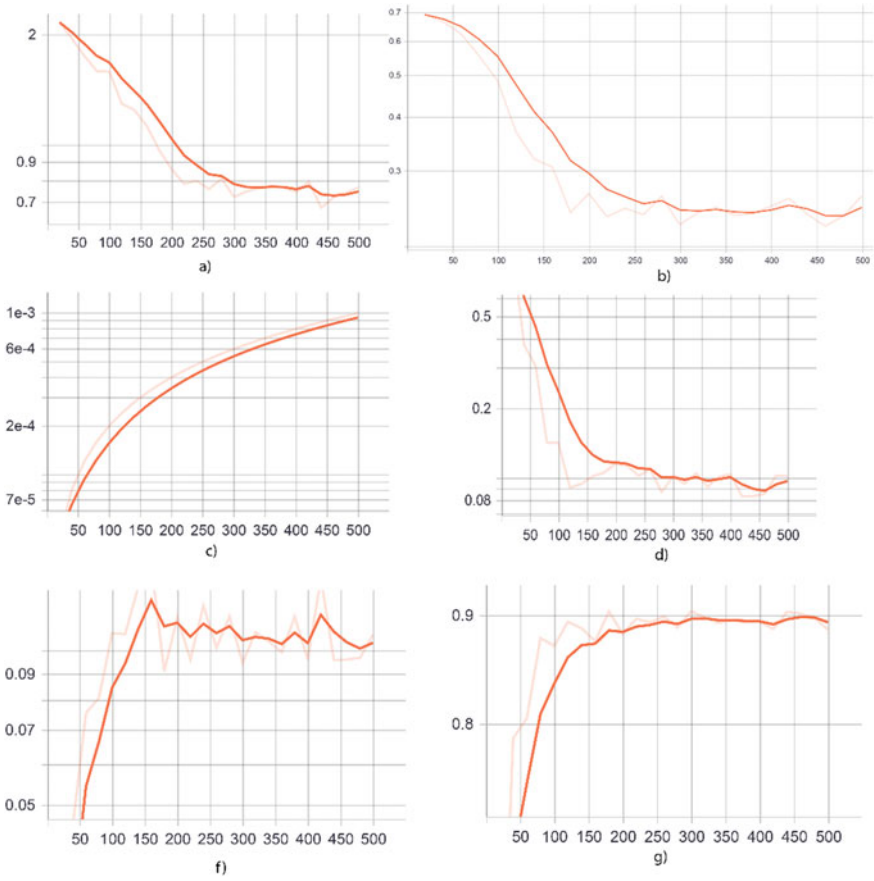
Fig. 2 Model estimation **a** total loss function, **b** loss function for masks only, **c** learning rate, **d** FN, **f** FP, **g** accuracy

Table 1 mAP$_{coco}$ for validation and test sets

AP		AP50		AP75		APs		APm		APl	
Val	Test	Val	Test	Val	Test	Val	Test	Val	Test	Val	Test
53.4	59.2	81.9	86.4	58.8	65.5	37.0	41.3	61.3	65.0	70.3	74.4

The model determines the mask best for large objects, so it should not be used to search for people on a small scale.

2.5 Proposed Algorithm Description

Motion analysis involves obtaining a path that has been passed by an object. The proposed method reveals a visual representation of the motion. In sports analytics, it is important to see the trajectory of the athlete on the field, not the coordinates of the pixels where he was. Therefore, the result should be an image showing the path. To do this, data from the neural network is used. Accordingly, each video frame is processed in a prepared neural network. For each received prediction of the person, it masks saved. The matrix for storing all masks has the dimension of the processed video. In the end, each value describes the number of all human appearances in the corresponding pixel of the resulting image. This algorithm is shown in Fig. 3.

For visualization, it is necessary to turn the received matrix into the image. Obviously, for videos longer than 255 frames, it is not possible to convert the number of human appearances to pixel brightness. Therefore, all values greater than 1 should be placed in the range from 1 to 255. To do this, each number is divided by the largest number in the matrix, multiplied by 255, and rounded to an integer.

For a color representation, a color map is applied to the image. A color map is a set of pairs for each possible pixel brightness value. It is created for each channel separately, so 3 separate sets for red, green, and blue channels should be created. The green color shows the athlete's short stay in a certain location, yellow shows the average, and red shows the longest.

The last step is mask blurring for smoother motion. Applying convolution with Gaussian filter (9) as a kernel returns a blurred image.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-\left(x^2+y^2\right)}{2\sigma^2}} \qquad (9)$$

where

x, y—the distance from the start of axes to half the size of the kernel minus one in both directions,

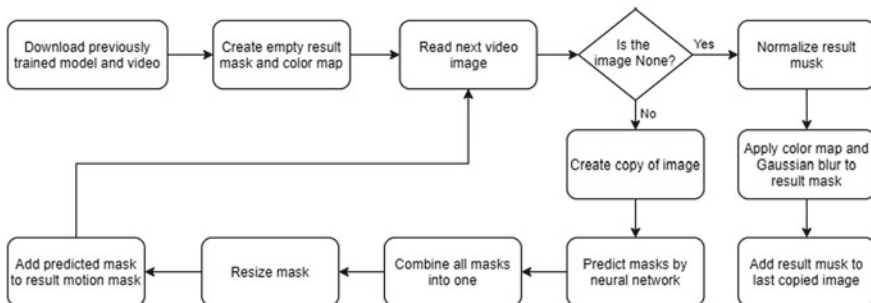σ—standard deviation of the Gaussian distribution.

Fig. 3 UML diagram of the proposed algorithm

The mask overlay the source video frame get another informative way to visualize the movement. That can be achieved with Formula (10):

$$dst = \alpha \cdot img1 + \beta \cdot img2 + \gamma \qquad (10)$$

where

α, β, γ—coefficients in a range from 0 to 1. Will be used 0.7, 0.5, and 0.
img1, img2—video frame and color mask.

The proposed algorithm returns a static image that represents the total trajectory of the player. This method involves the use of fixed cameras so that the result can be interpreted. For moving cameras, it is more appropriate to depict the operation of the algorithm on video using the last few masks predicted by the model.

2.6   Algorithm Operation Examples

The Python programming language was used to implement the process described above. The "OpenCV" library is used for convenient work with the image.

Several short videos from different sports, such as tennis, gymnastics, and football, were selected to test the algorithm and visualize its results. The top left contains an original photo, top right—photo with color map, bottom left—binary mask, and bottom right—color mask.

The first test video consists of 449 frames of 1920 × 1080 size. The process of creating the result took 230 s, which equals about 2 frames per second. The video shows a short part from a football match [20]. Results of the algorithm are presented in Fig. 4. The camera filmed the game from a large distance, so the players on the field were small.

The model has successfully determined people's positions. Since several players were close to each other, the result is difficult to interpret due to the crowds. It can be identified that in this video the movement is not active and most of the players are standing in their places most of the time.

Next is a video of a tennis match (see Fig. 5).
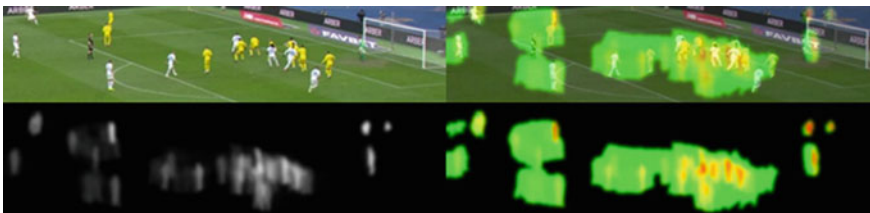
Fig. 4 The results of the proposed method for a fragment of a football game
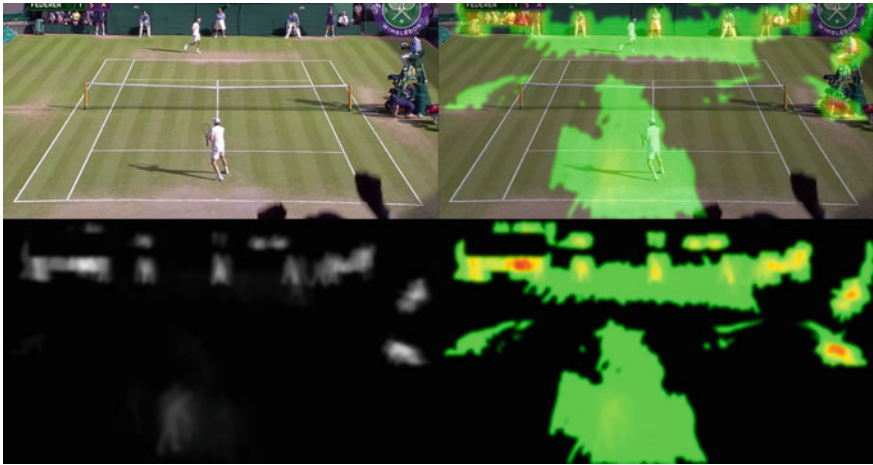
Fig. 5 The results of the proposed method for a fragment of a tennis game

The video consists of 1108 frames, size 1920×1080 [21]. Processing took 568 s, which is equivalent to 2 frames per second. The results show that the algorithm recognized the viewers rather than the players. The trajectory of the player closest to the camera was determined, but in comparison with static spectators, he was in each place for a short time, so his path is all green. The combination of the color mask and original frame helps to determine the area of his movement on the field.

To avoid the previous problem with viewers, for the video of the gymnast at the Olympic Games [22], only the largest mask was used for visualization (Fig. 6). Speed and size of video the same as previous ones. In a 100 frames video, the girl moves from one side of the bar to the other. The algorithm depicted her path correctly and clearly. It is possible to estimate the speed of movement: at first quite slowly, as
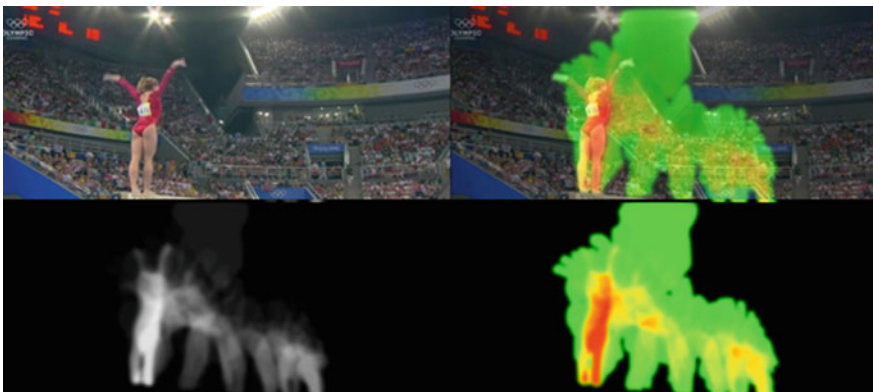
Fig. 6 The results of the proposed method for a fragment of a gymnast's performance

evidenced by the semi-saturated color, in the middle the speed increased, which is highlighted in green, at this time the girl gained speed for jump.

In the end, she completes the jump and stops at one point, which is highlighted in red. The jump itself can also be determined.

This result was the best in terms of interpretation, so it can be concluded that the described method of motion estimation is suitable primarily for short videos with a small number of non-static objects.

3 Conclusion

Presents examples of testing the algorithm and visualization of its result. method work on some short videos from different sports, such as tennis, gymnastics and football. The resulting image of these examples clearly shows the change in the position of objects over time. Runtime is approximately 2 FPS. The final images can be used to analyze sporting events that happened in the past. Potentially, the algorithm can be useful for athletes whose results are critically dependent on their location. For example, to determine the leader in a race where it is difficult for a person to determine the winner. The method works with pixels that accurately show the boundaries of objects, and will not make a mistake, because it calculates who crossed the finish line first.

The algorithm shows the possibility of using neural networks to segment people in order to represent their movement. The prospect is the development of an information system that can be integrated into the software environment of sports analytical projects.

References

1. Shinkaruk O, Blazhko N (2020) Development of motor skills in athletes of different sexes and their importance in the system of training in cheerleading. Theory Methods Phys Educ Sports 1:34–41. https://doi.org/10.32652/tmfvs.2020.1.34-41
2. Harrouss OE, Moujahid D, Tairi H (2015) Motion detection based on the combining of the background subtraction and spatial color information. Intell Syst Comput Vis (ISCV) 5:1–4. https://doi.org/10.1109/ISACV.2015.7105548
3. Shafie AA, Hafiz F, Ali MH (2009) Motion detection techniques using optical flow. Int J Electr Comput Eng 3(8):1555–1557
4. Quattoni A, Wang SB, Morency LP, Collins M, Darrell T (2007) Hidden conditional random fields. IEEE TPAMI 29:1848–1853
5. Rodriguez MD et al (2008) Action MACH a spatio-temporal maximum average correlation height filter for action recognition. IEEE Conf Comput Vis Pattern Recogn:1–8
6. Laptev I (2005) On Space-Time Interest Points. IJCV 64:107–123
7. Actions in context. https://hal.inria.fr/inria-00548645/file/MarszalekLaptevSchmid-CVPR09-ActionsContext.pdf. Last accessed 21 March 2021

8. Khlevna I, Koval B (2020) Fraud detection technology in payment systems. In: IT&I 2020—Information technology and interactions. Proceedings of the 7th international conference "information technology and interactions" (IT&I-2020). Workshops proceedings, Kyiv, Ukraine, 2–3 Dec, 2020. CEUR Workshop Proceedings, pp 85–95

9. Wang Y, Mori G (2009) Human action recognition by semilatent topic models. IEEE TPAMI 31:1762–1774

10. Khlevna I, Bura Y (2020) House price modeling by machine learning. In: Snytyuk V (ed) Information technology and interactions (satellite): conference proceedings, 4 Dec, 2020, Kyiv, Ukraine. Taras Shevchenko National University of Kyiv, Kyiv, pp 124–126

11. Niebles JC, Wang H, Fei-Fei L (2008) Unsupervised learning of human action categories using spatial-temporal words. IJCV 79:299–318

12. Leveraging temporal, contextual and ordering constraints for recognizing complex activities in video. IEEE Conf Comput Vis Pattern Recogn:1–8. https://doi.org/10.1109/CVPR.2007.383074

13. Understanding videos, constructing plots learning a visually grounded storyline model from annotated videos. https://www.cis.upenn.edu/~jshi/papers/storyline.pdf. Last accessed 21 March 2021

14. Mask R-CNN. https://arxiv.org/abs/1703.06870. Last accessed 21 March 2021

15. Microsoft COCO. https://arxiv.org/abs/1405.0312. Last accessed 20 March 2021

16. CocoApi, https://github.com/cocodataset/cocoapi, last accessed 2021/03/18.

17. Detectron2. https://detectron2.readthedocs.io/en/latest/. Last accessed 18 March 2021

18. Apache license. http://www.apache.org/licenses/. Last accessed 18 March 2021

19. CocoEval. https://cocodataset.org/#detection-eval. Last accessed 20 March 2021

20. Dynamo—movement. Studio and full match. https://youtu.be/u13zdNe-U0g?t=4960. Last accessed 20 March 2021

21. The best game ever? Murray v Federer. https://youtu.be/CGRzfUccmNE?t=452. Last accessed 15 March 2021

22. Women's individual artistic gymnastics all-around final—Beijing 2008. Throwback Thursday. https://youtu.be/d2E85PtfykA?t=952. Last accessed 18 March 2021

Quasi-Inflection-Based Part-of-Speech Tagging of Texts for Expert Formulations Content Ambiguity Detection in Modeling Domain Knowledge

Vitaliy Tsyganok , **Mykhailo Dubok** , **and Olha Tsyhanok**

Abstract The quality of the recommendations provided as a result of the application of decision support systems largely depends on the quality and reliability of the knowledge provided by experts. Solving the problem of automatic ambiguity detection in the textual expert formulations is a significant step towards increasing the reliability of knowledge and the adequacy of the models on the basis of which decision support is provided. Most approaches to automatic ambiguity detection rely on the use of part-of-speech tagging as the first step in ambiguity detection. The article proposes an automatic part-of-speech tagging method based on quasi-inflections (variable word components), the accuracy of which is comparable with the present rule-based approach implementations. The advantages of the rule-based approach include a significant reduction in the required amount of information, simple implementation of tagger improvements and a high degree of components' portability (rules, dictionaries, quasi-inflections, exceptions). Comparing the reported accuracy of part-of-speech taggers, Markov models and the transformation approach achieve an accuracy of up to 97%. At the same time, the accuracy of the rule-based approach varies from 97 to 100%. The proposed method's novelty is the use of quasi-inflections as the main and only method for determining the part of speech and grammatical characteristics of the word. Testing was conducted on 7 sets of textual formulations. The proposed method showed an accuracy of 98.70%. The paper provides a comparison between the accuracy of the proposed method and a Ukrainian Universal Dependencies model using Stanza for part-of-speech tagging.

Keywords Decision-making support · Domain knowledge model · Part-of-Speech tagging · Quasi-inflection · Text ambiguity · Ambiguity detection

V. Tsyganok (✉) · M. Dubok
Institute for Information Recording,, National Academy of Sciences of Ukraine, Kyiv, Ukraine
e-mail: tsyganok@ipri.kiev.ua

V. Tsyganok
Faculty of Information Technology, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

O. Tsyhanok
Department of Foreign Philology and Translation, Kyiv National University of Trade and Economics, Kyiv, Ukraine
e-mail: olzyg@ukr.net

1 Introduction: Analysis of the Problem Situation

One of the stages of the decision support process in various fields is the group construction of the domain knowledge (DK) model in the form of an appropriate knowledge base. Based on the knowledge gathered in such a model, both explicit (well-known to professionals) and expert (gained through the experience and intuition of certain narrow specialists), decision options for decision makers (DM) can be generated and evaluated. Since a significant share of knowledge in any field belongs only to experts, the use of expertise is very important to fully and adequately reflect all the properties of the domain knowledge in the knowledge base and further use them by decision support systems (DSS) to generate high-quality recommendations for a DM [1, 2].

For DSS of all classes, the stage of obtaining (collecting) expert information is mandatory [1], which precedes the stages of concordance [3, 4] and aggregation of information of different types [5, 6], and it is at this stage that experts provide knowledge both in the form of quantitative and qualitative assessments and in the form of textual formulations. It is during the formulation of a list of criteria, goals, factors, etc. that a misunderstanding can begin, which will result in a misinterpretation when building a knowledge base: a hierarchy of criteria, relationships and/or influences between the components of the system that is being modeled.

Formulations of the main goal, sub-goals during its decomposition, programs, projects, measures, solutions as well as any expert formulations use natural (human) language, which is characterized by ambiguity at all language levels. Formulations that may be perceived differently by different stakeholders due to ambiguity threaten to reduce the adequacy of the models and increase development costs and development time when corrections and changes are needed. The adequacy of the created DK models directly affects the quality of DSS recommendations. Therefore, it is necessary to reduce possible ambiguity during DK modeling for supporting decision-making.

To reduce the ambiguity of textual information, the following methods are used to process it. The most common way is ambiguity resolution—an automatic process that determines which meaning (interpretation) is more likely. However, given the risk of misinterpretation and the number of people involved in the process who will operate with such an interpretation, this is not an acceptable option for a DSS. The non-automatic method of ambiguity processing is represented by four different techniques by which it can be implemented: ambiguity avoidance (writing instructions), ambiguity prevention (writing in a fixed format), ambiguity detection (automatic detection in written text) and ambiguity correction (semi-automatic means of correction that interact with users) [7]. Ambiguity avoidance is not very effective, because even with the instructions provided, users rarely really make sentences less ambiguous [8]. Ambiguity prevention, of course, sets limits that reduce the chance of writing ambiguously, but it is also a rather restrictive technique. A significant number of scientific publications are devoted to methods of recognizing ambiguity

by people, but there are far fewer automatic implementations of detecting ambiguities. It is expedient to detect ambiguity of formulations during construction of DK models for decision-making support.

The ambiguity detection method is going to be used in two functional subsystems of expert DSS: the subsystem for obtaining knowledge about DK from experts (when writing expert formulations) and the subsystem for processing and generalization of expert knowledge. Such DSS include, in particular, the Solon-3 Decision Support System [9] and the Consensus-2 Distributed Collection and Processing of Expert Information System [10].

2 Discussion

In the scientific literature, issues of content ambiguity have been raised in many perspectives.

Gleich et al. [11] used lexical analysis and part-of-speech (POS) tagging to detect ambiguities of all four types, but all adjectives and adverbs are considered in the method as potential ambiguity, which gives too many false positive results. In this context, it is important to create our own method of lexical analysis of the formulations provided by experts. The basis of this analysis is POS tagging.

POS tagging or grammatical tagging is the process of assigning POS tags to words in a text [12]. The tagging process consists of three stages: tokenization, morphological analysis (assigning possible tags) and ambiguity resolution (selection of the most probable tag) [13].

Traditionally, one distinguishes 2 main approaches to automatic morphological analysis: the rule-based (constraint-based) approach and the statistical or probabilistic approach, better known as stochastic [13–16]. Currently, the transformational approach is also specified as the third of the main ones [12, 17]. Some scholars believe that the third main approach is the distributive one [18].

The rule-based approach uses contextual information to limit the number of possible POS tags [19] or to define a POS tag of a word [20]. For example, if in the English text the word is preceded by an article and followed by a noun, the tag "adjective" will be assigned. In addition to contextual information, morphological information is often used. For example, if a word is preceded by an auxiliary verb and the word ends in -ing, the "verb" tag will be assigned. Some taggers take into account case and punctuation [15]. Such information is useful in different languages. For example, in English, the case helps to define what the word "us" is: a pronoun or the acronym for "United States" (US). In German, for example, all nouns begin with an uppercase letter.

In the statistical (stochastic) approach, the most probable tag is selected on the basis of statistical data obtained during the analysis of unambiguously marked text. The frequency of words or the probability of n-grams is used as a criterion for maximization. The most common algorithm for implementing the n-gram approach

is the Viterbi algorithm. A combination of probabilities of certain sequences of tags and frequency of words is more complicated [16].

Brill's transformational approach uses machine learning, combining rule-based and probabilistic approaches. Like a rule-based approach, transformational learning is based on rules. Like the probabilistic approach, rules are automatically retrieved from data (a text corpus). One of the most widely used transformational tagging tools is the tagger developed by Brill [21].

Distributive tagging is an approach that eliminates the need to use manual rules and marked learning corpora that may not be available for certain languages or fields. Unlike the rule-based approach and the stochastic approach, distributive tagging is performed completely unsupervised. Schütze [22] proposed the analysis of distributive word patterns by constructing a matrix of adjacency of terms followed by a singular decomposition of this matrix to identify latent dimensions. In the space of reduced dimension provided by the singular arrangement of the matrix, the tokens are indeed intuitively grouped by POS. Given the context, it is possible to achieve results similar to POS tagging.

When comparing a rule-based approach and a stochastic approach, the probabilistic approach is often chosen because of the possibility of automatic learning and lower time costs [14]. Also, the stochastic approach is more effective in the analysis of long sentences than the rule-based approach [23]. Moreover, the probabilistic approach, in particular using the Markov model, is better to use in cases where the set of possible tags (tagset) is small in size. If the set of tags is large, the rule-based approach is more effective [24].

Comparing the claimed accuracy of POS taggers, Markov models and the transformation approach achieve an accuracy of up to 97% [21, 25]. At the same time, the accuracy of the rule-based approach varies from 97 to 100% [26].

E. Brill identified the following advantages of the rule-based approach over the stochastic approach: a small amount of stored information; clarity of a set of content rules; ease of finding and implementing improvements; greater degree of portability from one set of tags, corpus' genre or language to another [20]. However, the degree of portability of the rules is still low, because one needs to create different rules for each language.

Given the need to obtain correct rules that correctly determine the POS tag, the possibility of automatic learning is not a priority. Text formulations in DSS are of different, but mostly small length. Considering the volume of sets of possible tags, in addition to the common POS in each language (noun, verb, etc.), the size of each set depends on the specific language. The accuracy of each of the three approaches (rule-based, statistical and transformational) is roughly identical, but the ability to achieve a slightly higher score is important when choosing a tagging approach, as the next stage of automatic ambiguity detection—syntactic—strongly depends on POS tagging. Therefore, it is the rule-based approach that is a priority for solving the problem in the field of expert decision support. However, a hybrid approach at a certain stage of development is not excluded, for example, the use of probability as a last resort in cases where the number of possible tags is not reduced to one, despite the use of rules. In studies [15], the usefulness of the hybrid approach was

experimentally tested: having combined a morphological analyzer with statistical data from a corpus, the average efficiency was increased by 15.53% compared with the use of only a morphological analyzer.

In addition to the used approach, POS tagging is classified into supervised (automated) and unsupervised (automatic) tagging.

The system, introduced by Brill in 1992, is considered by the author and some scholars as a kind of rule-based approach [14]. After all words in the dictionary are assigned the most frequent tag, the module guesses the word tag according to its affixes (prefix, suffix). Then context rules are applied [20]. This method is acceptable for use as a basis.

In addition to the accurate definition of the correct tag, a rule-based approach can be used to identify a list of possible tags, then the definition of a particular POS is done by another module [15]. However, if information about POS of neighboring words is provided, this approach is able to perform both tasks. Thus works an existing constraint-based tagger by reading a sequence of words and alternative tags and passing it to a grammar converter, which leaves only one tag for each word based on contextual information [13].

Without resorting to parsing, contextual analysis is possible by analyzing word sequences [16]. For example, in the sentence "in large cities the level of air pollution is much higher" it is possible to distinguish the sequence of words "in large cities", which can be analyzed independently of the other words in the sentence. Thus, there is accurate information that the word "in" is always a preposition in all texts except colloquial and literary styles. If it is known that the word form "large" is an adjective, and the word form "cities" is unknown, then this sequence may correspond to the sequence of tags "preposition + adjective + noun". The same sequence of tags can be assigned if it is known that the word form "cities" is a noun, and "large" is unknown. Such sequences of words are called n-grams.

Brill demonstrated that a rule-based approach without any knowledge of syntax can be as effective as a stochastic approach. The tagger developed by him initially assigns the most probable tag to each word based on an analysis of a large marked corpus without regard to context. Then two procedures are performed: (1) words that are absent in the training corpus and begin with a capital letter are considered proper names; (2) the rest of the words that are not in the training corpus are assigned the most frequent tag for the last three letters of the word. This simple algorithm has a low error rate of 7.9%. Training implies making a list of words with their most frequent tag. Next, rule templates are used, which significantly reduce the number of errors made before applying the templates, and lead to a small number of new errors. The process of testing new templates is easy, as templates that generate rules that give mostly incorrect results are not included in the final list of templates. Applying only 71 templates, the error rate was reduced to 5.1%. Among these templates, 66 of them reduced the number of errors, 3 did not work, and 2 increased the number of errors. Brill also noted the ability of a rule-based approach to make more accurate POS tagging in idioms through automatic learning. Thus, the only knowledge gained non-automatically is the procedure of identifying proper names [20].

The disadvantage of Brill's approach is the low speed of operation compared to the stochastic approach. Increasing the speed while maintaining efficiency is possible by converting the program based on transformations into a deterministic finite converter [14]. However, this negates an important advantage of the tagger—the simplicity of its construction, which Brill emphasized [20].

Brill's approach, later called transformational, is still a hybrid of rule-based and stochastic approaches and achieves an accuracy identical to the probabilistic approach—97% [21].

Since the detection of ambiguity of formulations largely depends on the accuracy of POS tagging, it is advisable to use a rule-based approach with automated acquisition of rules.

Based on the above analysis, there is an urgent need to develop a method that has high accuracy. Since statistical and transformational approaches achieve an accuracy of no more than 97%, a rule-based approach is promising.

3   Research Objective

Improvement of the reliability of knowledge and adequacy of models, on the basis of which decision support is provided, through automatic ambiguity detection in textual formulations of experts. The development of a highly accurate method of POS tagging is one of the components of detecting ambiguity of textual formulations.

4   Formal Statement of the Problem

What is given: $W = \{w_i\}$, $i = (1, n)$—set of words of expert formulations, $P = \{p_j\}$, $j = (1, m)$—set of names of parts of speech, $D = \{d_k\}$, $k = (1, l)$ is a set of dictionaries, where words have appropriate grammar information, and n, m and l are the powers of the corresponding sets (number of elements in sets).

Needed to define: Mapping $W \rightarrow P$.

4.1   The Proposed Method

To solve the problem, a method is proposed, which, unlike existing methods, is based exclusively on the use of lists (dictionaries) of rules and quasi-inflections. Due to the use in the method of constructing dictionaries of automatic retrieval of elements of the set W that do not reflect $W \rightarrow P$, automatic inversive (reverse) alphabetical sorting of unmapped elements and automatic retrieval of all elements of the set W containing a quasi-inflection, whose POS correspondence is being checked, a high level of compliance is achieved in a certain mapping.

The method of constructing dictionaries of rules and quasi-inflections is a procedure that involves automated verification of the validity of rules and quasi-inflections. It can be performed arbitrarily or systematically. If the procedure is performed arbitrarily:

1. The rulemaker (RM) automatically retrieves the first element from the set W, which does not have the mapping $W \rightarrow P$.
2. The RM manually determines the corresponding element of the set P.
3. The RM manually hypothesizes that all elements of W, which contain the quasi-inflection, isolated from this element of W, must be matched to the same element of P.
4. The RM manually creates a corresponding rule of the type Condition $\rightarrow$ Result in the form of "if RuleType text tag", writing in a text file with the rules (described below).
5. The RM automatically (using the software) obtains a text file with all the unique elements of W that meet the condition of the rule, sorted in inverted alphabetical order.
6. The RM manually reviews all the obtained elements of W in case of mismatch of the element of the set P.
7. If a discrepancy is found, the RM manually

   a. clarifies the rule by increasing the quasi-inflection limit to the limit where the subset of W will not contain elements that do not correspond to the element of the set P specified in the rule, or
   b. adds additional rules before the current rule to cover all elements of the subset of W that do not match the element of the set P specified in the rule.

8. The RM repeats the procedure again for the next element of W, which does not have the mapping $W \rightarrow P$.

If the procedure is performed systematically:

1. The RM automatically retrieves all unique elements of the set W that do not have the mapping $W \rightarrow P$, sorted in inverted alphabetical order.
2. The RM manually allocates in order an inseparable sequence of elements W, which correspond to 1 element of the set P.
3. The RM finds a common quasi-inflection in all elements of the selected sequence.
4. The RM manually creates a rule based on a common quasi-inflection by writing it in a text file with the rules (described below).
5. The RM repeats the procedure from step 2 for the next sequence of elements of the set W.

The procedure and methods for compiling a dictionary of quasi-inflections are identical to compiling a dictionary of rules. The only difference is that for the dictionary of quasi-inflections the only condition is that the word should end with a certain quasi-inflection.

Solution method:

1. The preliminary stage (described above).
2. Check the element of the set W for full match. In case of match, go to step 5.
3. Obtain an element of the set P, i.e. the analyzed class (POS), via analysis by rules.
4. Obtain the analyzed subclass, i.e. grammatical characteristics present in each element of the set P, by analysis.
5. By gradually truncating the final letters, try to find the basic form of the element of the set W in all elements of the set D. In the absence of a match, go to step 6.
6. Add information from appropriate dictionaries to the element of mapping $W \rightarrow P$.
7. Repeat steps 1–5 for each element of the set W.
8. Return the mapping $W \rightarrow P$.

The method is implemented in the form of an automatic POS tagger based on quasi-inflections, using a rule-based approach.

When using the tagger, the user enters text in a dialog box, launches automatic tagging and receives information cards for each word. Each information card contains: (1) the entered word; (2) its analyzed class and subclass, if they could be identified; (3) the lemma; (4) a list of possible dictionary options, if any. Each dictionary option contains its own class and subclass. For example, for the entered word "вулиць" the information card will contain the analyzed class "іменник", the analyzed subclass "мнЖ" (plural of feminine), fuzzy match with the lemma "вулиця" and one dictionary option that provides a feminine noun singular.

In order to make the tagger work correctly, the preliminary step should be performed. It consists of creating a database of rules for determining the main classes (POS), which is carried out by the RM, and filling in the dictionary, which is compiled by a lexicographer or imported. The list, names and structure of files containing dictionaries and rules are not fixed and can be changed as needed. At the preliminary stage, a list of quasi-inflections for determining subclasses (grammatical characteristics) is also compiled.

The procedure for compiling rules can be arbitrary, for example, in case of accidental discovery of a new unrecognized word, or systematic.

The systematic approach implies automatic obtaining a list of all unrecognized words, from which the compiler has the ability to analyze and further derive rules in an automated mode.

In both cases, the basis for compiling a rule is an unknown word form, which is demonstrated to the person who compiles rules (compiler) in the development mode. The compiler creates a rule of type "if RuleType text tag", where "RuleType" is a type of check with a certain list ("ends", "equals", "does not equal", "begins"), "text" is the part of the word form to which the rule applies, "tag" is an assigned POS from a defined list ("іменник", "прикметник", "дієслово", etc.). For example, according to the rule "if EndsWith ба іменник", if the word form ends with "ба", the tag "іменник" is assigned. Thus, the hypothesis is made that all words ending with the letters "ба" are nouns.

Each hypothesis should be tested before making full use of the new rule. To confirm or refute the hypothesis, when checking the rule "if" is changed to "ifLog" for tracking, after which the compiler launches an automatic analysis of the text corpus. All cases of the rule execution are recorded. To filter the text, duplicate word forms are removed from the final report. After analysis, the results are displayed in inverted alphabetical order. Also, an extendable list of exceptions that are not recorded in the report is applied.

To increase the accuracy of the rules, i.e. to limit their application, the compiler can combine them using the logical operator "AND". For example, the rule "if EndsWith аві AND NotEquals праві іменник" assigns the tag "іменник" to all words that have the quasi-inflection "аві" except the word form "праві".

The POS of most words, namely those that cannot belong to several POS or to one POS with different grammatical features (for example, different cases), is uniquely established using only quasi-inflections, i.e. the last letters of words that may be smaller than, equal to or greater than the morpheme ending, in particular a quasi-inflection may be equal to the whole word. That is, a quasi-inflection is a variable part of the word, starting from the end.

To reduce the number of rules, a certain class is considered as the main, and the rest—as exceptions. For example, by default, words ending with the letter "г" in the Ukrainian language are considered to be nouns. But words that have the quasi-inflection "міг" or "сяг" are assigned the tag "дієслово". The basic rule is placed last in the list so that the exception rules are checked first. Thus, at the time of publication, the developed system uses 804 rules, which determine a POS.

In addition to grapheme analysis using rules, a fuzzy match is used for words that differ from the lemma form, such as "слова" (the lemma is "слово"). This study implemented an algorithm that provides for the gradual truncation of the last letters of word forms in a textual formulation and a dictionary, which is the implementation of stemming (trimming) of variable parts of words [24], which is used for many natural language processing needs, including content monitoring systems [27]. To avoid false matches, for example, the noun "раз" in the dictionary for the adverb "разом" in the formulation, it is possible to compare the analyzed POS with the POS in the dictionary. Also, to increase the number of correct matches, grammatical features are compared. For example, if the noun "президент" occurs in the text, and the word form is compared in the dictionary with the existing noun "президенство", it is necessary to compare the gender of nouns. Thus, the coincidence will not occur, because by analyzing the word form "президента" masculine will be obtained, and in the dictionary the noun "президентство" has the neuter gender. Fuzzy match can cover up to the last 6 characters.

In this implementation, for both full and fuzzy match, 16 dictionaries with a total volume of 5030 registered words are used. If the dictionary contains a word from the expert text formulation, used in the lemma form, the analysis by the rules and fuzzy match are not carried out. Brill's approach follows the same principle [17]. The disadvantage of this principle is the incorrect definition of the tag in cases where the word form occurs in the text not in the lemma form, but there is a formal match with another word. For example, in the phrase "промислового робота", the word

"робота" will be assigned a feminine tag instead of a masculine one, because there will be no fuzzy match.

Recognition of the number and gender of nouns in the formulation is also realized by analyzing a quasi-inflection. To better match the plural with the correct lemma form, separate quasi-inflections are distinguished for each of the three types of the plural: (1) plural whose lemma form is masculine; (2) plural the lemma form of which is the feminine; (3) plural the lemma form of which is the neuter gender. At the initial recognition of the plural, the noun is checked for the presence of a quasi-inflection of any type of plural.

In order to compile rules and lists of quasi-inflections, our own Ukrainian-language corpus of official style texts with a volume of more than 137,000 words was compiled. The corpus includes textual formulations in 7 goal hierarchy structure files, which were entered into the Decision Support System "Solon-2" [28], texts of 14 UN conventions (on the Law of the Sea, on Contracts for the International Sale of Goods, on the Rights of the Child, Against Corruption, against transnational organized crime, on Human Rights, on the Elimination of All Forms of Discrimination Against Women, on Biological Diversity, on the Rights of Persons with Disabilities, against Torture, etc.), one contract, a court decision, more than 40 passages.

Because speech is a dynamic phenomenon, all rules are based on factual cases, not theoretical principles. For example, despite the theoretical knowledge that in official style texts most words are nouns, not all quasi-inflections get basic rules according to which the word form is assigned the tag "іменник". In some cases, most words ending in a quasi-inflection can be adjectives or verbs. Moreover, there are letters that words cannot end with. No word forms ending with the letters "г" or "ц" were found in the training corpus. Thus, following the de facto principle of compiling making, it is possible to avoid redundancy.

5 Results

The novelty of the method is the use of quasi-inflections as the main and only method for determining the POS and grammatical features of the word.

To test the effectiveness of the proposed method, test was conducted on the basis of textual formulations in 7 goal hierarchy structure files provided to the Decision Support System "Solon-3" [9] and used in collective modeling of DK in the System of distributed collection and processing of expert information "Consensus-2"[10], with a total volume of 4378 words. The results of manual tagging were compared with the results of automatic tagging, considering manual tagging as correct. The markup uses 17 main class tags: "іменник", "прикметник", "прийменник", "сполучник", "абревіатура", "дієслово", "прислівник", "займенник", "дієприкметник", "дієприслівник", "частка", "число", "скорочення", "числівник", "помилка" (misspelling), "сполучнеСлово" (e.g., "як"), "інше" (for example, foreign words). Word class tags for manual tagging are coordinated with the information provided by

the resource "Словники України online" of the Ukrainian language and information fund of the NAS of Ukraine. Each text unit was assigned a tag based on the context.

During the testing, 59 automatic tagging errors were detected, while in the manual tagging there were 2 words with the tag "error", i.e. these are words that were misspelled in the text. To offset their impact, the number of text words that contain errors should be subtracted from the number of automatic markup errors and the total number of text words.

POS tagging accuracy "is defined as the ratio of the number of word forms correctly tagged over the total number of word forms tagged" [29]. Because some textual formulations contain errors, the following formula was used to determine the accuracy of A:

$$A = 100\% - \frac{x - y}{n - y} \times 100\%$$

where A is the accuracy, x is the number of errors in the automatic tagging, y is the number of applications of the class "помилка" in the manual tagging, n is the number of words.

Substituting the available values, we have:

$$A = 100\% - \frac{59 - 2}{4378 - 2} \times 100\% = 98,697\% \approx 98,70\%$$

Thus, the tagger showed an accuracy of 98.70% for POS tagging of text formulations in the test corpus. Words that have received the "error" tag in manual markup are not taken into account when determining accuracy, as a POS tagger is not a program for detecting and correcting errors. The tag obtained by analysis was chosen as a priority. If the analyzed tag is missing due to a full match or inability to determine, the POS tag was obtained from a dictionary.

Most of the words whose class could not be determined are abbreviations.

The demonstrated high accuracy of the method can be achieved with strict adherence to the described method of compiling dictionaries of rules and quasi-inflections.

Such accuracy is comparable with state-of-the-art POS tagging results (Table 1):

6 Validation of the Method

Currently, the format of the developed POS tagger is partially compatible with the format of TreeTagger, which has an accuracy of about 96% [30, 31], and incompatible with the format of POS tagging Universal Dependencies [32]. Since the Universal Dependencies project is the most promising and universal format for presenting sentence structure using the dependency grammar, it is necessary to either change the format of POS tagging of the proposed method or compare the output data of the

Table 1 POS tagging accuracy

Technology (method, approach or tagger)	Accuracy (%)
Ukrainian-language model in Universal Dependencies: the used tagger	91
TreeTagger	96
Markov models (probabilistic/statistical)	97
Transformation approach	97
Ukrainian-language model in Universal Dependencies: Stanza check	96.77
Ukrainian-language model in Universal Dependencies: claimed	97.5
Proposed method	98.70

Universal Dependencies format with the data obtained by the proposed method by establishing correspondence between the units of POS tagging of two formats.

When the latter option is selected, the text is first analyzed by both the proposed method and an external tagger using the Universal Dependencies format, and then the results of POS tagging are compared to match the names of POS. For example, class "іменник" matches "NOUN" in the format of Universal Dependencies. In cases where the match is broken, for example, an external tagger recognizes the noun as an adjective and marks it with the tag "ADJ", one can compare the correctness of the results. Such cases may occur because the accuracy of the compared POS taggers differs. For example, the accuracy of the POS tagging stated by the authors of the Ukrainian-language model in Universal Dependencies format is 97.5%, while the accuracy of the morphological analyzer stated by the authors is 91%. At the same time, POS tagging of the Stanza neural analyzer [33], written in Python, using the same model has an accuracy of determining POS of 96.77%.

It is possible to compare the results manually. The relevance of this is that POS tagging itself can also be ambiguous. Thus, the proposed method can be used to improve the overall process of ambiguity detection.

The Stanza neural analyzer with the help of separate modules allows to sequentially perform tokenization, POS tagging, lemmatization, parsing and named-entity recognition (for example, "Земля", "Київ", etc.). The accuracy of parsing directly depends on the accuracy of POS tagging.

Stanza modules for tokenization, including multi-word tokenization, POS tagging and lemmatization were used for comparison. The POS tagging module requires modules for tokenization to be loaded. The lemmatization module is used to facilitate manual comparison, but is not required for POS tagging.

Tokenization is the first step as well in the developed on the basis of the proposed method tagger. But then lemmatization and POS tagging can constantly call each other. This is due to the fact that an incomplete match in the dictionary also finds a lemma of a certain word form and can use it to increase the accuracy of POS tagging. At the same time, POS tagging is used to improve lemmatization.

Having received 2 annotated texts in different formats, their comparison is performed, where POS names of one format correspond to POS

names of another format, for example the name "ADJ" in Universal Dependencies corresponds to the name "прикметник" in the developed tagger, "ADV"—"прислівник", "NOUN"—"іменник", "VERB"—"дієслово", "ADP"—"прийменник", "PRON"—"займенник", etc.

The comparison revealed minor differences due to the format rules. First, in Universal Dependencies, acronyms usually belong to a certain POS and are indicated as a feature, and in the proposed method, an acronym is an independent class on a par with a POS. For example, the word "ЗМІ" is assigned the tag "NOUN" in Universal Dependencies, but the tag "абревіатура" in the proposed method. Second, punctuation in Universal Dependencies is always considered as a separate token and is denoted by the "PUNCT" tag, and in the proposed method, punctuation is considered either part of a form (but not a lemma), or a separate token if not adjacent to words, and then denoted by the "пунктуація" tag. For example, if the text contains "цілей:", Universal Dependencies distinguishes 2 tokens, and the proposed method—only one. Third, the symbol "*" (star) has the tag "SYM" in Universal Dependencies and "пунктуація" in the proposed method. However, there are also cases of incorrect analysis, especially of acronyms. In particular, for the word "ПТІС" the Stanza model identified the tag "PUNCT" twice and "NOUN" once, for the word "ПТ"—in most cases the tag "NOUN", the tag "PUNCT" once, the tag "PROPN" once. At the same time, in the proposed method in all cases the word has the tag "абревіатура". Tokens "132", "134", "140", "141" are identified by the Stanza model with the tag "ADJ", and the proposed method assigns the "число" tag. No other differences were found.

7 Conclusions

The method of POS tagging based on quasi-inflections is presented, which has an accuracy of 98.70%, comparable with the existing modern methods that use a rule-based approach, but at the same time differs by the consistency of means used, specifically quasi-inflections and fuzzy match. Given the small number of tools needed to achieve high accuracy, the proposed method is suitable for use in the field of automatic detection of ambiguity of expert formulations in DK modeling.

The proposed method and the analyzer developed on its basis are largely portable. Portability is the ability to use the developed system of dictionaries, exceptions, rules and their types and number-generic lists of quasi-inflections for other languages, but their lists will be different for each language. Completing each component does not require changes to the program code.

The proposed method and the tagger developed on its basis are largely portable. The portability implies the ability to use the developed system of dictionaries, exceptions, rules and their types and number-gender lists of quasi-inflections for other languages, but the lists will be different for each language. Filling each component does not require changes to the program code.

Another advantage of the proposed method, as well as any implementation of the rule-based approach, is the lack of need to store large statistical tables.

Further research is planned to be performed using the proposed method at the next stage of automatic ambiguity detection—syntactic one, which is based on information obtained at the lexical stage, in particular POS tagging.

References

1. Totsenko V (2002) Methods and systems for decision-making support. In: Algorithmic aspect. Naukova dumka, Kyiv
2. Saaty T (1996) Decision making with dependence and feedback: the analytic network process. RWS Publications, Pittsburgh
3. Tsyganok V, Kadenko S (2010) On sufficiency of the consistency level of group ordinal estimates. J Autom Inf Sci 42(8):42–47
4. Tsyganok V (2013) Providing sufficient strict individual rankings' consistency level while group decision-making with feedback. J Model Manag 8(3):339–347
5. Totsenko V, Tsyganok V (1999) Method of paired comparisons using feedback with expert. J Autom Inf Sci 31(7–9):86–96
6. Tsyganok V, Kadenko S, Andriichuk O, Roik P (2018) Combinatorial method for aggregation of incomplete group judgments. In: Proceedings of 2018 IEEE first international conference on system analysis and intelligent computing (SAIC), Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine, pp 25–30
7. Alomari R, Elazhary H (2018) Implementation of a formal software requirements ambiguity prevention tool. Int J Adv Comput Sci Appl 9(8):424–432
8. Winkler S (2015) Ambiguity: language and communication. De Gruyter, Berlin/Munich/Boston
9. Certificate of state registration of copyright to work № 8669. Computer program "SOLON-3 decision support system" (SOLON-3 DSS), Ministry of Education and Science of Ukraine State Department of Intellectual Property (2003). Accessed http://www.dss-lab.org.ua/Solon-3.htm
10. Certificate of copyright registration for work № 75023. Computer program "A system for distributed collection and processing of expert information for decision support systems—"Consensus-2"", Ukraine (2017). Accessed http://www.dss-lab.org.ua/Consensus-2.htm
11. Gleich B, Creighton O, Kof L (2010) Ambiguity detection: towards a tool explaining ambiguity sources. In: International working conference on requirements engineering: foundation for software quality 2010, LNCS, vol 6182. Springer, Heidelberg, pp 218–232
12. Pisceldo F, Adriani M, Manurung R (2009) Probabilistic part of speech tagging for Bahasa Indonesia. In: 3rd International MALINDO workshop, Singapore, pp 1–6
13. Chanod J, Tapanainen P (1995) Tagging French—comparing a statistical and a constraint-based method. In: 7th Conference of the European chapter of the ACL, Dublin, Ireland, pp 149–156
14. Roche E, Schabes Y (1995) Deterministic part-of-speech tagging with finite-state transducers. Comput Linguist 21(2):227–253
15. Altunyurt L, Orhan Z, Güngör T (2007) Towards combining rule-based and statistical part of speech tagging in agglutinative languages. Comput Eng 1(1):66–69
16. Altunyurt L, Orhan Z, Güngör T (2006) A composite approach for part of speech tagging in Turkish. In: Proceeding of international scientific conference computer science, pp 1–6
17. Sajjad H (2007) Statistical part of speech tagger for Urdu. National University of Computer & Emerging Sciences, Lahore, Pakistan
18. Chew P, Bader B, Rozovskaya A (2009) Using DEDICOM for completely unsupervised part-of-speech tagging. In: Proceedings of the NAACL HLT workshop on unsupervised and minimally supervised learning of lexical semantics. ACL, Boulder, Colorado, pp 54–62

19. Karlsson F, Voutilainen A, Heikkila J, Anttilla A (1995) Constraint grammar: a language independent system for parsing unrestricted text. Mouton de Gruyter, Berlin
20. Brill E (1992) A simple rule-based part of speech tagger. In: Proceedings of 3rd conference on ANLP. ACL, Trento, Italy, pp 152–155
21. Brill E (1995) Transformation-based error-driven learning and natural language processing: a case study in part-of-speech tagging. Comput Linguist 21(4):543–565
22. Schütze H (1995) Distributional part-of-speech tagging. In: Proceedings of the 7th conference of the European chapter of the ACL, Dublin, pp 141–148
23. Weischedel R, Meteer M, Schwartz R, Ramshaw L, Palmuzzi J (1993) Coping with ambiguity and unknown words through probabilistic models. Comput Linguist 19(2):359–382
24. Tapanainen P, Voutilainen A (1994) Tagging accurately—don't guess if you Know. In: Proceedings of the fourth conference on applied natural language processing. ACL, Stuttgart, Germany, pp 47–52
25. Hardie A (2004) The computational analysis of morphosyntactic categories in Urdu. PhD Thesis, Lancaster University, UK
26. Voutilainen A (1995) Morphological disambiguation. In: Karlsson F, Voutilainen A, Heikkilä J, Anttila A (eds) Constraint grammar. De Gruyter Mouton, Berlin, Boston, pp 165–284
27. Lande D (2014) Elements of computational linguistics in legal informatics. Research Institute of Informatics and Law of the National Academy of Legal Sciences of Ukraine, Kyiv, Ukraine
28. Certificate of state registration of copyright to work PA №4137. Computer program "SOLON-2 Decision Support System" (SOLON-2 DSS), Ministry of Education and Science of Ukraine State Department of Intellectual Property (2001). Accessed http://www.dss-lab.org.ua/Solon-2.htm
29. Dybkjær L, Hemsen H, Minker W (2007) Evaluation of text and speech systems. Springer
30. Schmid H (1994) Probabilistic part-of-speech tagging using decision trees. In: Proceedings of the international conference on new methods in language processing, Manchester, UK, pp 44–49
31. Schmid H (1995) Improvements in part-of-speech tagging with an application to German. In: Proceedings of the ACL SIGDAT-workshop, Dublin, Ireland, pp 47–50
32. Gerdes K, Guillaume B, Kahane S, Perrierm G (2019) Improving surface-syntactic universal dependencies (SUD): surface-syntactic functions and deep-syntactic features. In: Proceedings of the 17th international conference on Treebanks and Linguistic Theories (TLT). SyntaxFest, Paris, pp 1–7
33. Qi P, Zhang Y, Zhang Y, Bolton J, Manning C (2020) Stanza: a python natural language processing toolkit for many human languages. In: Proceedings of the 58th annual meeting of the ACL: system demonstrations. ACL, pp 101–108

Dynamic Malware Detection Based on Embedded Models of Execution Signature Chain

Ihor Karpachev and **Volodymyr Kazymyr**

Abstract Today, the most mobile devices, such as tablets, e-books, watches, fitness bracelets and etc. operate under the Android OS. At the same time the Android OS can be considered as unstable environment mainly because there are a lot of untrusted sources of mobile applications potentially containing dangerous code which is executing side-by-side on the Android Runtime (ART) as user's applications. In this paper new methodology is considered for determining the accuracy of malicious software (malware) assessments to provide the application functional safety. Dynamic functional tracing tool is used to build API call chains. Sequence of API calls chains compared used local and global alignment techniques from bioinformatics. Code samples from Google Play and Malware Genome Project were used for emulating safe and dangerous API calls.

Keywords Mobile applications · Functional safety · Android OS · Malware · API calls · Application execution signature chain

1 Introduction

Mobile devices are at the epicenter of modern growing demand of people staying interconnected and being able to solve their everyday tasks online. That is why, mobile devices captured more than half of the market of computing and communication systems, which led to extreme urgency of general protection of mobile applications, user data and functional security.

The rapid development of mobile applications, developed for Android OS, has highlighted the vulnerabilities of this system and necessitated the improvement of tools that can provide the appropriate level of functional safety. Analysis of current

I. Karpachev (✉)
Datascope Systems Ltd., Access House, Aviation Park, Flint Road, Chester, UK
e-mail: ikarpachev@stu.cn.ua

V. Kazymyr
Chernihiv Polytechnic National University, 95 Shevchenka Str., Chernihiv 14035, Ukraine
e-mail: vvkazymyr@stu.cn.ua

publications [1, 2] indicates an exponential growth of cyber threats for mobile applications of the Android system, while automatic methods of detecting and classifying malicious algorithms do not allow to avoid this type of threat. It should be noted that machine learning approaches of artificial neural network algorithms that allow to determine the level of threat based on basic parameters of program code, for example, when analyzing API calls, also do not provide adequate cyber-attack tracking efficiency due to sensitivity to code syntax or code obfuscation.

Most of existing security measures provided by OS Android are based on preventive actions and system restrictions in order to ensure platform security in general. Considering significant delay in resolving security issues and potential risk to lose private data or even disruption in functional safety of the system there is a necessity of an extra safety unit on top of existing general security system, which will help to notify user of potentially malicious software at the executing stage in runtime.

Functional and information security are two fundamental components of mobile application's safety that complement each other. The definition of information security is to ensure the availability, integrity and confidentiality of system data, while the functional security unit ensures the correct execution of the functions of the control system and transfer of control objects to a safe state in the event of system failures. The system analysis of mobile applications must include a functional safety unit that monitors potentially dangerous conditions and identifies relevant events that could lead to data loss, access to the confidential data of third parties or blocking device access by third parties.

2   Related Work

A systematic analysis showed that the problems of Android security are related to not perfect privileges system, which is convenient and effective when user installing new applications (Android OS display application permissions) [3–5], but leaves attackers access to functional nodes and sensitive data. Large amount of research has studied how to detect malware prior to installation. These researches can be roughly divided into static and dynamic analysis categories. For instance, TaintDroid, DroidRanger, DroidScope [6] can monitor behavior in runtime, where systems like Kirin [6, 7] identify malicious software using static analysis. Both methods have pros and cons, for instance dynamic runtime identification creating extreme overhead and not always can be directly applied to mobile device. By contrast, static analysis software is not causing such runtime performance issues but cannot be scaled properly due to fact that they are mainly build on manually crafted detection patterns.

Data leakage was discovered by Ulm university researchers where default Google Sync and Google Calendar apps transmits token in unencrypted way. By using this token user can login to any Google Service. This issue was fixed in Android OS 2.3+ by using software patch [8]. Google as a main contributor of Android every new version of OS restricting and improving access to resources such as

file system, executing permissions category, transitive kernel access, these actions reducing operational surface for attackers but not completely eliminates it.

In contrast with Google there are a lot of 3rd parties researches have been done in area of improving security of existing Android OS, one of them is DroidScope which is type of a Virtual Machine Introspection (VMI), which can be classified as a dynamic analysis framework. Contrasting other dynamic analysis frameworks, it's not located inside the emulator but creating OS virtual machine and OS-level semantics by locating outside the emulator. In this case, even in the kernel, the privilege escalation attacks can be detected. It makes any attackers task very hard. DroidScope is built, with rich set of APIs available to human analysts.

Other interesting solution is DroidRanger detection framework, which is based on two-stage analysis in order to detect zero-day malware and existing well-known malware. Detect framework normally utilize permission-based behavior footprint, building behavioral scheme and compare with existing knowledge database.

Another modern approach which is using neural projection method is used in order to provide characterization of malware families [9]. Method is interesting not only because machine learning techniques has been applied to such problem, but also decision tree is built in order to visualize and analyze such malicious software even better. In general, method is not applied to one application but to the family of applications instead.

There is another category of static malware detection tool—android permission-based analysis. The basis of such systems is analysis of permissions based on latent semantic indexing (LSI: Latent Semantic Indexing). LSI is commonly used technique to match queries to set of documents. Today, LSI can be considered as a standard method of matching higher order textual terms to documents (or any other compared entity, where textual blocks can be extracted). The technique is based on the calculation of a matrix in which rows are given by code elements and words, and columns—by documents. In order to search for relevant documents, the matrix is reduced using the method of singular value decomposition (SVD: Singular Value Decomposition) [10]. This method is applicable as a malware preventive tool due to performing prior to installation.

Traditionally most recent researches in this area are focused on application layer such as worms, viruses, cross-service attacks and similar. Nearly all of these vulnerabilities can be fixed by providing patches and updates to OS itself or specific application, but this leaves time room for attackers between discovering type of attack and actual deploying updates to physical device or server as mentioned previously. Software and hardware developers are normally reacting to the issue when it's already happened and have to find solution after system and/or user already got some damage. Automatic API calls chain generation can proactively protect and notify developers or users of potentially dangerous application behaviors in advance.

3   Functional Safety Analysis of Android OS

The API is the most sensitive component of the Android OS, which is protected through a system of access permissions (Android Permission), which determine the privileges of users and developers. To run critical functions of the software application for the developer, it is necessary to access the application manifest file Android-Manifest.xml with the subsequent consent of the user, who gives the application the appropriate privileges. Once permissions are granted and the application is installed on the mobile device, revoking them while the application is running can be quite difficult (in case of system permissions) and the system does not restrict the application from using its privileges. Most methods of automatic detection of malicious code by Android are divided into two general categories [11]:

- detection of malicious code based on signatures,
- detection of malicious code based on machine learning.

Signature-based approaches are based on finding specific patterns in bytecodes and API calls, which are easily circumvented by transforming program code at the bytecode level, machine-independent low-level code generated by the translator and executed by the interpreter. Machine-learning approaches remove patterns of program behavior (mainly, permission requests and critical API calls) and use standard machine learning algorithms to perform malicious code classification [12]. However, as the removed features are related to the syntax of the program, this type of detector is also not a reliable tool for protection against all type of threats. In terms or characteristics, today the effectiveness of software developed on the basis of these approaches depends largely on the quality of the developed detection scheme, characteristic of a particular behavior of software applications, so their specialization can be considered quite narrow.

A typical method of dealing with vulnerabilities in Android is an automatic patch system. Given that Android applications have not developed tools for automatic and effective vulnerability correction, a number of studies have been conducted to build algorithms for automatic patch generation for standard software applications running according to the client–server paradigm. Such algorithms analyze the program code, determine the cause of the vulnerability that led to the misuse of Android software and mobile device hardware, and then create a source code patch that can be used to temporarily fix the vulnerability without interfering with the operator process. To fix Android applications, it is necessary to use the bytecode rewriting technique to implement the patch code [13]. The main idea of this process is to overwrite the DEX file in order to establish all API calls and ensure compliance with the relevant security rules. Temporary blocking of Android applications can also be used to control potential security and privacy breaches. In addition, algorithms have recently been developed to allow hints to be placed in a bytecode program and thus protect access to resources. On the other hand, to solve problems associated with significant vulnerabilities in software applications, it is necessary to develop new techniques

that would effectively control the confidentiality of the information flow carried out by OS applications.

A characteristic feature of Android software applications is the rapid growth of program code due to updates and patches. At the same time, due to limited resources on mobile devices, there is a strict limitation on the size of applications, and thus the process of overwriting the bytecode requires an optimization phase.

4 Modelling of Application Execution Signature Chain

Development of a method for ensuring the functional safety of Android mobile device is a complex task. Every android application generates digital trace when using Android SDK in order to communicate with Android OS resources. Application execution signature chain (AESC) is a sequence of functions called from android application to Android OS via SDK. The aim of AESC is to model the chain of API calls and match it with patterns during malware detection.

The mathematical model of the chain matching can be represented in such way.

Let's denote $P = \{P_k\}$, $k = \overline{1, K}$—a set of API calls pattern related to malware application. Every pattern is a chain of API calls $P_k = \left(p_j^k \right)$, $k = \overline{1, K}$, $j = \overline{1, n}$. During detection we will get a chain of API calls which we will denote as a vector $C = (c_i)$, $i = \overline{1, m}$. The task is to find pattern that corresponds to chain C. In order to achieve it we repeatedly capturing the chain—C^q which is built within constant time quantum Q.

The API calls detection will be performed in two phases with use of two known algorithms from bioinformatics [14]. On the first phase the algorithm of local alignment by Smith-Waterman is used to select the suitable patterns from set of P solving the problem of finding similar areas in two sequences. This algorithm is convenient when comparing two sequences whose lengths differ significantly. Considering that at the beginning of quantization the lengths of the API calls chains may be insignificant, this circumstance is decisive. In this case $P_k = \left(p_j^k \right)$ and $C^q = \left(c_i^q \right)$ are the input sequences in cycle over k. On the every step of the cycle alignment is done using a similarity matrix $M_{m \times n}$ with elements $M_{i,j}$ that are calculated by the rule:

$$M_{i,j} = \max\left\{ M_{i-1,j-1} + w; M_{i-1,j} + g; M_{i,j-1} + g; 0 \right\}, \ i, j > 1 \qquad (1)$$

where w—score for match (+1) or mismatch (–1) of symbols in row i and column j, g—penalty for gap when transition is along row or column (g = –1). If variants give negative values then result equals 0.

Matrix is filling by calculation of the neighbor values (diagonal, upper and left) of the current cell beginning from zero cell. Note, that it is impossible to get negative values in the matrix since 0 is as lowest value. For the appropriate alignment the trace back procedure is used beginning from the cell with the highest value and moving

Fig. 1 Example of local alignment

<div align="right">

AWHSABCKVB   AWHSABKVB
I I I   I I I
AB –K   ABK
a) with gaps   b) without gaps

</div>

Fig. 2 Example of global alignment

<div align="right">

AWHSABKVB
I I I I I I   I
AWH–ABK – B

</div>

to pre-positions while the cell with score 0 will be reached. An example of such alignment is shown in the Fig. 1.

Among all the patterns $P = \{P_k\}$, those that have the maximum score (without gaps) are selected. Such procedure is repeated each time the API call chain length increases by adding newly discovered calls. This reduces the number of patterns examined. As a result, one or more patterns $P^* \subset P$ will be selected.

On the second phase of detection the algorithm of global alignment by Needleman and Wunsch is used. It provides a comparison of sequences along their entire length and, in our case, is necessary to assess the coincidence of the fixed API calls chain with the selected pattern P^*. As in the previous algorithm, alignment is performed by constructing a similarity matrix $M_{m \times n}$. However, the rule for calculating cell values has a different form

$$M_{i,j} = \max\{M_{i-1,j-1} + w; M_{i-1,j} + g; M_{i,j-1} + g;\}, i, j > 1. \qquad (2)$$

The value of the lower right cell of the matrix will be the best alignment score of the two sequences. To construct the alignment itself, it is necessary to restore the calculation path using the trace back procedure. Figure 2 represents the example of global alignment.

Using the result of the global alignment, it is possible to calculate the percentage of coincidence of the evaluated chains. For example from Fig. 2 this is 78%. In future such estimate is used to define the stop area in the software detection method.

5 The Method of Dynamic Building Android Application Execution Signature Chain

Dynamic building of API call chains can be divided into several parts:

1. Building a database of API call chains patterns, malware application description, type of malware in order to notify user in human-readable format. Process of building patterns will be based on Malgenome dataset
2. All execution signatures will be divided into two major groups primary and secondary. Primary AESC is a chain which can cause disruption in work, so

a straight threat to functional safety—this is primary task of current work. Secondary AESC is a sequence which can lead to entire or partial data loss or unauthorize access by third party—therefore is a direct task of information security.

3. Dynamic application analysis could be made by any dynamic function tracing tool as long as it's suitable for building AESCs. In the scope of current work Frida framework was used.

4. Accumulation and fragmentation of dynamic API calls will be done during specific time interval or quantum of time Q, in order to find sequential pattern match from existing database.

5. Sending fragment of AESC with metadata (fragment of AESC + number of the sequence n) to server via API which will be injected by Frida framework.

Server API main task is a searching of fragmented AESC matching. If fragment match with any AESC from database—user notification required as well as further action in order to ensure functional safety.

Building of AESC patterns database. In the scope of current work Malgenome project was used as a source for building malware signature database. Malgenome dataset provides wide range of malware application's sequence of functions call and permissions requested by each malware application. In database each AESC contains API call chain, human readable format description and how it potentially could harm functional safety. If most of the API calls belongs to android permissions which potentially can cause disruption in functional safety—this AESC mark as primary, all other will be marked as secondary, because they can be identified as a threat for information security.

Simple example of building AESC could be provided by malware application which unlock or change android initial screen lock password by deleting or modifying data/system/gesture.key file (method allowed malware application to lock android OS and block actual user from using Android OS resources). Next step is to show modal dialog which cannot be closed and drawn over other apps (these dialogs normally contain threat or extortion message). First user reaction—full device reboot, but this will not help due to android.permission.REBOOT permission combined with Android Broadcast Receiver which is used by malware application. So even after full reboot malware application will get onBootComplete event and will be able to show blocking dialog again. Malware application required filesystem read/write permission in order to being able execute current behavior, along with draw over other apps permission. Related API functions are coded by pseudo symbols (Table 1) according to full set of required permissions that listed in Android manifest file:

- android.permission.WRITE_EXTERNAL_STORAGE
- android.permission.READ_EXTERNAL_STORAGE
- android.permission.SYSTEM_ALERT_WINDOW
- android.permission. ACTION_MANAGE_OVERLAY_PERMISSION
- android.permission.RECEIVE_BOOT_COMPLETED
- android.permission.SYSTEM_ALERT_WINDOW
- android.permission.REBOOT

Table 1 API calls sequence and pseudo symbol

Sequence number	API function	Pseudo symbol
1	FileInputStream.getChannel	A
2	FileChannel.map	B
3	File.delete	C
4	FileInputStream.close	D
5	Ljava.lang.Class.getField	E
6	File.exists	F
7	AlertDialog.Builder	G
8	alertDialog.setCancelable(false)	H
9	android.os.IBinder.bindService	I

Table 2 Malware application description in database

Id	APK	Description
1	Blob	App may block device screen, with reboot auto launch feature

Table 3 AESC of malware application

AppId	AESC	Description
1	ABCDEFGHI	Primary

Straight after reboot API call sequence of malware application will be repeated and intercepted by Frida. In malware database new record of AESC will be added (Tables 2 and 3):

Usage of Malgenome dataset provide wide range of similar API call chains and android permissions required to implement malware functionality. This approach can help to identify zero-day malware which is not flagged or reported as a malware in Google Play Store yet.

Dynamic analysis of APK using Frida. Frida is dynamic function tracing tool of android application—APK [15], which helps to define application behavior in runtime by injecting code into binary Dalivik executable files (DEX). Captured AESC will be normally executed during of time quantum Q and further analyzed by server application. At the first stage Q experimentally selected 60 s as initial time frame (value less then 1 min creates high network load and cause AESC to be more fragmented). This time frame will cause AESC to be fragmented and necessity of each request to be *stateful* and contains token t (sequence identity) of potentially malware application and sequential number of call chain n for server to match fragmented parts into uninterrupted AESC. Token can be obtained by applying function f to application A:

$$f(A) = t_{unique} \tag{3}$$

In scope of a device the name of application package is unique, but it can be same for different devices. Generated hash from device id and testing application package name will give fully unique token t. Next step, after identifying token and capture fragmented API call chain we can build server request containing pair (r, C_r), where r—sequential number of quantum Q, C_r—API call chain captured within the quantum. In simplest case server seeks full pattern match. If match found server notifies client of potential malicious nature of tested application and how it can disrupt functional safety. If there is no match found—no action needed. In more complicated scenario—if AESC has got fragmented match, server temporary record sequence number, token and chain (t, r, C_r). Every request with same token will be *stateful*, i.e. if during next request no match found state will be flushed, otherwise server will accumulate fragmented chains until full match found.

6 Experiments and Results

6.1 Evaluation Measures

In order to evaluate detection performance, known metrics derived from confusion matrix will be used [16]:

$$\text{True positive rate} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

$$\text{False positive rate} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{5}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{6}$$

$$\text{Recall} = \text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{7}$$

$$F - measure = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{8}$$

where true positive or TP and true negative or TN are values classified as malware correctly. False negative or FN and false positive or FP are values classified as malware incorrectly. Based on these metrics there are two classes: positive and negative measures which will help to understand true positive and false positive rate better (Table 4).

True positive rate or TPR is the value of predicted software classified correctly. False positive rate or FPR is the value of predicted classified incorrectly. These metrics sometimes shows in percentage. Precision, which is also called positive

Table 4 Confusion matrix classes

Class	Positive prediction	Negative prediction
Positive	TP	FN
Negative	FP	TN

predictive value is used for the rate of relevant results rather than irrelevant. Recall is a sensitivity for most relevant results. In order to combine recall and precision into single metric we used F-measure (sometime called F-Score). This value helps to measure two values at the same time (it uses harmonic mean in place of arithmetic mean). F-measure is a value that estimates the entire system performance and required because sometimes model can have high precision and low recall or vice versa which makes it hard to evaluate and compare these values separately.

Experiment results. Initially main experiments were planning to conduct by dynamically changing value of quantum Q and recording delta in evaluation measures, but due to implementation details of method of dynamic API call chains would lead to measuring method performance only, which is outside of scope of current work. There were conducted 3 experiments with 502 applications across all tests with 81.8% of malware applications among them. There are two main dynamic parameters: number of malware applications and percentage of malware application within single experiment. First experiment contains 100 applications (94% malware apps), second contains 306 applications (85.62% of malware apps) and third contains 117 applications (20.5% of malware apps). Results of three experiments depicts in Table 5.

Obtained results shown that with increasing percentage of malware applications, inside single experiment scope, there are insignificant decreasing of TPR value by 1% simultaneously with significant decreasing of FPR by 96.66%. Surprisingly, experiments shown that method behavior better in mixed environment where malware application more than 15% and with increasing amount of benign application is not changing significantly. Most of the application markets including Google Play Store are mixed environments where vast majority of the app is not malware which is similar to environment in experiment 3 with low FPR in comparison with experiment 1.

Results comparison. Current experiment results can be compared with results [17] obtained in similar work of android malware detection by analyzing device's network traffic which has been implemented by using machine learning algorithms (MLA). In series of two experiments, they assessed file classifiers: the decision tree (J48), Bayes Network (BN), K-nearest neighbors (KNN) and multi–layer perceptron

Table 5 Experiment results

Exp	Number	Malware (%)	TPR (%)	FPR (%)	Precision	Recall	F-Score
1	100	94	98	0.66	0.9587	0.98	0.9692
2	306	85.62	99	0.022	0.9961	0.9961	0.9961
3	117	20.5	96	0.021	0.92	0.9583	0.9387

Table 6 MLA and AESC results comparison

Method	Feature selection	TPR (%)	FPR (%)	Precision	Recall	F-Score
MLP	Yes	88.25	11.75	0.887	0.883	0.88
MLP	No	93.03	6.97	0.932	0.972	0.944
KNN	Yes	99.65	0.35	0.997	0.997	0.997
KNN	No	98.63	1.37	0.986	0.986	0.986
Exp. 3	–	96	0.021	0.92	0.9583	0.9387

(MLP). MLA with feature selection provides good statistic value of TPR, but with quite high FPR which degrade overall F-measure and makes usage of this method not always reasonable. Table 6 compare results with and without feature selection on some MLA techniques that performs on same malgenome dataset with different number of applications.

For instance, MLP with or without feature selection shown high TPR value 88.25% and 93.03% respectively with relatively low FPR value 11.75% and 6.97%—which means high error rate (by comparing F-measures from MLA experiments and experiments conducted in current work we can see that only harmonic mean showing overall method results), author of this evaluation also mention that detection time of malware application is extremely critical in case of mobile device. That is main reason for measuring method performance in "Evaluation of machine learning classifiers for mobile detection" [16] where was the attempt to decrease detection time by using feature selection. Method of pattern matching of AESC has no this drawback because all heuristic analysis and CPU intensive tasks are moved to server side where computing potential is much higher than on mobile device, so response time will consist from network speed and time required for full AESC matching or fragmentation. Another advantage of remote server is logic encapsulation along with impossible reverse engineering procedure in comparison with android client–based solution where even code obfuscation is not reliable way of hiding implementation. Once third-party will understand how this defensive unit works it will be easy to block or interrupt normal malware detection procedure.

7   Conclusion

OS Android required continues improvements and upgrades in terms of security due to high user coverage and growing popularity. Dynamic analysis systems are showing high efficiency and benefits in preventing damage from modern malicious software by constant runtime analysis of application's interaction with OS environment. There is no single solution to security issues where even typical obfuscation algorithms can significantly complicate the analysis of malicious software code and the process of generating an application execution signature chain. Based on obtained results, analyzing of dynamic API call chains method showed less error frequency

in a mixed application environment. Despite the fact that reaction time is much higher in comparison with OS updates and application's patches, this is not guarantee full device and user protection. At this point damage already can be done, that is why developed method, like any other dynamic analysis method will work better with some static analysis tool which makes system hybrid and taking best of two approaches. Proposed solution is operating on application layer which slows down overall detection. Also, physical layer of Android runtime can be extended by adding this dynamic analysis capabilities, so shifting this task to hardware level can improve overall security task performance.

References

1. AL-Zadjali BM (2015) A critical evaluation of vulnerabilities in Android OS (Forensic approach). Int J Comp Appl (0975–8887) 130(5):38–42
2. Lasheras F, Comminello D, Krzemien A (2019) Advances in complex systems and their applications to cybersecurity. Complexity, Hindawi, pp 1–2
3. Allen G (2015) Android security and permissions. In: Beginning Android, pp 343–354
4. Elenkov N (2014) Android security internals: an in-depth guide to android's security architecture, 1st edn. No Starch Press
5. Au KWY, Zhou YF, Huang Z (2012) PScout: analyzing the android permission specification. In: CCS'12: Proceedings of the 2012 ACM conference on computer and communications security, pp 217–228
6. Arp D, Spreitzenbarth M, Hubner M, Gascon H, Rieck K (2014) Drebin: effective and explainable detection of android malware in your pocket. In: Proceedings of annual symposium on network and distributed system security (NDSS). The Internet Society, pp 23–26
7. Ahmed O, Sallow A (2017) Android security: a review. Acad J Nawroz Univ 6(3):135–140
8. Chatterjee S, Paul K, Roy R, Nath A (2016) A comprehensive study on security issues in android mobile phone. Int J Inno Res Adv Eng (IJIRAE) 3:62–72
9. Vega R, Quintian H, Cambra C (2019) Delving into android malware families with a novel neural projection method. In: Hindawi Complexity, pp 1–6
10. Shahriar H, Islam M, Clincy V (2017) Android malware detection using permission analysis, SouthEastcon, pp 1–6
11. Kapratwar A (2016) Static and dynamic analysis for android malware detection. In: SJSU Scholar Work, pp 6–10
12. Wen L, Yu H (2017) An Android malware detection system based on machine learning. AIP Conf Proc 1864:020136
13. Zhang M, Yin H (2016) Automatic generation of security-centric descriptions for android apps. In: Springer briefs in computer science android application security, pp 77–98
14. Jones N, Pevzner P (2004) An introduction to bioinformatics algorithms. Massachusetts Institute of Technology. A Bradford book
15. Frida dynamic instrumentation toolkit for developers, reverse-engineers, and security researchers. https://www.frida.re/. Last accessed 2 Feb 2021
16. Tharwat A (2021) Classification assessment methods. Appl Comput Inform 17(1):168–192
17. Narudin F, Feizollah A, Anuar B, Gani A (2014) Evaluation of machine learning classifiers for mobile malware detection. Springer-Verlag Berlin Heidelberg

Cybersecurity User Requirements Analysis: The ECHO Approach

Vasilis Katos ⓘ **, Duncan Ki-Aries** ⓘ **, Shamal Faily** ⓘ **, Angel Genchev** ⓘ **, Maya Bozhilova** ⓘ **, and Nikolai Stoianov** ⓘ

Abstract Cyber defense requires research and investment in advanced technological solution as well as in the development of effective methods and tools for identifying cyber threats and risks. This implies a need for a well-defined process for user requirements elicitation. The paper presents a structured approach for the identification of cybersecurity knowledge and elicitation of user needs, based on the development of specific use cases. Employing use cases is an effective way to identify the cyber security gaps. Example use case descriptions of the attacks on a general computer network are given. The proposed use cases are analyzed within CAIRIS platform. The modelling process confirms that CAIRIS is a powerful tool to enrich the context of threat models and UML class diagrams. Also, the modelling with CAIRIS could support using security-by-design principles. The research is conducted under the activities of "The European network of Cybersecurity centres and competence Hub for innovation and Operations" (ECHO) project.

Keywords Cyber defense · User requirements elicitation · Use case analysis · CAIRIS modelling

V. Katos · D. Ki-Aries · S. Faily
Fern Barrow, Talbot Campus, Poole, Bournemouth, Dorset BH12 5BB, Great Britain, UK
e-mail: vkatos@bournemouth.ac.uk

D. Ki-Aries
e-mail: dkiaries@bournemouth.ac.uk

S. Faily
e-mail: sfaily@bournemouth.ac.uk

A. Genchev · M. Bozhilova · N. Stoianov (✉)
Defence Institute "Prof. Tsvetan Lazarov", Bul. "Prof. Tsvetan Lazarov" 2, Sofia, Bulgaria
e-mail: n.stoianov@di.mod.bg

A. Genchev
e-mail: a.genchev@di.mod.bg

M. Bozhilova
e-mail: m.bozhilova@di.mod.bg

1   Introduction

Cyber defense of critical systems and citizen is a challenging task at the national and regional level. Cyber attacks can devastate Critical Infrastructure organizations, such as those part of the Health Care, Energy, or Security sectors, where damage can lead to loss of life. Because such attacks may not be limited to a single organization or state, collaboration is necessary to address common cyber threats and challenges.

Recognizing the need for collaboration to address increasing cyber threat, the European Union (EU) established a programme to create a European cyber security ecosystem. The ECHO (European network of Cybersecurity centres and competence Hub for innovation and Operations) project is one of the four pilot projects, funded by the European Commission, to create a cybersecurity competence network [1].

ECHO involves 30 industry and research partners from 14 member states. It aims to strengthen the EU's proactive cyber defense, improve the technological capabilities of a secure digital market, and protect the European citizens against cyber attacks more generally. With such a diverse range of partners and ambitious objectives, a consistent approach for eliciting and specifying requirements is important. The techniques used or capturing requirements needs to be easy for different partners to adopt, and requirements need to be managed using affordable tools that are likely to remain maintainable both now and in the future.

The use-cases are widely used as a technique for eliciting requirements for software systems, because of the benefits they provide. They can be used to elicit and specify requirements a user's perspective, they are effective for communicating with stakeholders, bringing hidden requirements in the minds of the stakeholders to the surface where they can be specified [2]. Nasr et al. [ibid] illustrate such benefits by using use cases to specify requirements for the embedded systems in the avionics domain. Faily et al. [3] demonstrate the usability of the use cases approach designing security and usability into a non-trivial software system for a research and development project.

Because use cases are an effective way to identify the user needs and cyber security knowledge gaps in different sectors, Task 2.1 of ECHO used use cases to understand the scope and complexity of different problem domains, share knowledge of domain-specific business knowledge and threat models, including the modus operandi of actors and attackers. In doing so, they helped standardize and streamline the story-telling and narrative of attacks in the sectors of concern, using a structure that enabled (i) requirements elicitation for information sharing models, (ii) the development of the curricula and modelling for enhancing cybersecurity/cyberdefense skills through the federated cyber range, (iii) inputs for the identification of sector-specific and inter-sector security challenges, (iv) inputs for the identification of technological challenges and opportunities, (v) the scoping and development of demonstration cases.

In this paper, we present our approach for use-case driven requirements analysis for ECHO. In Sect. 2, we present the definitions adopted for the specific task's needs. In Sect. 3, we provide an overview of the main features of the CAIRIS platform [4],

which we used to model, analyze and validate the ECHO use cases. We illustrate the approach taken to elicit storylines and use cases in Sect. 4, and our approach for subsequently modelling and analyzing them with CAIRIS in Sect. 5. We conclude in Sect. 6 by briefly describing benefits and implications of our approach.

## 2 ECHO Approach for a Use Case Definition

The ECHO concept for representation of the cybersecurity vulnerabilities includes three components: scenario, storyline and use case. Each identified sector contains a single scenario, which aggregates multiple storylines. The storyline includes a set of use cases based on the common type of attacks or the common infrastructure (asset). This set of use cases is about user needs that should be addressed. Figure 1 presents the hierarchical relation between the three terms—scenario, storyline, use case.

The words scenario, storyline and use case are being used differently in different areas and perceived differently by different authors. For the goals of the ECHO project, we have used the following definitions:

- Use case—A use case describes an interaction between attackers and the system/systems under attacks. Each use case has the following mandatory attributes: a name, a unique identifier, and a step-by-step description of a basic course of action. Some use cases also describe the system's states at each step, exception conditions and variant paths. (adapted from "A technical discussion on Use Case Best Practices", 06/11/03, IBM) [5]. Although the definition originates from Software Engineering, the clear way in which it describes the attacker-system interactions and corresponding outcomes makes it suitable for cyber security needs analysis.
- Storyline—A storyline groups several use cases sharing common infrastructure.
- Scenario—A set of storylines, related to the activities of malicious actors aimed to damage, theft or destroy assets in a specific sector (domain).

One of the first challenges we encountered when developing a use case template was the ambiguity over how to write use case descriptions, and choosing the right level of detail. Several books have been written on appropriate ways to write a good use case description. Based on an analysis of the best practice, we concluded that

Fig. 1 Relations: scenario—storyline—use case
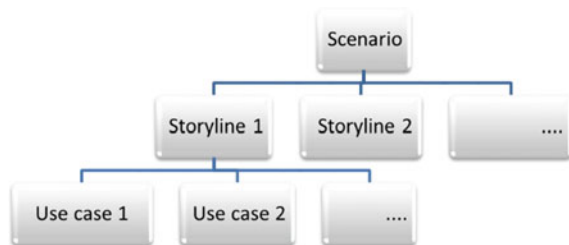
the usability of use cases is greatly enhanced by adopting a common structure for writing the descriptions. To address the project's objectives and needs, the template for ECHO use cases description was proposed. A template example for particular application is presented in Sect. 4.

3   Use Case Modelling with CAIRIS

Many tools support the management of use cases, but few support the management and analysis of complementary security design concepts used by ECHO.

CAIRIS (Computer Aided Integration of Requirements and Information Security) is an open-source platform for designing both security and usability into system specifications [4], and is a tool exemplar for how security and usability engineers might collaboratively address security and privacy problems at the earliest stages of the design process [6].

CAIRIS does not prescribe any particular design technique or methodology, and supports a wide variety of usability, security, and specification concepts. These include many of the concepts proposed in our approach. It supports the specification of use cases, assets and all the elements feeding into a risk analysis process. It also supports threat modelling using Data Flow Diagrams (DFDs) and attack trees. Faily et al. [7] present an approach for reasoning about tainted data flows in design-level DFDs.

CAIRIS supports a number of features which make it useful as a collaborative design platform. First, it supports the automatic generation of visual models, which can be dynamically created as model elements are added, updated, or deleted. Many visual models scale poorly because the bounded rationality bias makes it difficult to manage models that have become too complex to comprehend. By generating models, providing the ability to filter model, and providing different views of the same system, CAIRIS overcomes this problem. The most relevant visual models for our objectives include:

(I)   Asset models: these are based on UML class diagrams;
(II)   Task models: these are augmented UML use case models, which include additional information on task usability and misuse case risk impact;
(III) Risk models: these quickly visualize the elements contributing to risk, and can be categorized by metadata, e.g. ATT&CK tactic [8];
(IV)   Goal models: based on the KAOS [9] modelling language, these can show the system goals that exemplar system need to satisfy, and obstacles that obstruct these. These obstacles are attack trees.
(V)   DFDs: these illustrate data flows between entities, processes and data stores in exemplar systems, and trust boundaries these cross.

Figure 2 illustrates the CAIRIS asset model, and shows how security information is used to augment the UML class model elements. The actor figures indicate personas that interact with tasks that use assets in this particular context. The shading of red
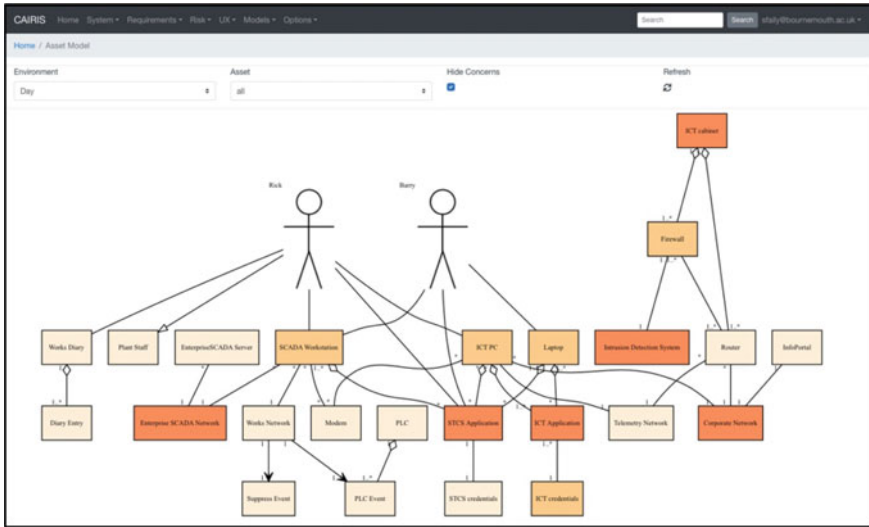
Fig. 2 CAIRIS asset model example

indicates the asset attack surface based on vulnerabilities the assets are currently exposed to; the darker the shade of red, the more exposed the attack surface is.

CAIRIS has been designed for interchangeability. CAIRIS model files are based on XML, and are intentionally both human and machine-readable. CAIRIS supports import/export is supported for a number of different file interchanges, specifications can be generated in PDF, OpenDocument Text and Word, and CAIRIS has a well-documented REST API [10].

Third, CAIRIS has been released under a permissive open-source license allowing both ECHO partners and user to consume models without any cost. CAIRIS can be freely used and extended as part of the ECHO project without any restrictions. Finally, unlike many open-source projects, CAIRIS has been heavily documented and, as part of the CAIRIS source code, and its documentation and tutorials are revised frequently as features are added and updated.

4 An Example Storyline and Use Cases Definitions

The approach used is based on end-users threats in corresponding sectors. The ECHO end-users are organizations, whose main business is in the selected sector of interest. They support the project with expert knowledge about the cybersecurity issues in their sector. Therefore, their knowledge of attacks are considered to be among the highest probable threats in the particular sector. The identified threats are modeled by use cases. The subsequent paragraphs present an example of use case concept applied to the military domain.

Storyline: Attacks on a HQ General computer network.

Story: There is an EU-supported peace-keeping mission in a zone of conflict. For the needs of the mission, a HQ is established. It is situated in a military base. A general computer network is deployed within the military base. Every officer has a workstation (mobile computer) with Ethernet connection to the network. The network topology is given in Fig. 3. An Active Directory (AD) domain controller is used to manage the user accounts in the network. There are adversary governments, who want to create a bad image for the peace-keeping mission about being unsuccessful. So, they have decided to support a local terrorist organization (noted as TORG) with information about the technology needed to steal information from the HQ. TORG, for its part, is particularly interested in headquarters plans and the personal information of military staff. There is a hostile state (noted as HST), which is interested in compromising the EU mission image. HST agents support TORG with knowhow. The Internet is provided by a DSL modem via domestic operator who provides Internet also for the needs of the national airport, located in that area. In the dormitory and the recreational places of the military base, a Wi-Fi network is deployed.

The HQ has planned to organize a convoy for fuel and other supplies to the local airport. It is scheduled for a predefined date. They have written the necessary documents and orders including a letter for the port authorities.

Storyline's context: The storyline and its associated use cases will take place in the defense domain. An attack on the network would allow sensitive data ex-filtration towards the opposing military forces (TORG in this case).

Storyline's objectives:

Through realizing this storyline and its use cases:

- The strength of the general network security will be tested
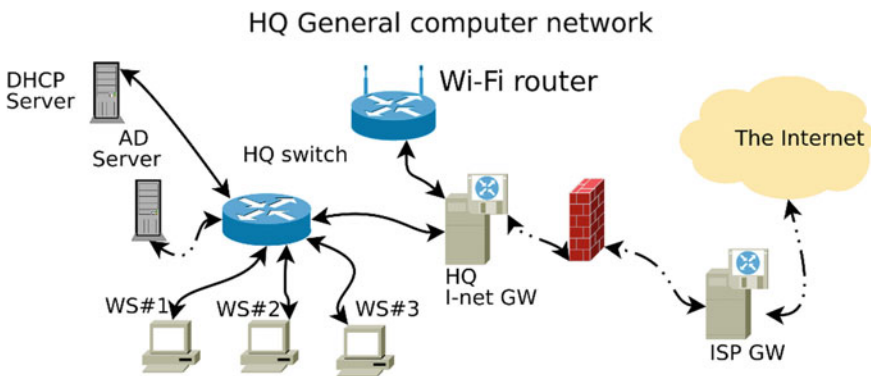- Possible future attacks towards computer networks will be prevented and/or damage will be mitigated

Fig. 3 The HQ network infrastructure

- The cyber defense staff will be educated to know about and defend against the described types of attacks.

Cyber-attack/s description/classification:

- Threat:
  - Computer viruses
  - Spyware
  - Hackers (High technological opposing forces)
  - Personnel (human factor).

- Exploited vulnerability:
  - Wi-Fi firmware vulnerabilities
  - Possible Man-in-the-middle (MITM)
  - Personnel
  - Missing data encryption
  - OS command injection
  - Unrestricted download of dangerous file types
  - Reliance on untrusted inputs in a security decision
  - Cross-site scripting and forgery
  - URL redirection to untrusted sites
  - Path traversal
  - Weak passwords
  - Software that is already infected with a virus.

Assets/processes affected: data, files, passwords, hardware (firmware).

Type of attacker: Terrorist group, supported by a hostile state.

Purpose of attack: Compromising the public confidence in the peace-keeping mission via support of terrorist group activities (military attacks) with valuable information.

The storyline "Attacks on a HQ General Computer network" includes three use cases: UC01 "Wi-Fi router firmware attack", UC02 "Man-in-the-middle attack", and UC03 "USB flash stick malware attack".

4.1   UC01 "Wi-Fi Router Firmware Attack"

Name: Wi-Fi router firmware attack.

Initial stage: Everything works as desired.

Summary: Similar to the CIA's CherryBlossom [11]/Weeping Angel [12] toolkits are provided to TORG by its hidden supporters (agents from HST). In this UC, we call it "VapourMonger". The goal of TORG is to steal information. That's why they try to hack the Wi-Fi network and obtain information via faking common sites everybody uses. These sites are for example the venerable (in this UC) social network

MasseHook. Because the attack needs to be performed remotely, TORG can use drones which they manage to land in the vicinity of the base or people carrying the equipment in backpacks.

Logical steps:

Step 1: A brute force WPA attack is started by TORG using its VapourMonger toolkit.
   System state: Everything works as desired.
Step 2: Eventually the toolkit manages to find the Wi-Fi password on the 2nd day of the attack, and proceeds with Step#3.
   System state: One of the Wi-Fi access points is under attack.
Step 3: A wi-fi admin password is found. The Wi-Fi access point attack phase#2 might find the password very quickly if a default password is held or might take some time if changed. Once found, the toolkit proceeds to the next step.
   System state: The access point control panel page is under attack. Wi-Fi administrative password is known to the toolkit.
Step 4: The VapourMonger toolkit uploads a new firmware image in the access point. The malicious firmware supports MITM attack by re-routing the connections to the Internet via outside device back to the Wi-Fi router (because it needs Internet access). So, the toolkit starts a MITM attack.
   System state: Wi-Fi password + Wi-Fi control panel page password are known to the toolkit.
Step 5: A fake (but having valid SSL certificate) clone of MasseHook social network page collects all passwords to the real MasseHook social network. Once a password is collected, the next TCP connections from the same IP/MAC address are passed-thru without intervention to allow the user to have normal service.
   System state: Wi-Fi network password and Wi-Fi access point control panel page password are known to the toolkit. The firmware is altered. A MITM attack is in progress.

Initiating actor: TORG.

Supporting actors: The HQ staff, HST government.

Final system state: On success: data leak. TORG knows the profiles of the EU mission staff, where do they live, their habits, who are their friends and family members.

Inputs: TORG knows where the military base is. (In fact, everybody knows that). Timing of the steps; MasseHook page requests from the users which are re-routed to the fake MasseHook page.

Outputs: HQ staff profile pages; friends lists and passwords for MasseHook.

4.2   UC02 "Man-in-the-Middle Attack"

Initial stage: HQ use DSL connection to a local internet provider. Computer network in the HQ operates as usual. The network administrator needs to download some drivers for a new scanner.

Summary: MITM Attack is performed with an aim to penetrate in HQ's network, take control over workstations and AD controller, and download sensitive data.

Logical steps:

Step 1: A TORG attacker succeeds to mount a modified (for MITM) router on the communication line between the HQ and the Internet provider.
   System state: Everything works as desired, but the communication passes through the attacker's MITM device.
Step 2: A man-in-the-middle attack (MITM) is performed. Communication of HQ's network with Internet is altered. URL, DNS responses are modified to point to a fake (mirror) site with modified binaries.
   System state: The Internet access is controlled by the attacker.
Step 3: The administrator downloads a driver for his new scanner. His download request is redirected to a fake web site. The driver downloaded is an executable of self-unpacking installer type, crafted by the hackers to contain the real drivers and the advanced persistent threat (APT) malware.
   System state: The driver executable is present on the administrator's workstation (WS#1).
Step 4: The malware setup is executed on the administrator's workstation (WS#1). It installs the malware and the drivers and starts the malware.
   System state: The malware is active on the administrator's workstation.
Step 5: The APT establishes a connection with its C&C server outside indicating its readiness for commands.
   System state: The malware is active on the administrator's workstation. It's connected with the C&C server.
Step 6: By external commands, sensitive data from the Administrator's workstation is ex-filtrated. Hashes of user and Administrator passwords are sent to the C&C.
   System state: The hashes and files of the administrator's folders are present on the attacker's side. The other state—no change.
Step 7: Attacker cracks the hashes and sends commands to gain administrative rights on the Administrator workstation.
   System state: A process of the APT with elevated privileges is running on the administrator's workstation. The attacker has access to all active directory resources.
Step 8: Attacker succeed to download the plan for the petrol supply convoy and its schedule.

System state: Plans for the convoy route and schedule are known to TORG. Other processes: no change. The APT is ready for commands to disrupt the communication when TORG decides to launch a real attack on the convoy.

Initiating actor: TORG.

Supporting actors: Northland intelligence team, HQ's system administrators, Local Internet Provider.

Final system state: The HQ Infrastructure is infected. The APT is connected to C&C and serving requests from TORG.

Inputs: TORG intelligence team knows details about a local Internet provider and used cable connection to HQ; HQ's system administrator knows that he needs software drivers; Timings of the steps; The convoy plan.

Outputs: Exfiltrated plans; User and Administrator passwords; HQ's network topology.

4.3 UC03 "USB Flash Stick Malware Attack"

Initial stage: The described computer network infrastructure works normally without errors.

Summary: Local citizens are hired as support staff in the HQ. One of them—a poor citizen—Namir cleans the HQ office. A TORG agent recruits him for $200 with task to leave a USB flash drive in the HQ office. The flash stick is marked as EU unclassified information and is infected with malware. The unsuspecting officer on duty inserts the flash stick in his computer, which is Windows-based. Two months later, from a military unit, which defends the port at zone of conflict two platoon commanders are killed in Brussels on their annual leave. The TORG takes the responsibility.

Logical steps:

Step 1: The cleaner leaves the USB flash drive on the desk in the HQ office.
   System state: The HQ computer network infrastructure is functioning normally.
Step 2: An hour later an officer on duty notices the "registered" USB stick and inserts it into WS#2 to see what it contains.
   System state: The HQ computer still functioning normally.
Step 3: The worm accommodated at the flash stick is activated.
   System state: The malware is active on WS#2.
Step 4: The malware (on WS#2) scans all files in the Documents folders. Then it uploads them to an attacker supplied site on the Internet.
   System state: The malware is active on the user's workstation. The exfiltrated content is present on the attacker's host.

Step 5: System state: The malware is active on the user's workstation, waiting for new files to upload.

Initiating actor: The TORG hired employee.

Supporting actors: The officer on duty; HST cyber forces (with malware).

Final system state: A malware is active on WS#2. The uploaded content contains files of the HR with names and ranks of the staff with the plan for the annual leave of the staff.

Inputs: Timing of the steps; Plans (to be stolen); TORG knows who from locals has been hired by the EU mission.

Outputs: Exfiltrated files; Infected system.

5 An Approach for Modelling the Example Storyline Within CAIRIS

This section presents an approach for turning an ECHO storyline and use cases into a CAIRIS model that validates and visualises described kill chains. Our approach surfaces assumptions that hidden in the storyline document and puts into context the threat intelligence contributing to the different risks leading to the final outcome.

The approach has three steps that identify and progressively refine the risk model elements.

Step 1: Initial Asset Identification

The first step entailed identifying explicit and implicit assets from the storyline and modelling the relationships between them. We identified an initial set of seven assets by browsing the storyline and use cases: Workstation, Domain controller, DSL modem, DHCP Server, HQ personal information, Wi-Fi access point, Base network.

In later steps, as we understood the storyline and the context a bit better, we elected additional assets being attacked, elaborated existing assets, and identified relationships between the assets. The final asset model is shown in Fig. 4.

While modelling, it was also necessary to state assumptions in our model about things that were implicit. For example, references to the Wi-Fi password implied that a single credential was used to obtain access to the base network, which also included Internet access. However, deployment of active directory assumes additional credentials are needed to access resources on workstations.

Step 2. Rationalising the Attacker

We reviewed the storyline to understand the different system roles presented in the document. In Step 1, an 'external user' appeared to be present. In Step 2, the role inadvertently downloading the malware could be an administrator and, in Step 3, both roles are present. Given the context, we decided to define'trusted' and 'untrusted' roles.
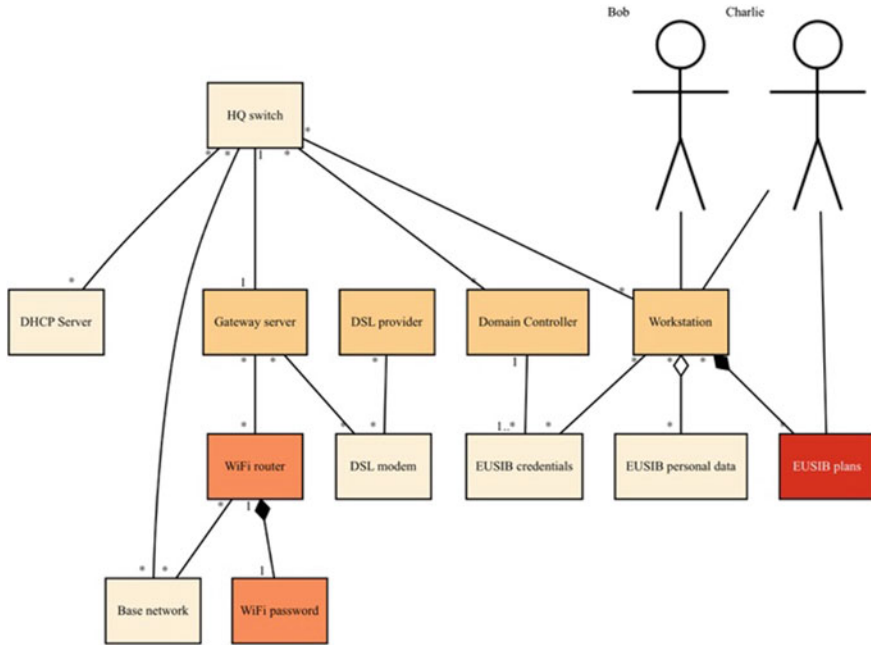
Fig. 4 Complete asset model for the storyline

It is then necessary to model attackers with the motivations and capabilities to carry out the attacks described. For this storyline, we identified three attackers:

- 'Ardit' (named after Ardit Ferizi) is an attacker with a certain amount of technical knowledge, and at least sufficient knowledge to properly operate the tools he has been given.
- 'Bob' is a network administrator.
- 'Charlie' is an officer on duty.
- 'Trudy' is a cleaner.

Of these attackers, only Ardit and Trudy have any malicious intent. The other attackers are motivated primarily by productivity.

At this stage, we also modelled tasks and skeleton personas implicit from the storyline. Two tasks were identified. One of these (Patch kit) entails a skeleton persona (Bob) following procedures for doing a monthly update of all the workstations. This task was associated with the Workstation asset.

The other task (Starts shift) describes the procedures carried out by another skeleton persona (Charlie) who starts a shift as a duty operations officer—an activity that doesn't preclude the checking of information on USB sticks. This task was associated with the Workstation asset and a newly identified HQ plans asset.

Normally, we would begin our analysis with some user research and the creation of personas. In this case, to avoid confusion, we name the attackers 'Bob' and 'Charlie'

in CAIRIS as 'Bob Attacker' and 'Charlie Attacker'. The personas associated with these tasks are named Bob and Charlie respectively, to allow these user models to evolve should user research subsequently be carried out to better understand the user goals and expectations for the personas' associated roles.

Step 3. Modelling the Kill Chains

Each use case describes a kill chain, where each use case step describes a risk. Once the risk in each step is realized, it affords a possible vulnerability or threat that provides the foundation for the risk in the next step. Modelling the risk in each step entails identifying an exploitable vulnerability and the assets being exploited, the attack carrying out the threat, the threat itself and the threatened assets, and a misuse case scenario that puts the risk in context. When we applied this approach to the storyline, we produced the Risk model in Fig. 5.

The next figures and paragraphs describe the UC01 "Wi-Fi router firmware attack" risks' modelling. The other two use cases' cyber security risks were modelled using the same approach.

In the first step, corresponding with the risk 'Crack Wi-Fi password'(Fig. 6). The threat described corresponds with a WPA brute force attack facilitated by the VaporMonger toolkit, and the vulnerability is the use of default passwords. Together, these form the basis of the 'Crack Wi-Fi password' risk, made possible by Ardit getting closer enough to the base perimeter to obtain the base Wi-Fi signal. On the basis of the brute force attack, the router password can be obtained in a short period of time.

The second step corresponds with the malicious re-route risk (Fig. 7). This is facilitated by the attacker obtaining the Wi-Fi password via the Crack Wi-Fi password risk and, using the VaporMonger toolkit, patching the firmware to poison the router's DNS cache. Like the previous risk, this attack is possible due to the use of a default password for the Wi-Fi router, which we assume would be known to the VaporMonger toolkit.

The third step corresponds with the Site redirection risk (Fig. 8). The Malicious re-route risk facilitates the Base cache poisoning threat, where the malicious re-routing ensures MasseHook network requests are redirected to a fake site that collects credentials before re-directing traffic to the authentic site. This is made possible by permissive checking of DNS tracking.

6   Conclusion

In this paper, we presented the approach used by ECHO for devising use cases for subsequent analysis and validation using CAIRIS. Our approach helped elicit sector-specific cybersecurity knowledge available within the ECHO consortium, identify the user needs within the different industry sectors.

The benefit of our approach for use case definitions was the subsequent specification of user requirements that the ECHO outputs should address. Relevant use cases
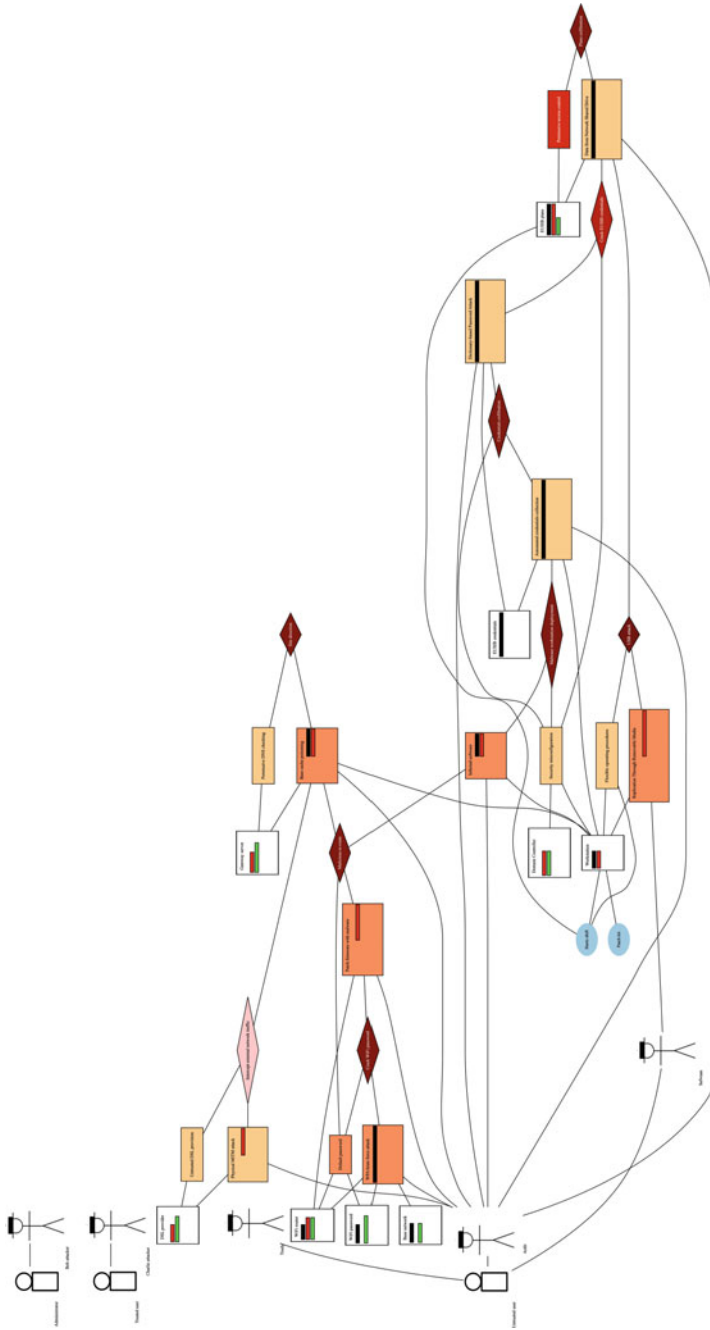
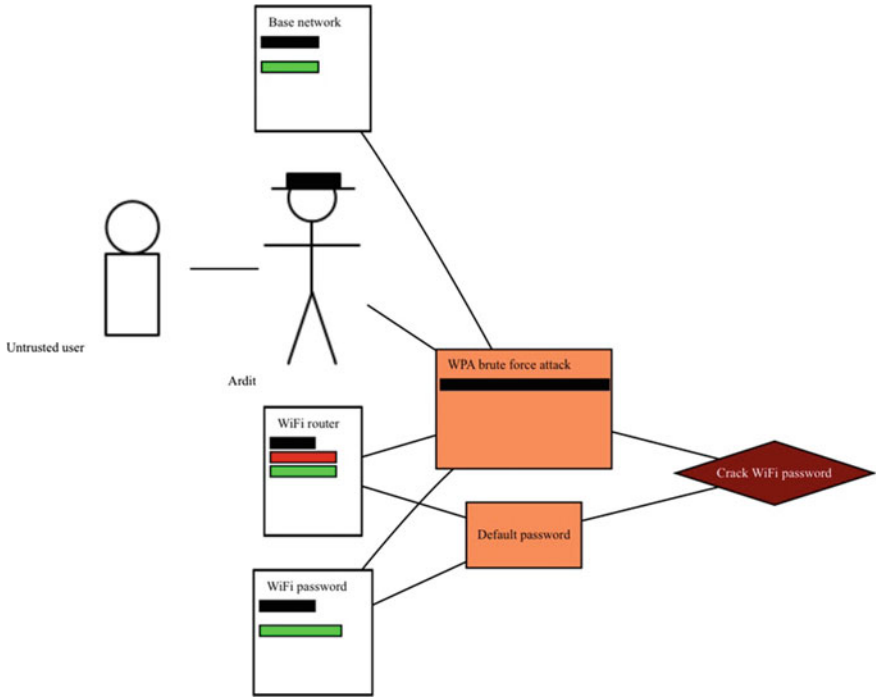**Fig. 5** Complete risk model for the storyline
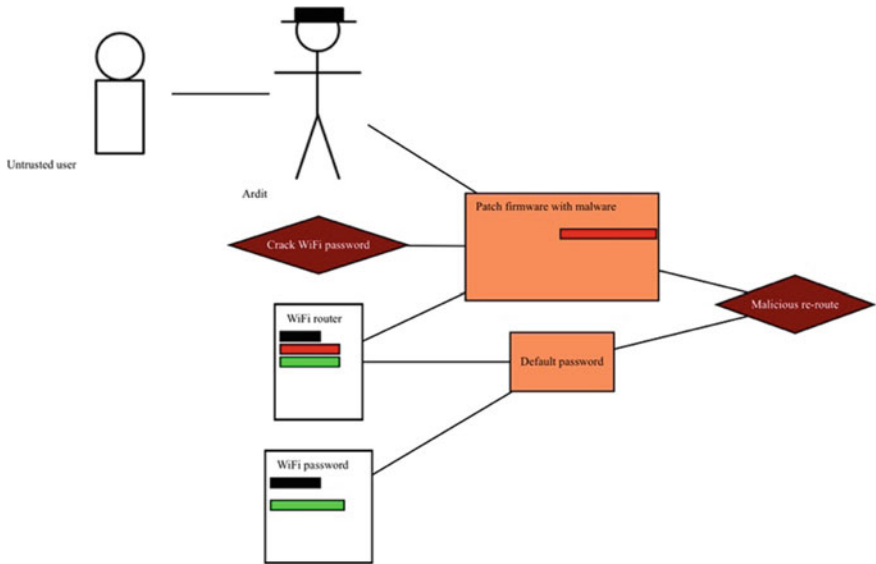
Fig. 6 Crack Wifi password risk

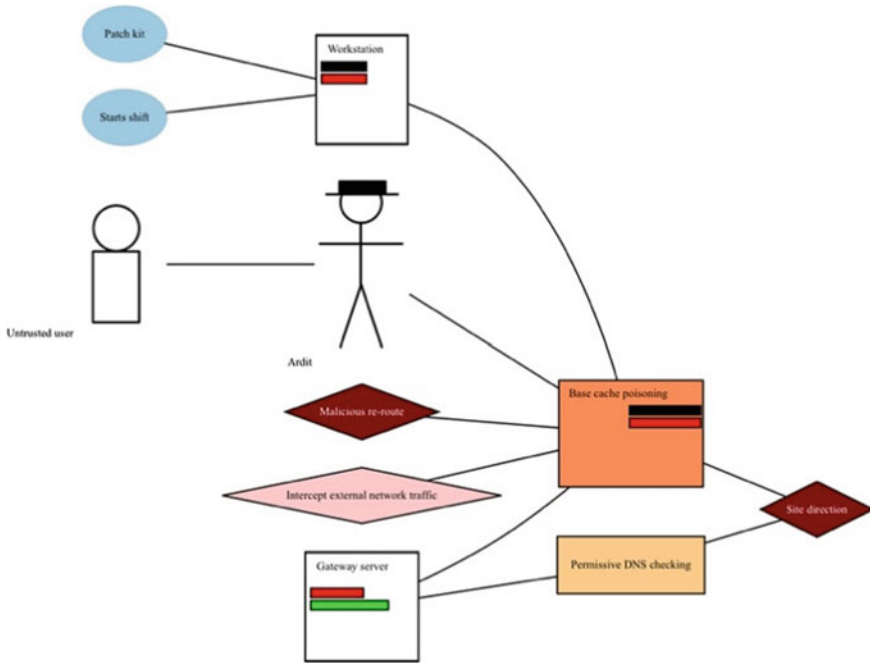**Fig. 7** Malicious re-route risk

Fig. 8 Site re-direction risk

description provides some assurance for the development of the project assets, and useful guidelines for the validation of the solutions developed.

Our approach has a wider application than only requirements elicitation. The use cases could and should be used to drive the development process. The research has proved that CAIRIS is an effective platform for enriching the context of threat models. We are currently using CAIRIS to apply security-by-design principles to the design of ECHO Early Warning System and ECHO Federated Cyber Range. Our experiences designing these systems will be the subject of future work.

References

1. European network of cybersecurity centres and competence hub for innovation and operations (ECHO). https://echonetwork.eu/. Last accessed 12 April 2021

2. Nasr E, McDermid J, Bernat G (2002) Eliciting and specifying requirements with use cases for embedded systems. In: Proceedings of the seventh IEEE international workshop on object-oriented real-time dependable systems, San Diego, CA, USA, pp 350–357

3. Faily S, Lyle J, Fléchais I, Simpson A (2015) Usability and security by design: a case study in research and development, USEC'15, San Diego, CA, USA, Copyright 2015 Internet Society, ISBN 1-891562-40-1. https://doi.org/10.14722/usec.2015.23012

4. Computer aided integration of requirements and information security (CAIRIS) web site. https://cairis.org. Last accessed 11 March 2021

5. A technical discussion on use case best practices, 06/11/03, IBM. https://www.eg.bucknell.edu/~cs475/F04-S05/useCases.pdf. Last accessed 10 March 2021

6. Faily S (2018) Designing usable and secure software with IRIS and CAIRIS, 1st edn. Springer

7. Faily S, Scandariato R, Shostack A, Sion L, Ki-Aries D (2020) Contextualisation of data flow diagrams for security analysis. In: Eades III H, Gadyatskaya O (eds) Graphical models for security, GraMSec 2020. Lecture notes in computer science, vol 12419. Springer, Cham. https://doi.org/10.1007/978-3-030-62230-5_10

8. https://attack.mitre.org/. Last accessed 11 March 2021

9. A KAOS tutorial. http://www.objectiver.com/fileadmin/download/documents/KaosTutorial.pdf. Last accessed 11 March 2021

10. CAIRIS API. https://app.swaggerhub.com/apis/failys/CAIRIS. Last accessed 11 MArch 2021

11. https://www.bleepingcomputer.com/news/security/cia-created-toolkit-for-hacking-hundreds-of-routers-models/. Last accessed 12 April 2021

12. https://en.wikipedia.org/wiki/Vault_7. Last accessed 12 April 2021

Fuzzy Game-Theoretic Modeling of a Multi-Agent Cybersecurity Management System for an Agricultural Enterprise

Valentyn V. Nekhai [ID]**, Elena Trunova** [ID]**, Iryna Bilous** [ID]**, Iryna Bohdan** [ID]**, and Mariia Voitsekhovska** [ID]

Abstract Today, in the context of growing cybercrime, corporate networks including agricultural networks require adequate protection against external cyber-attacks. A four-level model of cybersecurity management multi-agent system (MAS) of agricultural enterprise, based on the Cyber Situational Awareness paradigm, is proposed. For modeling the interaction between attack agents and protection agents, it is proposed to use the apparatus of fuzzy game theory, in particular—fuzzy game with "nature". The aim is not to find the best solution, but to find winning strategies that form a fuzzy knowledge base and obtain a fuzzy system security function. The main study purpose is to analyze and substantiate the introduction of fuzzy game models in the information security management system of agricultural enterprises corporate networks. The hypothesis of the study is to explore the possibility of using fuzzy game models, in particular games with "nature", in the cyber defense management systems of agricultural enterprises. The theoretical and methodological basis is the methodology of Game theory and the main Fuzzy logic provisions. The proposed models provide an opportunity to improve an agricultural enterprises cyber security management system. Using modern methods of corporate networks protection, the management of the agricultural enterprise will be aware of the real state of the external cyber threat. The proposed approach allows to consider fuzzy indicators to estimate the amount of losses in different scenarios, which can significantly reduce the large number of computational operations in decision-making in the process of adverse situations.

Keywords Cyber situational awareness · Information system · Control system · Fuzzy logic · Multi-agent system · Agent technologies

V. V. Nekhai · E. Trunova (✉) · I. Bilous · I. Bohdan · M. Voitsekhovska
Chernihiv Polytechnic National University, 95 Shevchenko str., Chernihiv 14035, Ukraine

1 Introduction

Information systems and decision-making methods based on knowledge are now becoming the leading factors of sustainable agricultural production. Cybersecurity management systems have become extremely complex and intensively automated. Accordingly, farmers need to know and understand the benefits of using intelligent information systems to solve various management problems. At present, agricultural enterprises are not intelligent enough, as a result of which they are not able to detect new types of attacks, to distinguish false alarms in normal situations, to interfere with the normal operation of the system.

The use of artificial intelligence methods in the field of cybersecurity is aimed at raising awareness of possible situations and automatic detection of threats, the so-called cyber situational awareness [1].

When focusing on knowledge, the automated system of the enterprise can be considered as a set of agents that act as separate intelligent systems. The operation of such systems is based on parallel and asynchronously operating "intellectual services", each of which is matched by a software agent. The interaction of agents occurs through the common bus of the enterprise. The design of services takes into account the possibility of interaction with other external systems and the possibility of embedding in other platforms. Prompt response to unforeseen external events, adaptive planning of units, control of resource use in real time—the main functions of the system [2].

The issue of automated control system stability and its security is largely determined by the reliability of the protection system, which depends primarily on the chosen strategy. It is possible to improve the management system of cybersecurity of agricultural enterprises by modeling a multi-agent management system of cyber defense of agricultural enterprises, which will result in scenarios of behavior of attack agents and defense agents in conditions of uncertainty. In this case, the actions of attack agents are defined as fuzzy linguistic statements.

2 Related Works

In most cases, the efforts to use intelligent systems in agriculture are aimed on solving certain problems of the agricultural enterprise. Evidence of a sufficient coverage of the functions of systems built on the basis of knowledge is confirmed by Talaviya et al. [3] in detailed review on artificial intelligence application in agriculture.

The use of multi-agent systems to ensure production processes is mostly specialized. For example, the control system of agricultural product is considered by Zhu et al. [4].

With the development of intelligent open systems, the issue of ensuring its security is becoming increasingly important. Nikander et al. [5] emphasize that the most common problems in farm-level cyber-physical environment are consequences of a

lack of foresight, farm staff errors or are caused by the environment. In the context of farm communication networks, human errors include topological problems occurred from poorly installed or mismanaged equipment, whereas the environment can cause problems via extreme temperatures, humidity, animal contact, or extreme weather.

The security of information systems is a serious problem today as existing threats take on new dimensions. Information security designed to protect important information assets from accidental or intentional damage. Cybersecurity is a set of procedures and systems to protect computer systems and networks from intentional and unintentional damage or threats.

In recent years, there is a tendency to expand the existing mathematical approaches to solving problems of cybersecurity management through Game theory methods implementation. Manshaei et al. [6] provided an overview and systematization of 128 works in six main categories: security of physical and MAC-level, security of self-organization of networks, intrusion detection systems, anonymity and confidentiality, security economics and cryptography. Each category identifies security issues, players and game models.

In particular, Do et al. in the paper [7] consider the existing theoretical and game approaches to the problems of cybersecurity and confidentiality. Game-theoretic models are used to combat various threats to network security and have proven their usefulness.

In the web security game, the attacker and the security defender can have too many strategies, which will make the problem computationally more complex and time consuming. Hence, fuzzy mathematics can be used to handle the situation. Ansari and Datta Gupta [8] suggested use Fuzzy game-theoretic approach based on the interactions between the security defender and the ubiquitous attacker. A non-cooperative game is formulated where the attacker tries to access and modify data assets of an organization and the defender tries to protect the resources. The suboptimal Nash equilibrium guarantees that irrespective of the attacker's strategy the defender's strategy is optimal. A Fuzzy logic and Game theory based adaptive approach for securing opportunistic networks against black hole attacks considered by Chhabra et al. [9].

Mentioned researches may serve as a foundation for further research using Fuzzy Game theory for information systems protection strategies development, including agricultural domain.

A method of choosing the best strategy for protection agents is proposed, when the situation that requires a decision allows its formalization in the form of a game with "nature" according to linguistic estimates and statistical values of probabilities of attack agents strategies.

3 Materials and Methods

The theoretical and methodological basis of the study is formed in the context of cyber-situational awareness in the multi-agent model. General scientific methods

were used, in particular analysis and synthesis, induction and deduction. Specific methods of Game theory were used for the research, in particular games with nature and the main provisions of Fuzzy logic.

3.1 Multiagent Management System of an Agricultural Enterprise

In general, the multi-agent cybersecurity management system of an agricultural enterprise is defined as follows:

$$MAS = (A, E_A, IA_A, I_A, BOS, ES, S, AC_{us}, AC_{at}), \quad (1)$$

where MAS is multiagent system; A—a set of agents; E_A—multiple operating environments of agents; IA_A—a set individual action agents; I_A—a set of interactions between agents; BOS—a set of basic organizational structures; ES—development strategy; S—a set of internal and external information flows (source of attacks); AC_{us}—a set of user actions; AC_{at}—the set of action intruders.

The greatest difficulty in the theoretical research and practical implementation of modern MAS are issues related to information security of agents and information resources they operate.

The set of agents A of the multi-agent system includes a subset of agents for cybersecurity:

$$A = ((AnA), (ConfA), (PA), (CountA), (TA), (UA), (IntA)), \quad (2)$$

which perform the corresponding functions (see Table 1).

A simplified composition of agent classes was used to model the MAS of an agricultural enterprise cybersecurity management. The model of the multi-agent system is presented as four-level (perception of the situation, understanding the situation, projection, permission) situational awareness of the internal state and the environment and an adequate response to the level of the identified threat [10].

A general view of the simulated MAS as situational awareness and communication between agents is presented in Fig. 1.

Note that when applying the MAS to each component of the agricultural enterprise is placed in accordance with the agent who can solve only some local problem. The solution to the global problem can be realized by implementing a set of agents and organizing effective interaction between them. The MAS will allow to consider the set of all possible strategies in conditions of uncertainty and choose the most acceptable (effective).

Table 1 Classes and goals of cybersecurity agents of a multi-agent management system of an agricultural enterprise

The sets of agents	Name	Functions
AnA	Analysis agent	analysis of the environment, detection of vulnerabilities, notification of the configuration agent, troubleshooting
ConfA	Configuration agent	elimination of vulnerability and confirmation of its absence
PA	Protection agent	detection of suspicious actions
CountA	Counteraction agent	elimination of the process of unauthorized actions, their sources and consequences
TA	Training agent	collecting, processing and disseminating vulnerability data for the training of other agents
UA	User agents	execution of permitted functions for the use of network resources
Int A	Attack agents	gaining access to classified information, interfering with the system, equipment failure, other damage
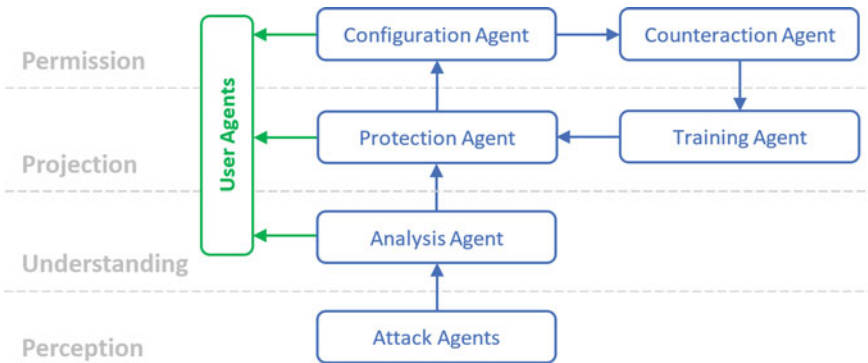
Fig. 1 MAS and agent relationships

3.2 Theoretical-Game Approach to Ensure the Stability of the Multi-Agent System

Multi-agent system of control and protection involves the use of attack agents ($Int A$) and defense agents (PA) as input data. To justify the choice of strategies to ensure the safety of agricultural enterprises, we propose to use the apparatus of game theory.

In modern control and protection systems there is a tendency to expand the existing mathematical approaches to justify the choice of strategies for ensuring the safety of agricultural enterprises, through the use of game theory. Theoretical-game approaches are used for: display the interaction of attack agents ($Int A$) and protection agents (PA), predict the actions of attackers and determine the appropriate actions of

protection; obtaining quantitative estimates of the information system level protection by predicting the actions of attack agents ($IntA$) and protection agents (PA).

If the defense agents know the matrix of the game and a finite set of attack agents' strategies, but do not know what strategy is implemented in a particular situation, the matrix game formalizes the situation of choosing defense strategies in uncertainty, each party has no information about the actions of the other party.

When the security system detects an intrusion into the network for which it does not have a ready-made appropriate strategy, it makes sense to apply a strategy based on the formation of abnormal traffic for this user, i.e. to create for the attacker to simulate the success of his attack by changing the data he views or copies, but its actions that lead to undesirable consequences for the network.

Set of possible strategies of attack agents and decisions of defense agents have a complex structure, and it is almost impossible to prove their completeness. In addition, it is considered more useful for protection agents not a discrete range of possible decision results, but a continuous one, with information about the distribution of the realization degree of these results. Traditionally used methods of decision-making in conditions of uncertainty do not provide such an opportunity. More appropriate to the essence of the task can be considered a game with "nature" in a fuzzy setting [11].

Consider the interaction of attack agents ($IntA$) and defense agents (PA) as a game with "nature", where the agent (PA) must choose a strategy in the conditions proposed by attack agents.

In particular, consider the case of linguistic evaluation (for example, large, medium, insignificant) of attack agents, which is formalized by fuzzy sets and point values of probabilities of attack agent strategies (statistical processing of sensory data).

To solve the problem, it is necessary: determine which of the possible options for determining the source data and in what combination will provide the most adequate model of the situation that requires decision-making (involvement of experts); develop the certain procedure for formulation of the input conditions of the problem that depend on the chosen option; create the appropriate libraries for each of the selected options with its own method of solving.

Elements of the payment matrix in the case of fuzzy problem statement can be set either in the form of fuzzy numbers or linguistic estimates. The transition to a fuzzy payment matrix requires the following tasks.

Formation of linguistic values set for the elements of the payment matrix.

To evaluate the selected strategies with the appropriate attack agents (payment matrix) set the term set of linguistic assessments of the level of security (behavior of security agents) $T = L(low)$, LM (below average), M (medium), MH (above medium), H (high). These estimates are formalized by fuzzy sets of triangular (trapezoidal, normal) membership functions.

$$\tilde{D} = \{\mu_L(x), \mu_{LM}(x), \mu_L(x), \mu_M(x), \mu_{MH}(x), \mu_H(x) : x \in [0, 1]\}. \quad (3)$$

When building a payment matrix for each alternative solution $a_i \in A$ and every state of nature $s_j \in S$ a set of fuzzy linguistic estimates will be matched $\tilde{D}_i = \{\mu_{ij}(x), i = \overline{1, m}, j = \overline{1, n}\}$, where $\mu_{ij}(x)$ belongs to the set $\tilde{D}$.

As a result, we obtain a payment matrix with elements in the linguistic form $\tilde{D} = \left\| \tilde{d}_{ij} \right\|$, where $\tilde{d}_{ij}$ is the value obtained by fuzzification, which is formalized by a fuzzy set with membership function $\mu_{ij}(x)$, $x \in [x_{\min}, x_{\max}]$.

In addition, it is necessary to specify the distribution of probabilities in the space of attack agents' strategies $P = \{p_j : j = \overline{1, n}\}$, $\sum_j p_j = 1$.

Then decision-making based on the linguistic evaluation of attack agents and the point values of the probabilities of attack agent strategies can be represented as follows:

$$G = \left\{ A, S, P, \tilde{D} \right\}, \tag{4}$$

where $A = \{a_i : i = \overline{1, m}\}$—protection agents set of strategies; $S = \{s_j : j = \overline{1, n}\}$—attack agents set strategies; $P = \{p_j : j = \overline{1, n}\}$—distribution of probabilities on the set of strategies of attack agents, $\sum_j p_j = 1$; $\tilde{D} = \{\mu_L(x), \mu_{LM}(x), \mu_L(x), \mu_M(x), \mu_{MH}(x), \mu_H(x) : x \in [0, 1]\} = \left\| \tilde{d}_{ij} \right\|$—the set of fuzzy estimates of the selected strategies with the corresponding attack agents (fuzzy payment matrix).

Fuzzy payment matrix of the game with "nature" according to the linguistic assessment of attack agents and point values of strategies probabilities of attack agents will be given by means of the Table 2.

Building an integrated assessment of strategies, which can be used to decide on the best strategy (lowest cost) for defense agents. Note that in this problem it is impossible to apply any of the known criteria, because they are designed to work

Table 2 Fuzzy payment matrix of the game with "nature" according to the linguistic assessment of attack agents and point values of strategies probabilities of attack agents

Attack agents' strategies probability $\tilde{p}_j$	Strategies of attack agents, s_j						Integrated strategy evaluation, γ_i
Protection agents' strategies a_i	s_1	s_2	…	s_j	…	s_n	
	p_1	p_2	…	p_j	…	p_n	
a_1	$\tilde{d}_{11}$	$\tilde{d}_{12}$	…	$\tilde{d}_{1j}$	…	$\tilde{d}_{1n}$	γ_1
a_2	$\tilde{d}_{21}$	$\tilde{d}_{22}$	…	$\tilde{d}_{2j}$	…	$\tilde{d}_{2n}$	γ_2
…	…	…	…	…	…	…	…
a_i	$\tilde{d}_{i1}$	$\tilde{d}_{i2}$	…	$\tilde{d}_{ij}$	…	$\tilde{d}_{in}$	γ_i
…	…	…	…	…	…	…	…
a_m	$\tilde{d}_{m1}$	$\tilde{d}_{m2}$	…	$\tilde{d}_{mj}$	…	$\tilde{d}_{mn}$	γ_m
Best strategy							$\min(\gamma_i)$

with numerical data, and to linguistic estimates and formalized fuzzy sets, you can apply only logical operations, including comparison of fuzzy sets [12].

Therefore, it is advisable to bring the form of fuzzy sets (membership functions) to a single option. For transformation, it's proposed the operation of replacing an arbitrary fuzzy set with a fuzzy set with an equivalent triangular (trapezoidal, normal) membership function, where the left and right boundaries and the center of gravity coincide with similar indicators of the original membership function.

To calculate the integrated assessment of the consequences of applying certain strategies, the input data for constructing an equivalent fuzzy set with a triangular membership function are: the boundaries of the carriers of fuzzy sets of the corresponding string and the coordinate of the center of gravity of the whole set of fuzzy sets $[x_{\min}, x_{\max}]$, x_{CG}—coordinate center scales.

The triangular dependence function is uniquely given by a triple:

$$\left[x_L = x_{\min}, x^*, x_H = x_{\max} \right], \tag{5}$$

where x^* is the unknown coordinate of the maximum of the membership function, which in this case is determined by the maximum of the values of the membership functions of fuzzy sets of the corresponding line (strategy) of the estimation matrix.

The x^* value can be calculated based on a known ratio to determine the coordinates of the center of gravity triangle with vertex coordinates $[x_L = x_{\min}, x^*, x_H = x_{\max}]$:

$$x_{CG} = \frac{1}{3}\left(x_L + x^* + x_H \right). \tag{6}$$

When calculating by relation (5) at arbitrary values of x_L, x_{CG}, x_H, the value of $x^* > x_H$ can be obtained, which is impossible under the conditions of determining the membership function.

Therefore, when calculating the value of x^* it is necessary to enter the appropriate restriction. Then:

$$x^* = \begin{cases} x_{CG} - (x_L + x_H), & x^* < x_H; \\ x_H, & x^* \geq x_H. \end{cases} \tag{7}$$

Another situation is possible when calculating by relation (5) for some combination of values x_L, x_{CG}, x_H it will be obtained that $x^* < x_L$, which is also impossible under the conditions of construction of membership functions. In this case, the following restriction must be introduced:

$$x^* = \begin{cases} x_{CG} - (x_L + x_H), & x^* > x_L; \\ x_L, & x^* \leq x_L. \end{cases} \tag{8}$$

Thus, the possible result of the application by the protection agent of some strategy a_k, if the exact strategy of the attack agent is not known, can be represented by an

equivalent fuzzy set:

$$\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_m \end{bmatrix} \to \tilde{D} \to \begin{bmatrix} A_1^{Tr} \\ A_2^{Tr} \\ \dots \\ A_m^{Tr} \end{bmatrix} \tag{9}$$

The final step in solving the game is choosing the best strategy from a variety of possible ones. For this purpose, it is necessary to compare the received integral estimations of strategies [13].

We formulate a fuzzy hypothesis $\tilde{H}_0$ about the possibility of considering one fuzzy estimate better than another, formalized by a fuzzy set with a triangular (trapezoidal, normal) membership function $\mu_{\tilde{H}_0}(x) = x$, $x \in [0, 1]$. Then the task of comparing fuzzy sets is to determine the degree of their correspondence (proximity) to this hypothesis. As already mentioned, the membership function can be considered as a set of material points $Y = \{y_k : k = \overline{1, K}\}$, with coordinates $\{x_k, \mu(x_k)\}, x_k \in [0, 1]$.

On this set for further calculations, it is necessary to choose a point—its generalizing characteristic. Such a point is the coordinate of the center of gravity. The integrated assessment conformity degree of the consequences of the choice of a particular strategy x_i is determined by the distance between the coordinates of the respective centers of gravity:

$$\gamma = \sqrt{\left(x_{CG}^H - x_{CG}^*\right)^2 \cdot \left(\mu\left(x_{CG}^H\right) - \mu\left(x_{CG}^*\right)\right)^2}, \tag{10}$$

where x_{CG}^H—is the coordinate of the center of gravity of the membership function of the fuzzy hypothesis; $\mu\left(x_{CG}^H\right)$—its value at this point; x_{CG}^*—is the coordinate of the center of gravity of the membership function of fuzzy integral estimation of an arbitrary strategy; $\mu\left(x_{CG}^*\right)$—the corresponding value of the membership function.

Obviously, the smaller the value of this indicator, the higher the correspondence of a particular estimate to the hypothesis $\tilde{H}_0$, the better the corresponding strategy will be.

4 Results

In the process of testing the model, a knowledge base of intrusion detection systems was created, which contains attack scenarios (attack agents). The strategies of security agents based on them differ from each other both in objects and parameters for monitoring and in time. For example, consider three scenarios (Table 3).

For reasons of convenience of calculations, the values of the elements of the evaluation matrix were reduced to the range [0,1].

Table 3 Fuzzy representation of the game with "nature" between attack agents and defense agents in 3 scenarios of attack agents

Attack agents' strategies probability	Strategies of attack agents, s_j			
Protection agents' strategies a_i	s_1	s_2	s_3	s_4
	$p_1 = 0.2$	$p_2 = 0.4$	$p_3 = 0.1$	$p_4 = 0.3$
a_1	$(L; LM)$	$(LM; M)$	(M)	$(M; MH)$
a_2	$(M; MH)$	(M)	$(L; LM)$	$(LM; M)$
a_3	(MH)	(LM)	(H)	$(L; LM)$

5 Discussion

As follows from Table 4, the best strategy should be considered a_3, which will be offered to the protection agent.

The proposed model of decision-making based on a fuzzy game with "nature" allows you to choose the best option in situations of uncertainty, when assessments of the consequences of the decision can be presented only in the form of vague, linguistic statements. Since none of the known criteria can be applied in this problem, as they are designed to work with numerical data, and only logical operations can be applied to linguistic estimates and formalized fuzzy sets, it is proposed to form an integral assessment of the consequences of the chosen solution membership function, which stores the entire amount of source information and simplifies the procedure of comparing the obtained fuzzy, integrated estimates.

Ten cyberattack scenarios and strategies for their prevention have been developed, each implemented up to 100 times. The results presented on Fig. 2 show the average values of the detection and prevention percentage of losses. The actions of the defense agents were chosen according to the result of solving a fuzzy game with the attacking agent.

This approach allows to consider fuzzy indicators to estimate the number of losses in different scenarios, which can significantly reduce the large number of

Table 4 Fuzzy representation of the game with "nature" between attack agents and defense agents in 3 scenarios of attack agents in the coordinates of the centers of gravity

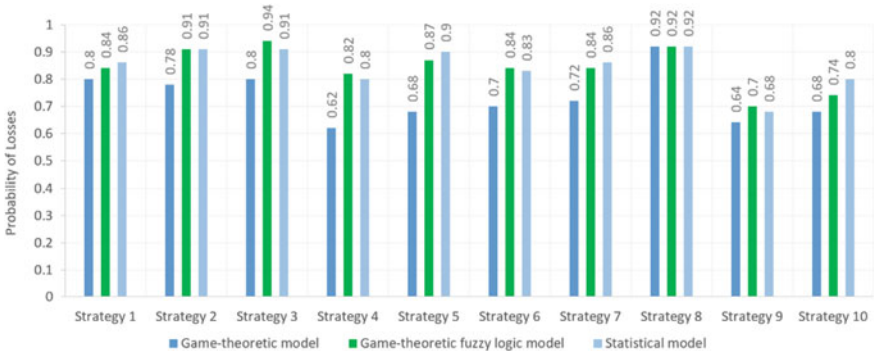| Attack agents' strategies probability p_j	Strategies of attack agents, s_j				Integrated (weighted) assessment of strategy, γ_i
Defense agents' strategy value $CG(a_i)$	s_1	s_2	s_3	s_4	
	$p_1 = 0.2$	$p_2 = 0.4$	$p_3 = 0.1$	$p_4 = 0.3$	
$CG(a_1)$	0.263	0.342	0.5	0.657	0.437
$CG(a_2)$	0.657	0.5	0.241	0.321	0.451
$CG(a_3)$	0.788	0.263	0.912	0.211	0.417
Best strategy (Strategic choice), $\min(\gamma_i)$					0.417

Fig. 2 Accuracy of detection of the proposed approach

computational operations in decision-making in the process of adverse situations. The probability of losses in the implementation of different scenarios according to the proposed methods is slightly different from the results of a discrete game with "nature" and the usual statistical assessment. Implementation time will be much shorter due to the use of expert judgment.

6   Conclusions and Further Research

This paper proposes a solution to the problem of agricultural enterprises management information systems protection by modeling decision-making systems based on game theory. A multi-agent cyber security management system has been developed, which is an integral part of the agricultural enterprise multi-agent security system. The cybersecurity system contains attack agents and defense agents, their place and functions in the system are determined on the basis of the cyber situation awareness concept. To ensure their effective functioning, a fuzzy game-theoretic approach is proposed, in particular fuzzy games with "nature", which ensures the stability of the multi-agent system.

In this approach, the evaluation of defense agent strategies is implemented on the basis of attack agent's linguistic evaluation, which is formalized by fuzzy sets and point values of attack agent strategies probabilities, which allows to choose winning strategies. They form the knowledge base on the basis of which it is possible to determine the strategy of agents-protection of an agricultural enterprise information management system.

For further research, a promising area will be the introduction of automated assessment of possible scenarios, based on a single knowledge base of a multi-agent system. This will ensure the functioning of security agents while preventing external cyberattacks. It would also be useful to consider combinations of linguistic estimates of

attack agent strategies and linguistic values of probabilities of attack agent strategies in different scenarios.

References

1. Pappaterra MJ, Flammini F (2019) A review of intelligent cybersecurity with Bayesian networks. In: 2019 IEEE international conference on systems, man and cybernetics (SMC). Published. https://doi.org/10.1109/smc.2019.8913864
2. Skobelev PO, Simonova EV, Budaev DV, Voshhuk GYu, Laryukhin VB (2018) Oblachnaya intellektualnaya sistema SMART FARMING dlya upravleniya tochnym zemledeliem (Cloud-based smart system SMART FARMING for precision farming management). In: Informaczionnye tekhnologii v upravlenii (ITU-2018). SanktPeterburg
3. Talaviya T, Shah D, Patel N, Yagnik H, Shah M (2020) Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. Artif Intell Agri 4:58–73. https://doi.org/10.1016/j.aiia.2020.04.002
4. Zhu Y, Li S, Liu S, Yue E (2011) Design of agent-based agricultural product quality control system. Comput Comput Technol Agri IV:476–486. https://doi.org/10.1007/978-3-642-18333-1_56
5. Nikander J, Manninen O, Laajalahti M (2020) Requirements for cybersecurity in agricultural communication networks. Comput Electron Agri 179:105776. https://doi.org/10.1016/j.compag.2020.105776
6. Manshaei MH, Zhu Q, Alpcan T, Bacşar T, Hubaux JP (2013) Game theory meets network security and privacy. ACM Comput Surv 45(3):–39. https://doi.org/10.1145/2480741.2480742
7. Do CT, Tran NH, Hong C, Kamhoua CA, Kwiat KA, Blasch E, Ren S, Pissinou N, Iyengar SS (2017) Game theory for cyber security and privacy. ACM Comput Surv 50(2):1–37. https://doi.org/10.1145/3057268
8. Ansari AQ, Datta Gupta K (2017) Fuzzy game theory for web security. Appl Soft Comput Web 183–189. https://doi.org/10.1007/978-981-10-7098-3_11
9. Chhabra A, Vashishth V, Sharma DK (2017) A fuzzy logic and game theory based adaptive approach for securing opportunistic networks against black hole attacks. Int J Commun Syst 31(4):e3487. https://doi.org/10.1002/dac.3487
10. Nekhai VV, Dorosh M (2020) Using the cyber situational awareness concept for protection of agricultural enterprise management information systems. Inf Secur Int J 46(2):168–181. https://doi.org/10.11610/isij.4612
11. Seraya OV, Katkova TN (2012) Zadacha teorii igr s nechetkoy platezhnoy matritsey (The task of the theory of games with a fuzzy payment matrix). Math Mach Syst 3:29–36
12. Piegat A (2010) Fuzzy modeling and control. Beltz Verlag
13. Chernov VG (2018) Comparison of fuzzy numbers on the basis of construction of the linear relation of the order. Dynam Difficult Syst XXI Century 2:81–87
14. NATO SPS Project CyRADARS (Cyber Rapid Analysis for Defense Awareness of Realtime Situation). https://www.cyradars.net. Last Accessed on 25 March 2021

Mathematical Modeling and Simulation of Special Purpose Equipment Samples

Issue of Airplane Modification Mass Variations Based on Their Structural Modelling

Viktor Riabkov ⓘ**, Ruslan Tsukanov** ⓘ**, Oleksandr Dveirin** ⓘ**,
Liudmyla Kapitanova** ⓘ**, and Maryna Kyrylenko** ⓘ

Abstract Transport category airplane modification creation process obtained wide distribution among all airplane-design companies all over the world, including Ukraine. It is the main development direction for the airplanes of this category, because new achievements are introduced the most quickly and the time of new airplanes implementation decreases considerably by means of modification changes. But, according to the analysis, some problems arise, among which it's necessary to mention the problem of the modification mass increase up to 27–30% in comparison with the base airplane. This leads to considerable worsening takeoff and landing characteristics of modifications. To minimize this mass increase, a new method is proposed for mass structurization to starting mass and takeoff mass, with estimation of their increase reasons, resulting in their difference minimization. Thus, the mass at a final stage of the modification designing is indicated as starting mass, but the modification mass at the certification stage is indicated as takeoff mass. The proposed method for the airplane mass variation estimation also allows to solve the other problems, related to takeoff and landing and economical characteristics of the modification.

Keywords Airplane modifications · Starting and takeoff masses · Takeoff run · Required runway length

V. Riabkov · R. Tsukanov (✉) · L. Kapitanova · M. Kyrylenko
National Aerospace University «KhAI», Kharkiv, Ukraine
e-mail: ruslan_xai@ukr.net

L. Kapitanova
e-mail: zzzmila888@gmail.com

O. Dveirin
Antonov Company, Kyiv, Ukraine

1 Introduction

Transport category airplane modification creation have been distributed in the world airplane designing and became the main direction of this airplane type development [1].

This process is also widely utilized in the Antonov Company. Since seventieth of the last century, on the base of such worldwide known airplanes, as An-26, An-12 and An-124, dozens of their modifications have been created, the most adequate requirements to that time [2].

American companies Boeing and Lockheed, and the Ukrainian Antonov Company are the most experienced in military transport designing [3].

These airplane-design companies are leading in development of methods and routines for creation of high-performance modifications of transport category airplanes. They achieved perfection in the implementation of modification changes in the airplanes under consideration.

But even for these companies, solving the mass problem for modifications having higher productivity remains actual problem, as it is integral index of their efficiency.

2 Analysis of References Data and Problem Statement

The main documents, on which the airplane design process is based, are aviation regulations Part-25 [4] and Part-33 [5], where all the requirements to new developed transport category airplanes are specified. But they do not contain conditions for the case, when some part of modification parameters is taken from the base model, but the other part should be changed taking into account modern requirements.

Requirements, which should be met during aviation engines creation, are stated in the normative publication [5]. But the conditions of the main engine integration in the system «airplane-engine» are not stipulated in this document, which does not allow to optimize the «modification-engine» system.

Operational and economical conditions of transport category airplane operation are clearly defined in the requirements of ICAO [6] and IATA [7] international organizations. But mass parameter as integral index of their efficiency is absent in these documents.

The problem of airplane mass formation in combination with other key parameters, such as thrust-to-weight ratio, speed performance etc. is considered in the survey [8].

But it is again for new developed airplanes, when limitations to some of these parameters are designated not so strictly.

Among Ukrainian publications, it is necessary to note the work [9], where the authors gave dependencies for mass determination, but only for newly designed regional airplanes, without considering other airplanes of types.

In publication [10], the problem of mass minimization in the process of wing panels manufacturing is considered, as a particular case of modification changes in wings of already designed airplanes, but generalization for other airplanes of types has not considered.

In publication [11], mass variation, as a result of tightened strength requirements and inevitable reinforcement of particular components, is indirectly considered, but variation of other airplane parameters has not mentioned.

In publication [12], particular variation of specific component is shown, its operation improvement is also shown, but impact of its variation onto airplane as a whole and its mass is not shown.

Thorough research as for estimation of structural changes in airplane is given in the publication [13]. The research shows necessity of the modification changes, but their influence onto mass variation has not considered.

Particular changes of the modification mass are also considered in the publication [14], which also does not give an answer on the following question: in what manner can their influence on the modification mass variation be taken into account and structured. In publication, formation of the most common airplane parameters are considered, when there are no strict limitations, which exist in the modification development process: what should be left without any changes, and which parameters should be improved, and how this influences the modification mass, as the most integral index of its efficiency.

The important conclusions follow from the given analysis of references and publication [15]:

- When making transport category airplane modifications, inevitable growth of their starting and takeoff masses takes place at all stages of their lifecycle;
- For the time being, there is no structural approach to evaluate reasons of the modification mass growth from starting mass value to the mass value with which the modification is certified;
- This leads to decrease in specific indexes of efficiency per mass unit.

3   Goal and Tasks of Survey

The goal of the publication is investigation of the transport category airplane mass variation in the processes of its designing, prototype testing, series production, and the mass evaluation at the airplane modification certification.

To achieve this goal, it is necessary:

- To perform analysis of modification changes features on the examples of military transports;
- To develop method and models for the modification mass structurization into starting mass and takeoff one;
- To simulate the interrelation between starting mass and airplane key parameters;
- To determine the reasons of airplane modification takeoff mass growth;

- To perform investigation of takeoff mass variation influence on the takeoff and landing parameters of certified modifications.

4 Features of Modification Mass Variation (on Example of Military Transports

During development of military transports, the problem of their mass formation taking into account their purpose appears:

- Presence of cargo cabin of big sizes with cargo doors of corresponding dimensions;
- Presence of strong cargo deck (floor), equipped with tie-down points and capable to withstand heavy loadings (both concentrated, and distributed) caused by transported cargo;
- Presence of cargo door structural elements, performing the function of loading devices (ramps, loading bridges etc.) and being simultaneously airplane structural components;
- Wide range of carrying capacity and possibility of wide variation of mass of the transported cargo and uplifted fuel ratio;
- Wide range of center-of-gravity position, reaching 18–20% MAC (Mean Aerodynamic Chord);
- Possibility of cargo and troops paradropping with the mass of paradropped monolithic cargo reaching the airplane carrying capacity;
- The main properties of military transport, finally, are derivative from masses of its components, and, consequently, technical and economic efficiency is completely determined by the structure and values of component masses, which reflect all other properties;
- Creation of complex technical system as military transport is rather long-term cycle. The time from the development beginning till «peak» of operation makes tens of years together with manufacturing and operation cycles. As a result, prediction of all characteristics and, consequently, masses for this period is required;
- Within this period definite changes to both initial conditions and content of the military transport project take place, which influences its output performance and, first of all, mass of components.

These problems become more acute by the fact, that the more airplane size is, the more considerably the «square—cube» law manifests; as mass ratio of the empty airplane, without making special actions, has the expressed tendency to substantial increase.

The problem of mass performance is partially solved by means of base airplane modification designing scientific-and-technological advance over a period of its application taking into account. But even in this case, there are some objective and subjective reasons for mass considerable variation before its certification.

5 Method and Models of Modification Mass Structurization into Starting and Takeoff Masses

As the modification mass variation takes place during all its lifecycle, it requires introducing new terms: starting (initial) mass and takeoff mass (with which the modification is certified).

Starting mass means the mass of base model and the mass of modification at the final stage of designing.

It is usually estimated in the terms of airplane mass balance:

$$\overline{m}_s + \overline{m}_{PP} + \overline{m}_f + \overline{m}_{eq} + \overline{m}_{pl} = 1 \quad \text{or} \quad \sum \overline{m}_i = 1, \tag{1}$$

where $\left(\overline{m}_s;\ \overline{m}_{PP};\ \overline{m}_f;\ \overline{m}_{eq};\ \overline{m}_{pl}\right) \rightarrow m_i/m_0$—are the mass ratios of structure, power plant, fuel, equipment and systems, transported cargo, included in the existence equation;

m_0—is the starting mass of base airplane or the takeoff mass of modification.

Takeoff mass means the mass of modification at the certification stage that is accounting all improvements, arisen in the process of manufacturing and flight tests.

6 Modeling of Starting Mass Interrelation with Modification Key Parameters

To analyze the reasons of the modification starting mass variation, the equation of its weight balance is presented in the form of the system of the mass components

$$\overline{m}\ (S, PP, F, EQ, PL) = f\ (S, PP, F, EQ, PL(m_0;\ \in P_i)), \tag{2}$$

where $\in P_i$—is a set of determinative parameters.

It is a mass mathematical model of base airplane and its modifications.

In this model, starting mass (m_0) plays a role of the main «controlling» parameter or «dimension» parameter. The set of other parameters P_i describing the project according to their specific physical sense can be subdivided into four groups or bases of data included in the project; each of them has definite semantic information:

- Specific requirements to the project. This group of parameters can be defined as «REQUIREMENTS» data base and denoted as «R». It defines designing boundary conditions, which implied by specified technical requirements;
- Group of parameters, determining physical possibilities of achieving the required properties and qualities, necessary to solve the airplane creation problem. All physical specific indexes, achieved for the considered level of technical state of science and industry (material specific strength, aero-gas-dynamic factors,

specific consumptions, and other specific masses of properties and qualities) can also be included. This group of parameters can be defined as «LEVEL» data base and denoted as «L»;

- Group of parameters describing «dimensionless» layout-geometrical model and airplane structure that is its shape. This group, characterizing relative airplane volumetric-geometrical data, can be defined as «IMAGE» group of parameters and denoted as «I»;

- Small group of parameters, which is a characteristic of «I» group «saturation» with specific mass expanses to implement the parameters, included in the groups «R» and «L». This group characterizes mass and energy density of the airplane (including specific loading on the lifting surface, starting thrust-to-weight ratio, load per mid-section unit area) can be called «DENSITY» and denoted as «D».

Experience of transport category airplane modification creation shows, that mass plays special role in their technical and economic efficiency, which predetermines these groups relation with technical and economic efficiency of all decisions. The condition is met within each of the groups

$$\frac{\in P_i(R, \ L, \ I, \ D)}{m_0} = \frac{TER}{Starting \ mass}, \tag{3}$$

Thus, TER (Technical and Economic Requirements) are natural measurer of expenses necessary to achieve the specified set of properties and qualities of airplane.

But, no matter how good the airplane was arranged and optimized at the initial stage of designing, dialectics of the creation process inevitably leads to variation of both numerator and denominator of this ratio. Hence it's clear, that the project parameters «drift» from the optimal ones and, as a result, its efficiency worsening.

Thus it's necessary to take into account, which solution of the condition (3) can exist only, when specified requirements correspond to the available possibilities of their implementation.

Elicitation of this fact that is modification required characteristics comparison with available technical possibilities and reduction of the set of parameters $\in P_i$ in a single specific image, is a compromise problem of the new modification development.

Hence, there is a necessity in the method for solving equation of existence (1) in respect to the modification analysis problems, which allows to retrace the relation of $\in P_i$ with mass characteristics variation (see Fig. 1).

When solving the system of weight balance, we use the approach of its conditional subdivision into required and available masses with specific combination of R, L, I, D parameters, which describe the airplane version being designed:

$$\eta_{req} = m_{pl}(m_0, \ R, \ L, \ I, \ D) + m_{eq\,1}(m_0, \ R, \ L, \ I, \ D) \tag{4}$$

$$\eta_a = 1 - \big[\overline{m}_s(m_0; \ R; \ L; \ I; \ D) + \overline{m}_{pp}(m_0; R; L; I; D)$$
$$+ \overline{m}_f(m_0; \ R; \ L; \ I; \ D) + \overline{m}_{eq2}(m_0; \ R; \ L; \ I; \ D)\big], \tag{5}$$
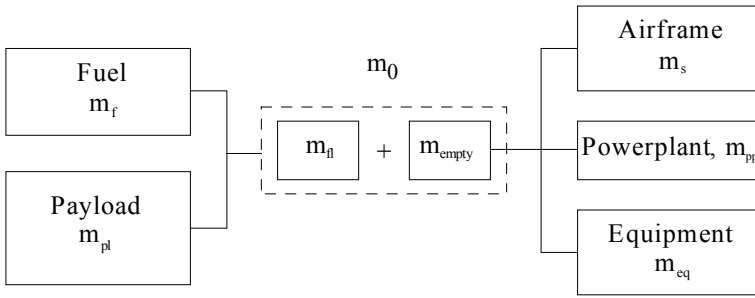
Fig. 1 Structure of modification mass components

which can be characterized for each specific combination of R, L, I, D parameters describing the airplane as:

η_{req}—is the required airplane useful-to-takeoff load ratio under specified requirements and set of functions of systems and equipment;

η_a—is the available airplane useful-to-takeoff load ratio under specific scientific and technical level of industry.

In other words, the required mass (m_{req}) determines that part of starting mass (m_0), which should be given for constant values of R, L, I, D.

But the available useful-to-takeoff load ratio shows, which part of mass (m_0) can be given to meet the technical and economical requirements under the existing conditions of their implementation.

The equation of existence can be solved, that is the airplane can be created for the specified requirements, only in the case when the conditional required mass is equal to the conditional available mass. It is also clear, that change in the R, L, I, D parameters combination that is their variation in the process of airplane modification with greater productivity, shifts the solution of the equation of existence to a new point, which corresponds to a new starting mass (m_0) of modification (see Fig. 2).

Starting mass increment is determined by the expression:

$$\eta_{req} = \frac{\sum m_{req}(m_0; \in P)}{m_0}$$

$$\eta_a = 1 - \frac{\sum m_a(m_0; \in P)}{m_0}$$
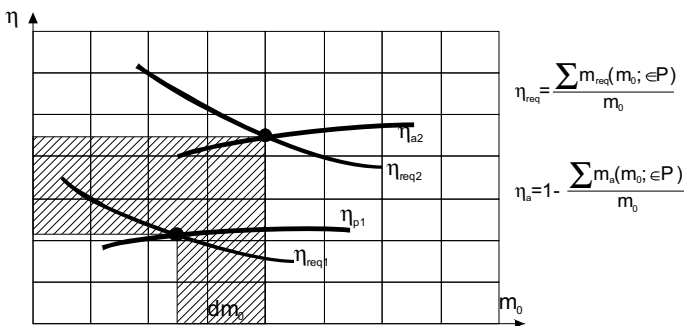
Fig. 2 Increment of starting mass (dm_0) in the process of modification creation
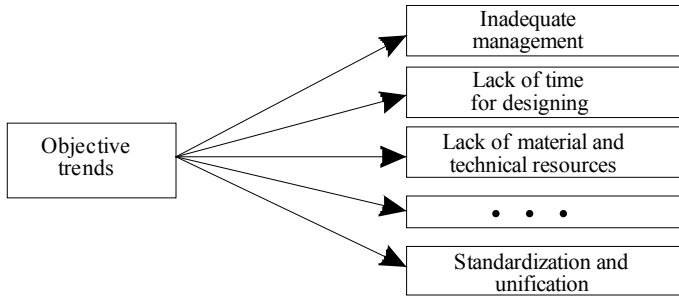
Fig. 3 Objective reasons of starting mass variation

$$dm_0 = \left(\sum \frac{\partial m_{lf}}{\partial P_i} \cdot dP_i - \sum \frac{\partial m_{emp}}{\partial P_i} \right) / \left(\frac{\partial m_{emp}}{\partial m_0} - \frac{\partial m_{fl}}{\partial m_0} \right), \tag{6}$$

where $dm = dm_{fl} - dm_{emp}$.

When estimating the starting mass increment on the base of expression (6), it is also necessary to take into account its inevitable growth, caused by some subjective and objective reasons (see Fig. 3).

Let us make subdivision into two subgroups within widened interpretation of equation of existence for the modification. All the fixed mass parameters attribute to the first group, variable mass parameters caused by the modification variations attribute to the second one. In this manner the proposed method becomes universal, applicable for modification changes at all stages, even in the process of finished airplane operation.

Systems of Eqs. (1), (4), (5), (6) essentially are a mathematical model of the modifications, describing relation between groups of parameters (R, L, I, D), the starting mass and its changes in the designing process. It is necessary to mention, that such systems are substantially concrete for the specific airplanes and even for the specific designers companies. These systems give possibility to retrace common laws, for example, expressed shift of mass ratio minimum to the side of greater starting mass with the modification productivity growth.

7 Airplane Modification Takeoff Mass Increase Reasons

Absence of necessary resources. Solving of many problems during modification development, presence of sufficient material and technical resources to the beginning of new item production, their insufficiency dictate certainty of design-manufacturing solutions, that leads to the takeoff mass increase.

Standardization and unification increase necessity. High level of standardization and unification of the modifications increase their technical and economical indexes, but inevitably results in the modification takeoff mass growth.

Manufacturing deviations increasingly depend on lack of resources and corporate culture. Latent foul-up, manufacturing technology violation, part structure change or material replacement contribute to the airplane takeoff (empty) weight growth. Examples of latent foul-up can also be: increased thicknesses of cast part walls; absence of lightening holes; uncut or not rounded cones in sheet and pressed parts; tolerances exceeding during mechanical processing; fasteners replacement etc. They lead to the modification takeoff mass growth.

Static and fatigue tests of new modification are obligatory during its designing. Results of the static and fatigue tests (later in time) of structure strength discovers possible mistakes in calculations and designing, imperfectness of methodology, assumed loading schemes of airframe components. Elimination of insufficient static and fatigue strength of structural components is performed commonly due to their reinforcement, which also leads to the certified modification takeoff mass increase.

Flight tests. Mass reserve can be required according to the results of factory, state, certification, operational and other types of tests. Mass reserve can be required to perform one or several characteristics: payload, flight, takeoff and landing performance, stated reliability indexes and other contracted data.

Mass reserve absence to the moment of flight test beginning can result in worsening some operational characteristics, for example, flight speed, engine optimal power settings, kilometric fuel consumption, cruising flight altitude etc.

Application of the method for modification mass structurization into starting and takeoff ones, and the models for their difference minimization is confirmed by the quantitative variation of An-26, F-27 and B-727 airplanes modification masses, shown in Figs. 4 and 5.

8 Research of Takeoff Masses' Variation Influence on Takeoff and Landing Parameters of Certified Modifications

Objectivity of the specified reasons influence on takeoff mass variation when designing modifications with increased flight and hourly productivity is also justified by the data shown in Fig. 6.

As it is clear from Fig. 6, takeoff mass increase trend for the heavy airplane category is the most pronounced from zero till 38%.

So considerable increase in modification takeoff masses firstly results in decrease of their takeoff and landing characteristics, such as takeoff run and landing run, and also required runway length.

It is typical for all weights of transport category airplanes (Table 1).

As it is clear from the Table 1, the takeoff run increases almost by 500 m, and taking into account possible aborted takeoff, required runway length can increase by 1000 m.
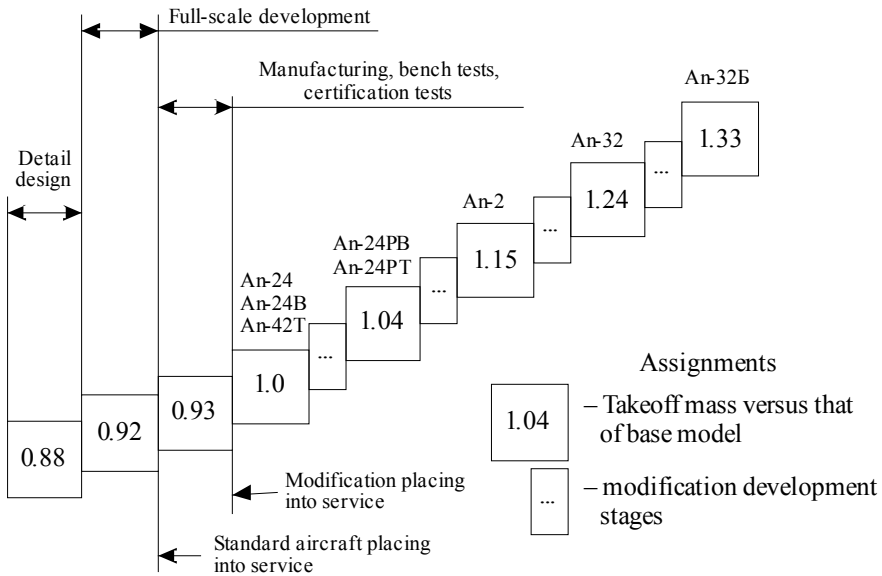
Fig. 4 Variation of starting and takeoff masses of An-26 airplane and its modifications according to stages of their creation and operation

Objective and subjective increase of starting and takeoff masses influences negatively on the estimation of airplane specific efficiency pursuant to the most important indexes.

Thus, specific workability per unit of takeoff mass is

$$\overline{A} = \frac{A}{m_0} = \frac{m_{pl}}{m_0} = \overline{m}_{pl}L, \tag{7}$$

but specific fuel efficiency is determined by the expression

$$\bar{t}_e = \frac{m_f}{m_0\overline{A}} = \frac{\overline{m}_f}{m_0} \cdot \frac{1}{L}, \tag{8}$$

where L—is the flight range.

Index «b» relates to the base airplane, and index «m» is to modifications. It is clear, that if to consider the relations $\frac{\overline{A}_m}{\overline{A}_b}$ and $\frac{\bar{t}_{em}}{\bar{t}_{eb}}$, they characterize improvement or worsening of technical and economic indexes at variation of the starting (or takeoff) masses of modification.

Some approaches [8, 9] are proposed to solve these problems.

These programs include the most important stages of this airplane lifecycle: consideration of various versions of solutions at the designing stage and strict monitoring over the solutions adoption at the manufacturing stage.
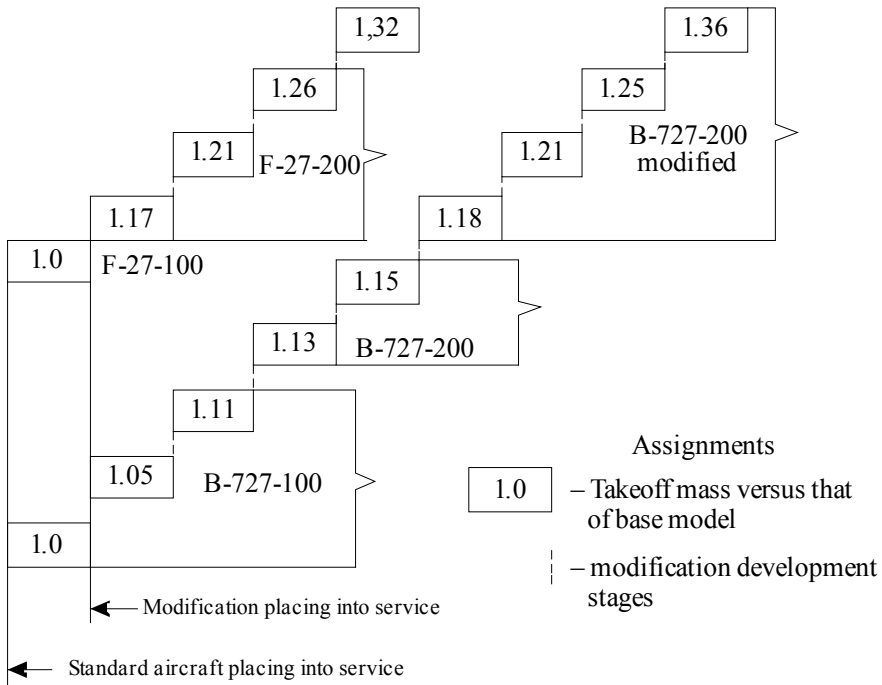
Fig. 5 Variation of starting and takeoff masses of F-27 (Holland), B-727 (USA) airplanes and their modifications relative to base model according to stages of their development and operation

Very effective and modern direction of «keeping» the increase of starting and takeoff masses of more productive modifications is the path of deep modification changes. They are performed in wing geometry (to increase its lift-to-drag ratio) and in power plant (to expand thrust-airspeed characteristics of main engines). Harmonization of these changes already at the stage of more load-lifting modification designing provides minimal increase of starting and takeoff masses [1, 3].

9 Discussion of Research Results of Modification Mass Formation Based on Structural Modelling

A new method for structurization of transport category airplane modification mass into starting and takeoff is proposed. The masses are limited by low value by the base airplane mass, and by the upper value—by the required competitiveness parameters of newly developed airplane (para. 5).

The developed models (2)–(4) allow to estimate the airplane modification starting mass variation and discover of its growth reasons (see Fig. 3) at its designing stage.
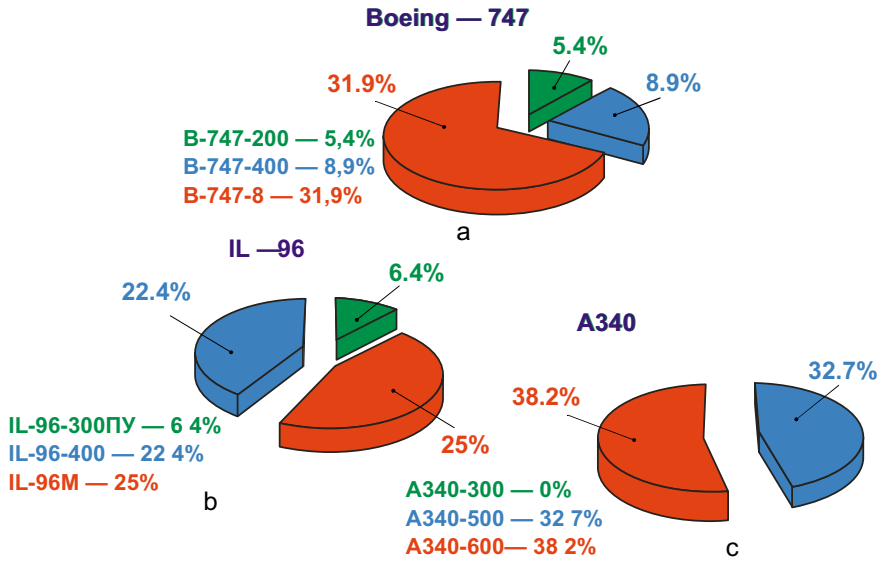
Fig. 6 Variation of modification takeoff masses of certified heavy transport category airplanes: **a**—Boeing-747 airplane modifications, **b**—IL-96 airplane modifications, **c**—A340 airplane modifications

Table 1 Influence of takeoff masses of A330 and A340 airplane modifications on their takeoff and landing characteristics

Airplane model and its modification	Takeoff mass, kg	Engine model	Takeoff and landing parameters	
			Takeoff run, m	Landing run, m
A330-200	243,450	GECF-6-80E	2256	1509
A330-300	243,450	GECF-6-80E	2652	1745
A340-200	275,000	CFM56-5C	2652	1790
A340-300	276,500	CFM56-5C	3048	1864
A340-500	368,000	RRTzent500	3186	1960
A340-600	368,000	RRTzent500	3186	2012

The reasons have been discovered, and expressions (5)–(8) have been given, which allow evaluation of the takeoff mass increase at all time periods till the modification certification (see Fig. 3).

Quantitative estimation of inevitable increase of starting and takeoff masses can also be given on the examples of such widespread transport airplanes, as An-26 (Ukraine), B-727 (USA) and F-27 (Holland).

The modification masses structurization into starting and takeoff allows estimating quantitatively their influence on:

- Takeoff and landing characteristics of such airplane modifications as A330 and A340 (see Table 1).
- Important competitiveness indexes—the specific productivity per unit of takeoff mass and the specific fuel efficiency.

It is necessary to stress, that the given models of starting and takeoff masses increase evaluation, as opposite to the models used nowadays, allow determination of takeoff mass of the certified modification with the stated parameters of efficiency at the designing stage [1].

10   Conclusions

1. By the analysis of already created modifications of military transports, it was established that:

   - The modification mass considerably depends on the airplane purpose. The mass increase in comparison with the base airplane, on which base the modification was created, is noted;
   - The modification mass increase takes place at all stages of its lifecycle: during designing, after prototype testing, in series production, and in the certification process. The total increase of the modification mass in comparison with the base airplane reaches 27–30%.

2. To minimize modification mass growth, the method and models for structurization of the modification mass into starting and takeoff has been developed. Thus, the starting mass is presented in form of models of its dependence on the main parameters, implemented in the modification. It was established, that the starting mass increase is stipulated by the requirement of necessary carrying capacity and flight range of the modification, and also by:

   - The level and experience of the company developing the military transport modification;
   - Lack of time for designing.

   Takeoff mass growth is determined by:

   - Absence of necessary resources;
   - Mass increase of component items, supplied by co-executors;
   - Necessity of the modification of standardization and unification increase.

3. In evidence of the application of the proposed method and the models, quantitative evaluation of the starting and takeoff masses of modifications of so efficient airplanes as An-26 (Ukraine), F-27 (Holland), and B-727 (USA) is given.

4. Influence of takeoff mass variation of B-747, IL-96 and A340 airplane modifications on worsening the takeoff and landing parameters such as: takeoff run, landing run, and on the specific indexes of their productivity and fuel efficiency

per unit of added mass has been estimated. That is required when evaluating a specific modification competitiveness.

References

1. Los A (2020) Sozdanie modifikacij—osnovnoe napravlenie razvitija voenno-transportnyh samoljotov. In: Kosmicheskaja tehnika. Raketnoe vooruzhenie, vol 1, pp 114–120. KB «Juzhnoe», Dnipro
2. Los A, Shmyrov V, Riabkov V (2020) Tendencii v razvitii operativno-takticheskih voenno-transportnyh samoletov. In: Otkrytye informacionnye i komp'juternye integrirovannye tehnologii, vol 87, pp 72–82. KhAI, Kharkiv
3. Los A (2016) An-178 high level roadmap. Technical act of the SE Antonov, Kiev
4. Electronic Code of Federal Regulations. Part 25—airworthiness standards: transport category airplanes. http://www.ecfr.gov/cgi-bin. Last Accessed on 27 April 2015
5. Electronic Code of Federal Regulations. Part 33—airworthiness standards: aircraft engines. http://www.ecfr.gov/cgi-bin. Last Accessed on 27 April 2015
6. ICAO. Global and regional 20-year forecasts, pilots, maintenance personnel, air traffic (Doc 9956). https://store.icao.int/en/global-and-regional-20-year-forecasts-pilots-maintenance-per sonnel-air-traffic-doc-9956. Last Accessed on 20 Jan 2021
7. IATA. Airline industry forecast. http://www.iata.org/publications/Pages/airline-industry-for ecast.aspx. Last Accessed on 21 Jan 2021
8. Komarov V (2017) Conceptual aircraft design: electronic textbook. Samara State Aerospace University, Samara
9. Kiva D, Grebenikov A (2014) Scientific backgrounds of integrated design of transport category aircraft. National Aerospace University «Kharkiv Aviation Institute», vol 1, Kharkiv
10. Sikulsci V (2017) Study of the process of shape formation of ribbed double-curvature panels by local deforming. Eastern-European J Enterp Technol 4(1(88)):43–49. https://doi.org/10.15587/1729-4061/2017.108190
11. Hill O (2014) Aircraft modifications: assessing the current state of air force aircraft modifications and the implications for future military capability. http://www.dtic.mil/dtic/tr/fulltext/u2/a462766.pdf. Last Accessed on 14 June 2014
12. Process for verifying the coherence between aircraft take-off parameters and an available runway length, Pat. 8380372B2 USA: Int CI. G06F19/00. Appl No. 12/537478: http://www.freepatentsonline.com/8380372.pdf. Last Accessed on 11 Sep 2015
13. Valdés R, Comendador F, Gordún L, Nieto F (2011) The development of probabilistic models to estimate accident risk (due to runway overrun and landing undershoot) applicable to the design and construction of runway safety areas. Saf Sci 5(49):633–650
14. Shirley C, Schetz J, Kapania R, Haftka R (2014) Tradeoffs of wing weight and lift/drag in design of medium-range transport aircraft. J Aircr 3(51):904–912
15. Anipko O, Mirgorod J, Prijmak A (2013) Perechen' pokazatelej svojstv i baza dannyh TTH transportnogo letatel'nogo apparata kak slozhnoj tehnicheskoj sistemy. In: Integrovani texnologiyi ta energozberezhennya, vol 1, pp 123–126, Kharkiv

Decision Support with Bayesian Influence Network During UAV Flight Control

Maksym Herashchenko (ID)**, Kirill Bashinsky** (ID)**, Yurii Kamak** (ID)**, Serhii Nesterenko** (ID)**, Serhii Rudnichenko** (ID)**, and Oleksandr Isachenko** (ID)

Abstract   When conducting unmanned aerial vehicle (UAV) test flights, the air traffic controller constantly monitors their performance, taking into account the state and trends of the environment and the technical parameters of test flights UAVs are performing. This person is the responsible one who makes decisions in the context of these tasks (decision-maker, DM). One of the tasks is to make a timely and reasonable decision on whether to perform an immediate landing of a UAV in the presence of adverse factors or to continue the test flight. The article describes the building process and the result as a possible version of the decision support system (DSS) based on the Bayesian network of influence. The Bayesian network of influence is described as a variant of the Bayesian network of trust, which uses nodes that simulate DM solutions and possible loss of working time in the event of UAV recovery after a possible accident. The simulation results revealed a synergy effect in the event of simultaneous action of several adverse test factors, in which the total negative effect of these factors is higher than the sum of the effects in the case where each factor acts alone. The practical results of research which can be used in simple cases are revealed. However, the article concludes that the effective operation of the proposed DSS is possible only in a fully functional environment of probabilistic modeling.

Keywords Unmanned aerial vehicle · Decision making · Bayesian network of influence

1   Target Setting

The development and production of unmanned aerial systems (UAS) require a significant number of tests at different stages of their life cycle. UAS testing includes as a

M. Herashchenko · Y. Kamak · S. Nesterenko (✉) · S. Rudnichenko · O. Isachenko
State Science and Research Institute of Armament and Military Technics Testing and Certification of the Ukraine Armed Forces, 1 Strilets'ka Str., Chernihiv 14033, Ukraine

K. Bashinsky
1285Th Office of Military Representatives in the Defense Ministry of Ukraine, 8 Moskovs'ka Str., Kyiv 01010, Ukraine

component a large number of flights of unmanned aerial vehicles (UAV). The organization and management of such flights are usually performed by a test crew headed by its commander. In these circumstances, this one is the responsible decision-maker (DM). For specificity, only that part of his area of responsibility is considered here, which includes mainly two tasks:

- timely implementation of all activities planned for the test as a set of experiments;
- implement the UAV flight manager functions in terms of ensuring the safe and productive performance of test flights.

When conducting UAV flights, DM constantly monitors the progress of their implementation, and at the same time takes into account the state and trends of the environment and the technical parameters of UAVs that perform test flights. One of the tasks of DM is to make a timely and conscious decision on whether to make the immediate UAV landing or to continue the test flight under the objective external conditions that occur at the time of the decision. Each of these decisions has both positive and negative effects, and these effects have for the most part the opposite effect on the intermediate and final test results. For these reasons, the decision-making process is subjectively difficult.

The need to make such a choice is accompanied by the uncertainty of the situation, which arises from that facts:

- at the same time there are many factors of the environment and the UAV state, which often affect the results of a decision in opposite ways;
- DM's tasks outlined above are competitive. It is desirable to solve them simultaneously, but in most critical situations they often become mutually exclusive.

For these reasons, there is a new task of creating a certain decision support system (DSS) for DM for the UAV test flight management, the purpose of which can be considered as the actualization of possible losses and risks in decision-making.

2 The Aims of Article

The purpose of this article is to describe a possible variant of DSS based on the Bayesian network of influence (BNI) [1], the use of which during test flights of UAVs will allow:

- increase the efficiency of tests;
- provide more safety for both personnel and technical means involved in the test.

3 Actual Scientific Researches and Issues Analysis

The UAS emergence and its rapid progress in recent decades have shown [2] that they are relatively inexpensive but highly effective means of solving a wide range of tasks, both military and civilian. In addition to reconnaissance and surveillance, which historically emerged among the first ways of the UAS use [3], their combat use as a means of delivering strike weapons has proved very effective, as shown by the experience of recent local military conflicts [4].

In the practical application of UAVs, their flight control is performed according to the usual standards [5]. However, at the stages of research and development, when UAV test flights are performed, these rules are applied in simplified versions, taking into account the special conditions of research polygons of organizations that conduct testing and certification of UASs. It should be noted that so far no proper attention has been paid to the development of adequate specialized information support for UAV test flight control. Specifically, this applies to DSS developments aimed at facilitating deciding by DM in test situations with uncertainty, lack of decision-making time, and the difficulty of taking into account the many external factors that often contradict each other.

The general principles and sequence of creating modern decision support systems in conditions of uncertainty are set out in many studies, examples of which are:

1. Work [6], which outlines the approach to the creation of DSS, based on the idea of DM training on the example of previously made decisions and analysis of their consequences. To do this, a computer environment is created to interpret scripts with the exception situations from the subject area (essentially a computer game), working with which, DM learns to make more effective decisions.

2. Study [7], which contains information on DSS that can be used in the work of situational control centers. To facilitate decision-making, the author proposes to use a one-dimensional utility function, which is calculated according to the methodology developed by him based on expert opinions and information from several databases. The work offers, along with the already known [8], another rather complex method of convolving the vector of partial efficiency criteria into one indicator, based on which DM is proposed to make decisions.

3. In [9] the development and definition of DSS characteristics based on fuzzy situational networks are described, in which the set of states of the controlled system and control influences are represented by a graph, the nodes of which correspond to the system existing at a certain time in a certain state with a certain probability. The numerical basis of the method is a matrix of results of pairwise comparison concerning the degrees of partial criteria consistency, the list of which is determined expertly. Such a model allows the performance of probabilistic calculations, based on which is a sub graph of the output with the maximum score, the end nodes of which are interpreted as images of the recommended optimal solutions.

The implementation of the above approaches can be illustrated by several examples of DSS described in the literature.

For example, the DSS described in [10] is included in the integrated combat control systems of the submarine (SM). The system is designed to assist the SM commander in an objective assessment the tactical circumstances and based on this assessment to make proposals for the best course of action. The working organization of the integrated command system contains the recommendations DSS about doing by the submarine the combat operations. The method of forming proposals to the commander is not described.

In [11], data concerning general-purpose DSS are shown, which provides DM solutions based on vector estimates, the individual coordinates of which (partial indicators) correspond to the DM's subject area perception. The model includes a description of management scenarios and a multi-criteria model for comparing alternatives. The mentioned vector estimates are generated by the system based on pseudo-statistical data obtained by simulation.

In [12], the attempt to describe DSS for the UAV operator is made, whose work is essentially reduced to the recognition of images of reconnaissance objects by filtering the images obtained, highlighting features, and comparing them with reference images stored in the database. These actions are required to calculate the coordinates of the selected target and "form recommendations to the UAV operator", the nature of which is not specified in the article. There is also no data on how to make recommendations.

The simulation application of modeling development of DM solutions of different levels can be illustrated by the developments of the Belarusian company "Belfortex", which based on a unified platform "Fort-2" [13] are made. This platform implements the principles of the High-Level Architecture (HLA) [14], developed in the United States in the mid-1990s on behalf of the United States Department of Defense.

An example is the "Complex of mathematical models for playing up the fighting of groups of troops FORT" [15]. This complex is a tool for commanding the armed forces at the General Staff level and, according to the description, it supports the simulation of the following tasks:

- substantiation of building an air force (AF) group and air defense forces (ADF);
- effectiveness study the enemy's tactical techniques in overcoming him the air defense system;
- assessment of combat capabilities and effectiveness of the created military group;
- directions of armament modernization substantiation for the AF and ADF.

In [15] it is stated that the system described can be optionally supplemented by algorithms of computational and optimization nature, namely:

- optimization of the choice of positions of anti-aircraft missile systems;
- automated fighting order construction;
- the active noise interference impact study on the results of hostilities;
- determining the effectiveness of electronic warfare and others.

Analysis of these data shows that the DSS developments authors focus mainly on providing DM as much information as possible about the situation (problem)

the DM wants to make the decision. Often it is done using some kinds of multi-criteria convolution methods. Postulated in the number of works taking into account the vagueness (lack of certainty) of the input information is reduced to solving not a stochastic but a deterministic problem, which, at best, is based either on expert knowledge or methods of pairwise ranking criteria, or a combination thereof.

This paper proposes a prototype DSS, which the above-described shortcomings of existing approaches and solutions take into account. Its methodology is based on a well-developed mathematical apparatus of probabilistic inference [1]. It is one of the variants of Bayesian networks of trust, namely networks of influence [1], which feature is the ability to model the DM's decision and predict the consequences. The use of Bayesian influence network technology allows combination expert experience with correct probabilistic modeling of the situation and, besides, ensures the adaptability of the model by refining its parameters during the usage statistics accumulation.

4 DSS Model Building

UAV test flights (TF) are usually carried out in the regulated conditions [5] of the military proving ground (aerodrome). There are representatives of the UAS organization-developer, as well as representatives of the customer in the test team (TT, Table 1). Latter are the staff of the state organization, which is responsible for the objectivity of the test results. Both these staff groups take part in the organization and conduct the testing.

In these circumstances, the TT head is the DM that manages the UAV's flights, i.e. DM is the one who decides to launch the UAV, gives operational orders on the course of the flight during its implementation, and decides to do immediate action in case of abnormal or threatening situations during the flight. Among the latter is the decision on whether to continue the TF, or whether to do the immediate UAV landing to avoid undesirable consequences.

DM decisions regarding the UAV flight and the adoption (or non-acceptance) any measures in connection with current situations are influenced by factors that can be conditionally classified into factors external to UAVs and factors that can be considered internal. Within this study as the main external factors may be considered the following:

Table 1 The test team staff

From the UAS organization-customer	From the UAS organization-developer
Chief of the TT	Ground Pilot
Specialists in the UAS features/subsystems research	Ground navigator
	Staff of logistic and providing of UAS

- the quality of the geopositioning navigation system (GPS), which characterized by the number of GPS satellites, which is observed stably by the UAS during the TF;
- availability and sufficiency of energy reserve to continue the flight on the UAV board. In most models of small and medium-class UAVs equipped with electric drive, this value correlates with the residual charge of the onboard battery.

The main internal factors are:

- quality and stability the video image transmission simplex channel from UAV video cameras to the ground control station (GCS) of the UAS;
- quality and stability the duplex telemetry channel UAV-GCS.

It should be noted that the factors listed in this list are considered here as the most significant. If necessary, under other conditions, this list can be modified for example, due to other parameters of UAV flight telemetry.

Let the weather get worse during the TF. Before the DM a dilemma appears—to continue the TF or to command to the UAV landing immediately?

The TF continuation, on the one hand, perhaps will allow to complete the measurements planned for this flight, which has a positive impact on the implementation of the test program. On the other hand, with further deterioration of the weather, the landing process itself can lead to the UAV accident or damage. This means significant material losses for its recovery and the inability to continue testing with it until its recovery.

Immediate UAV landing helps to avoid the threat of UAV accidents, the probability of which increases in bad weather conditions. But on the other hand, there will be a break in testing. TF, not completed, will have to be re-planned and must be performed from the beginning another time, which means unproductive waste of time and material resources.

Taking into account written above, it can be concluded that, depending on the DM decision, the resolution of the dilemma mentioned above actualizes at least two main factors that are desirable to minimize simultaneously:

- the risk of UAV accidents, which increases with the TF continuation while the flight conditions deteriorating;
- possible loss of working time in the process of performing tests due to downtime or elimination of the consequences of abnormal situations.

The problem is that DM does not have formalized guidelines in the process of making such a complex decision and, as a result, is often guided only by its intuition, which is unreliable.

Bayesian trust network may be build a in the environment of BayesFusion Genie Academic 3.0 according to the [1], supplementing it with nodes of decision-making and calculation the probable accident risk and the value of possible working time loss. Under these conditions, the Bayesian network of trust acquires new properties, as a result of which the term "Bayesian networks of influence" (BNI) is applied to

them. This network will perform the DSS function when it is open in the GeNiE environment, and with its support, it will be changing its state due to DM actions.

The process of constructing a BNI begins by creating nodes, which are probabilistic variables images, in the environment [1] for modeling Bayesian networks. A simple probability variable (in terminology [1] it corresponds to the Chance node type) is described by its states (outputs) set and a priori conditional probabilities its existing in each of them. These probabilities distribution is given by the conditional probabilities table (CPT), which is the node property. Probability values are determined either expertly or from statistics available.

In our case, these model nodes reflect the expected state of the factors that the DM pays attention to when the one is deciding to continue or terminate the TF.

These nodes are then connected by directional arcs (ribs), which simulate the causal relationships between nodes. An oriented acyclic graph of nodes and ribs forms a Bayesian trust network (BTN). To create additional functionality in BTN, the *Decision* type nodes (solution) and/or *Value* nodes (calculated value) can be included in it. Then such graph usually is called the Bayesian network of influence (Fig. 1). According to the graph theory [16], the node from which the directed edge originates is called an ancestor (or parent node), and the node in which such a directed rib ends is called a descendant, or a child node. Nodes that have no ancestors are called graph leaves.

The BNI in Fig. 1 contains the "*Decision*" node that simulates a discrete variable with two states: the *Fly* state means the decision to continue the TF and the alternative *Land* solution corresponds an order to perform an immediate UAV landing. Node "*Delay*" is introduced to create ability for calculating the probable loss of the test team working time in a particular decision in a given flight situation.

The "*Weather*", "*Satellites*", "*Video*", "*Telemetry*" and "*Battery*" nodes-leaves simulate a group of input flight parameters that the DM takes into account when making a decision. The experience of UAV operation and testing shows that these factors are statistically independent, that is, a change in any of them does not change
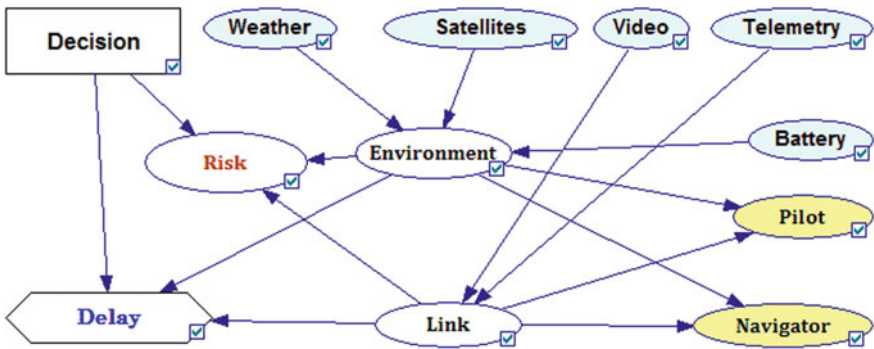
Fig. 1 Bayesian network of influence to support the decisions of the UAV air traffic controller. Nodes are shown visually as *Icon*

the prior probabilities distribution of the others. Therefore, there is no need to model any of their cross-effects.

The "*Environment*" and "*Link*" nodes are intermediate nodes whose purpose is to generalize the values obtained from the input group of graph nodes (leaves). The introduction of such nodes into the BNI structure is a typical technique. Its purpose is to alleviate the expert's load who determine the initial modeling parameters values in the CPTs. In our case, these are the probabilities in the "*Risk*" node CPT and the probable amount of time lost in the "*Delay*" node.

The ground UAV pilot and navigator perform operational UAV control during the flight, and, if necessary, DM can take into account their assessment of the technical feasibility of flight. These assessments are affected by current environmental factors ("*Environment*" node) and communication channels ("*Link*" node). Since in the model on Fig. 1 the impact of these nodes on the nodes "*Risk*" and "*Delay*" is set explicitly by the corresponding ribs of the graph, known output values of the nodes "*Pilot*" and "*Navigator*" create some information redundancy for DM. It makes decision-making more reliable.

To start modeling in BNI, a priori probability distributions must be expertly determined by dialog filling their CPTs in all nodes of the model. Examples of such tables are shown in Fig. 2.

Created according to Fig. 1 BNI with defined CPT of all nodes is ready for use as DSS.

The mathematical basis of working with BNI is the recalculation a priori joint probabilities for events-consequences under the conditions when the result of events-causes (part or all of them) becomes known. These probabilities are called a posteriori. This means that the DM "performed certain experiments" (or simply have read some indicators from the instruments). As a result some of the model variables changed their state from "such, that may be in a number of states with probabilities set by CPT" up to the state of "variables with a known outcome, the probability of which is now known as 100%". Such a change in the state of a probabilistic variable is called "its evidence with result defined".
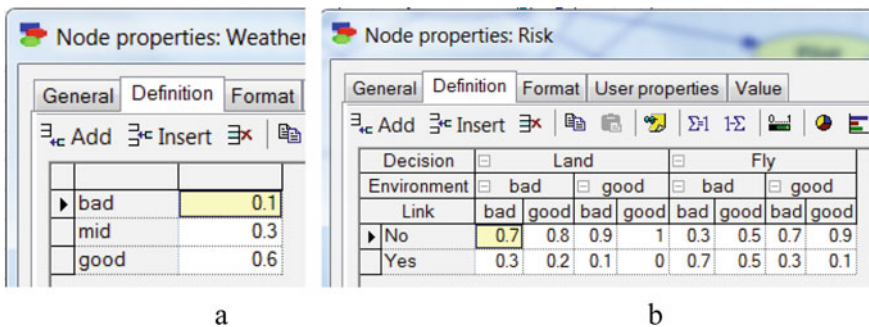
Fig. 2 Examples of the a priori conditional probability tables for BNI nodes, **a**—for the "*Weather*" node; **b**—for the "*Risk*" node

The known Bayes formula [17]

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \qquad (1)$$

establishes the dependence of the conditional probability $P(A|B)$ of the occurrence of event-consequence A in the information presented about the event-cause B (a posteriori probability of event A) on the conditional probability $P(B|A)$ occurrence of event B in the presence of information about event A and a priori probabilities $P(A)$ of event A and the full probability of event B. Therefore, after obtaining the value of variable B (its evidence) by the Bayesian formula, it can be calculated the a posteriori probabilities of variables that are direct descendants.

If, for example, for the proposed model (Fig. 3) mark.

$$P(Decision = Fly) = P(d); P(Environment = bad)$$
$$= P(e); P(Link = bad) = P(l); P(Risk = Yes) = P(r)$$

then for the Risk node, formula (1) takes on a more specific form

$$P(r|l, e, d) = \frac{P(d|r) \cdot P(e|r) \cdot P(l|r) \cdot P(r)}{P(d|r) \cdot P(e|r) \cdot P(l|r) \cdot P(r) + P(d|-r) \cdot P(e|-r) \cdot P(l|-r) \cdot P(-r)}. \qquad (2)$$

In this context, the notation $-r$ means a situation where the event r does not occur.

Formula (2) is cumbersome. It shows that the conditional probabilities calculation became much more complicated when the number of influencing factors of
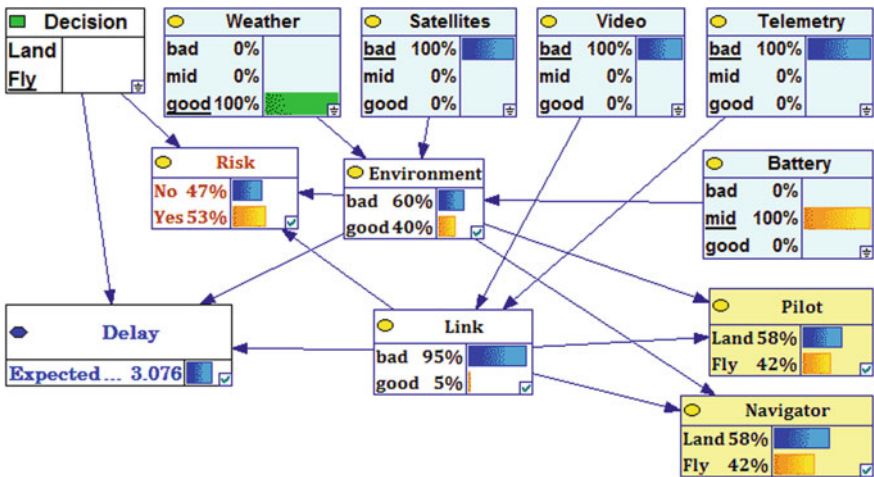
Fig. 3 BNI after performing Bayesian probabilistic inference. Visual display of nodes is switched to *Bar Chart* mode

the dependind event increases. Therefore, for the methods of probabilistic modeling practical use, one need to use appropriate software tools, one of which is the BayesFusion GeNiE. Difficult work on correct manual programming the formulas similar to (2), is automated in it. The user only needs to make a correct graphical model of the simulated variables interaction.

In essence, recalculating the conditional probability of a dependent event means taking into account the additional knowledge from the achieved result of the event-cause.

But the causes and consequences of the described calculation are themselves the causes of other events modeled by the structure of the edges of the graph of the BNI model. Therefore, the GeNiE software environment iteratively continues to recalculate the nodes a priori probabilities in their a posteriori values, while performing a complete bypass of the model oriented acyclic graph. This process is called "Bayesian probabilistic inference" [17]. In the BNI model, it will end when posterior state probabilities values will be calculated for all nodes. An example of a BNI with updated a posteriori states after Bayesian inference is shown in Fig. 3.

In Fig. 3, the nodes-sheets, marked ⬍, have certain definitions, and the probabilities of the nodes, marked ☑, are calculated during the execution of the probabilistic inference.

DM makes decisions using the DSS proposed here according to the following algorithm. The initial state of the model—it is open in the environment BayesFusion Genie Academic.

1. For leaf nodes "*Weather*", "*Satellites*", "*Video*", "*Telemetry*" and "*Battery*", evidence of those states of the corresponding factor, which DM observes at the current time, are entered. This is done by double-clicking on the desired result name of the desired node. For example, in Fig. 3 it is indicated: the weather is good (*good* = 100%), UAS sees GPS satellites poorly (*bad* = 100%), video from the UAV comes unsatisfactory (*bad* = 100%), telemetry refused (*bad* = 100%), the remaining battery charge is medium (*mid* = 100%).At the *Decision* node, double-click on the name of the decision, which, say, this person considers possible to make. In Fig. 3 selected solution is *Fly*—to continue the test flight.

2. The GeNiE application according to these data calculates and demonstrates the initial simulation: the risk of UAV accident is 53%, which can be considered serious. The possible estimated loss of working time of the TT to eliminate the consequences of the accident may amount to 3076 working days, as the serious level of possible damage to the UAV will require, accordingly, serious repairs.

3. The DM, based on this data, can try to change his possible solution from *Fly* to *Land* (perform an immediate landing). Then the initial values will change to (*Risk Yes* = 21%, i.e. small; possible loss of time *Delay* = 0.5 working days, as this is a possible loss of time to re-perform the urgently interrupted incomplete testing flight).

4. The DM officially notifies the subordinate of the decision one has reached, understanding the risk value and the possible consequences calculated by the

DSS. It is clear that all responsibility for the final decision remains with the DM, DSS acts only as an advisory system.

5. If necessary, go to point 1 and the cycle of DSS is repeated.

5 Results

Here was performed a study of the simulation results on the developed model (Fig. 3) of the different decision-making situations if DM decides to continue the test flight (*Decision = Fly*). Given the small number of input parameters in the model, the analysis of the risk of UAV accidents during the continuation of the test flight with different combinations of adverse factors was performed. The results of these model experiments are presented in Fig. 4.

If the limit of allowable risk accept as $R < 50\%$, then from Fig. 4 it is clear that with the simultaneous action of three or more adverse factors, the continuation of the flight is an irrational decision, the implementation of urgent landing UAV looks as more correct strategy.

An interesting issue is the study of the quantitative extent of the influence of individual factors on the results of DM decision-making in terms of ensuring trouble-free flight. From the experience of flights, it is well known that the greatest impact on the risk of accidents during landing has adverse weather, and this fact is taken into account by experts in determining the CPT for the nodes of the model in Figs. 1 and 3. To illustrate the quantitative measure of the influence of adverse weather, which is a property of the model proposed here, according to the simulation, Fig. 5 is shown.

Zone **a** of Fig. 5 shows how the risk increases when the influence of weather is added to the influence of single factors; the **b** zone corresponds to situations when the influence of weather is added to two factors; zones **c** and **d** respectively, when the influence of weather is added to three and four factors.
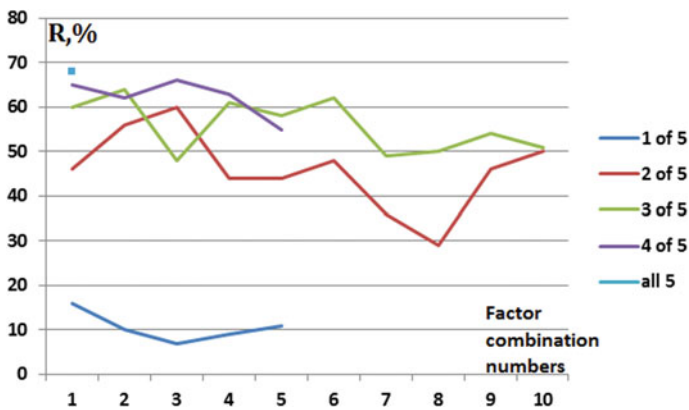
Fig. 4 Risks of UAV accidents during continued test flight under the influence of various combinations of adverse environmental factors and communication conditions
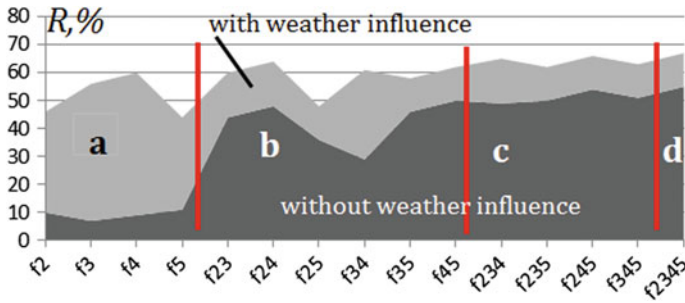
Fig. 5 Additional increase the UAV accidents risk during flight continuation, when bad weather factor is added to the combinations of adverse factors

The simulation results showed that the synergistic effect of each factor in addition to the already existing combination of factors is manifested in a qualitatively identical way, similar to that shown in Fig. 5. Synergism is especially pronounced when at least one additional factor is added to a single adverse factor.

The comments made above regarding the simulation results in Figs. 4 and 5 are practical but limited in nature, and therefore can facilitate DM decision-making only when one does not have access to DSS during control UAV test flights. But these recommendations do not reflect the full complexity the factors interaction in the process of testing, so the real benefit is available only in conditions when the DSS based on BNI will operate in a realistic simulation environment capable to perform Bayesian inference.

6   Conclusions

1. Bayesian trust networks are an adequate and effective mechanism for building decision support systems in conditions of unclear situations that promptly arise during the testing of military equipment.
2. The simulation results revealed a synergy effect in the event of simultaneous action of several adverse test factors, in which the total negative effect of these factors is higher than the sum of the effects in the case where each factor acts alone.
3. A practical recommendation in the absence of DM access to the DSS is to decide on the urgent landing of the UAV if at least three adverse factors begin to act simultaneously during the UAV test flight.
4. Despite the identification of the practical research results that can be used in simple cases, the actual work of the proposed DSS should be performed in a full-featured probabilistic simulation environment, an example of which is the application BayesFusion GeNiA.

References

1. GeNIe Modeler 3.0.R2. User Manual. BayesFusion LLC (2020) https://support.bayesfusion.com/ docs/GeNIe.pdf. Last Accessed on 15 Jan 2021
2. Hayat S, Yanmaz E, Muzaffar R (2016) Survey on unmanned aerial vehicle networks for civil applications. A communications viewpoint. IEEE Commun Surv Tutorials 18(4):2624–2661
3. Najiya K, Archana M (2018) UAV video processing for traffic surveillence with enhanced vehicle detection. In: 2018 second international conference on inventive communication and computational technologies, ICICCT, pp 662–668, Coimbatore
4. Walsh J (2020) The effectiveness of drone strikes in counterinsurgency and counterterrorism campaigns. Strategic Studies Institute, US Army War College
5. Rules of the air (2005) Annex 2 to the convention on international civil aviation (10th ed.). International Civil Aviation Organization
6. Yevgrafov I (2009) Development of a decision support system in emergency information conditions. PhD Dissertation. AS of Russia, VINITI, Moscow
7. Cherkasov A (2011) Development of mathematical and algorithmic support for adaptive decision support systems in situational centers. PhD Dissertation. Krasnodar, Kuban STU
8. Peregudov F, Tarasenko F (1997) The basics of system analisys (2nd edn.). Tomsk, NTL Publishing
9. Zernov M (2008) Methods and software for decision support based on fuzzy situational networks. PhD Dissertation. Smolensk, Sm. Branch of Moscow Power Engineering Institute
10. Diment A, Grigoriev A, Mikhailov V, Konson A, Lukichev V (2017) The role and place of the decision support system in the integrated combat control systems of submarines. Mar Electron 2(60):10–14
11. Buyanov B, Lubkov N, Poliak G (2006) Management decision support system using imitation modeling. Probl Manage 6
12. Koroliuk N, Yeremenko S (2015) Intelligent decision support system for control of unmanned aerial vehicles at the ground control post. Inf Process Syst 8(133):31–36
13. Software Development Based on the "Fort-2" Technology. http://www.belfortex.com/page/show/77. Last Accessed on 10 Jan 2021
14. STANAG 4603 Ed. 2 (2015) Modelling and simulation architecture standards for technical interoperability: high level architecture (HLA). NSO(AC/323)0234 (2015) NMSG/4603
15. A complex of mathematical models for playing with the combat actions of troops groupings "Fort". http://www.belfortex.com/page/show/9. Last Accessed on 10 Jan 2021
16. Burkatovskaya Y (2014) Graph theory, vol 1. Tomsk Polytechnical University Publishing, Tomsk
17. Tulupiev F, Nikolenko S, Sirotkin A (2006) Bayesian networks: logical and probabilistic approach. Nauka Publishing, St.Petersburg

Modeling and Analysis of the Main Statistical Parameters of the Testing System of Special Equipment

Ihor Korniienko , **Svitlana Korniienko** , **Volodymyr Dmytriiev** , **Anatolii Pavlenko** , **and Dmytro Kamak**

Abstract The purpose of the testing institute, which is considered in the article, is to conduct mass independent tests of special equipment. Usually, testing is a stage in large government programs, therefore there is a great responsibility in conducting tests with a high quality. Consequently, delays and especially disruptions of the testing procedures are also inadmissible. This requires finding ways to optimally plan all the operation processes of the testing institution upon receipt of an intensive flow of requests for testing special equipment. Modeling of the institute's activity was carried out with the help of the theory of queuing systems. Based on the general data set for the last few years, the central, second and third moments of the arrival rate of requests for testing and the test rate are calculated. The statistical probability of deviation of testing beginning times due to the influence of certain external and internal factors inherent to the testing process is determined and its statistical characteristics are calculated. The exponential law of distribution of the arrival rate of testing requests and the service rate is determined, that reduces the queuing system to the following kind $M/M/n$. Based on the analysis of the behavior of the input flow of requests for testing, the general scheme of queuing formation is adjusted. The limiting parameters of the testing system are calculated and the minimum requirements for the number of service channel are formulated. To ensure adequate capacity, requirements for the formation of reserves in the testing system are justified. A further solution to the problem of researching the effectiveness of the testing institution is to model

I. Korniienko (✉) · S. Korniienko
Chernihiv Polytechnic National University, Shevchenka 95, Chernihiv 14035, Ukraine

S. Korniienko
e-mail: cornel@ukr.net

V. Dmytriiev · A. Pavlenko · D. Kamak
State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Strilets'ka 1, Chernihiv 14033, Ukraine
e-mail: vadmitriev@ukr.net

A. Pavlenko
e-mail: agpav@ukr.net

D. Kamak
e-mail: dkam@meta.ua

the behavior, distribution and prioritization of sub-flow of requests for testing for different types of special equipment and search for optimal methods of distribution of personnel on testing teams according to the structure of the input flow of requests for testing.

Keywords Testing · Queuing · Special equipment · Input flow · Testing requests

1 Introduction

In recent years, the rearmament program of the Armed Forces of Ukraine has been intensively implemented. The priority of the program is the development of new domestic specimens of special equipment or a deep modernization of existing ones. This has led to a significant increase in the total number of testing requests for specimens of special equipment which can be carried out only by one government testing institution. In our case, the model of such a testing institution is studied. As it turned out, the growth of the flow of requests for testing increased the risk of late completion of tasks of testing special equipment at different stages of its life cycle. Delays at the testing stage might lead to further delays in the technological process of development and mass production of specimens of special equipment.

In order to service the increased input flow of testing requests, the structure of the testing institution was expanded and new specialized units were created. To some extent, this played a positive role, but, as practice shows, did not eliminate the risk of delays in conducting mass tests. Further increase and complication of the testing institution structure is unacceptable due to the increase in its maintenance costs. In order to minimize the risk of disruptions during the processing of test requests and to normalize the planned processes for preparation and testing, it was proposed to study the processes in the testing institution, which are somehow related to the all process testing. This study aims to identify problems that lead to test delays and to develop methods for optimal management of the testing system.

The currently well-developed analytical apparatus of queuing theory was chosen as a research tool [1]. The choice of this mathematical approach was based on the general idea of queuing theory: the existence of a requests input flow which arrives in the service system. According, a testing request is a requirement to be fulfilled by a testing institution. With, statistical estimations of the testing system productivity are received already at the analytical modeling stage. They are able to indicate problems during the servicing of the input flow of requests and allow obtaining optimal parameters for the system service of testing institute.

2 Background and Related Work

This work continues the series of publications covering the optimization of testing processes. Particularly in [2], the usage of queuing systems theory for modeling processes in a testing institution is substantiated. The main theoretical studies of queuing systems theory, which are presented in [1, 3], have found application in the research of dynamic systems in many areas, including computer science [4], social infrastructure, education and manufacturing [5, 6], solution of military-special problems [7], modeling of military-technical systems [8, 9], study of combat operations effectiveness [10], comparative assessment of military equipment [11].

In [2] the analysis of parameters of an input flow of requests for testing and its possible influence on the efficiency of functioning of the testing institution is carried out. The mathematical study of probability theory [12] and mathematical statistics [13] served as the basis for obtaining the parameters of the input flow of requests for testing. In addition, [2] substantiates and defines the unit of time that best meets the property of consistency for statistical data sets [12]. In [2] the statistical mathematical expectation value of the number of requests per week is determined to be $M^*[X] = 1.79$. Also, in [2] it was established that the probability of time arrived of testing requests is distributed according to exponential law and the number of requests per week is distributed according to the Poisson law with parameter $\lambda \Delta t \approx 1.87$. The established distribution of the time of receipt of the request according to exponential law is important and convenient for further modeling of the testing institution as a queuing system [14].

Studies of the behavior of requests in a queue with different importance and urgency [15, 16] served to form an approach to the differentiation of testing requests and the creation of a system of queues, which is also given in [2]. In [17] considered the problem of service rate control of a single-server queueing system, where given a model, which useful in providing a tractable alternative for the control of service centers with nonstationary arrival rates.

3 Research Method

3.1 Description of the Model of Testing Institution

The testing institution is a structure whose main activity is related to the systematic and infallible conduct of mass tests and certification of specimens of special equipment. The testing process itself at the institution can be represented by the queuing system shown in Fig. 1 [2].

Requests for testing specimens of special equipment are sent to the system input (to the testing institution), thus forming a system of queues. The waiting time of a request in the queue is a necessity related to formation time of a test team, preparation of the methodological component of the tests, training of operators who will operate
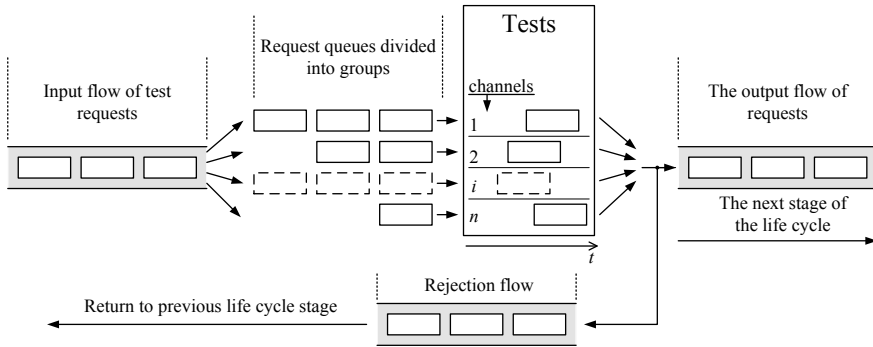
Fig. 1 Diagram of the testing institution queuing system

the equipment during the tests, preparation of the test site, measuring systems and waiting for testing. Moreover, we note that some requests do not require the above technological operations and, accordingly, their time waiting in queues is less. An example of such an occurrence is a specimen that was previously tested and all these operations were performed for it, but during the tests, the specimen was rejected due to deficiencies and returned for testing after their elimination by the developer.

The group of queues is formed because of the heterogeneity of specimens of special equipment, for the testing of which, as a rule, narrow-profile specialists are involved. That is, each queue consists of requests for testing the same type of specimens of special equipment. For such requests, the waiting time in the queue $\omega_n = \omega_n^t + \omega_n^c$, where ω_n^t – the waiting time for the implementation of technological processes of preparation for testing the specimen on request C_n; ω_n^c—waiting time of the request due to the occupation of the i-th service channel. For re-occurring requests which do not require preparatory action: $\omega_n = \omega_n^c$.

In [2] the main statistical parameters of the input flow are defined:

- mathematical expectation of the number of applications per week—$M^*[X] = 1.79$;
- variance and standard deviation of the input flow, respectively—$D^*[X] = 1.93$ and $\sigma^*[X] = \sqrt{D^*[X]} = 1.39$;
- Poisson's distribution parameter (input flow intensity) $\lambda \Delta t \approx 1.87$;
- coefficient of variation of time intervals between requesti $v(A(t)) = 1.08$.

In addition to the existing factors and force majeure factors that determine the nature of the input flow of testing requests, it can be expected that some contribution was caused by time deviations associated with the quarantine measures of the COVID-19 pandemic.

3.2 Input Flow of Testing Requests

Rationale behind the Study. Despite the results obtained in [2], the study of statistical parameters of the input flow of test requests has not lost its relevance and remains necessary for a couple of reasons:

- in [2] there was opted for a time interval of 1 week ($\Delta t = 1$) for statistical analysis of the random variable $\alpha(\Delta t)$—the number of test requests received during the time interval Δt. Note that a week is considered as a time interval of 5 working days. To obtain reliable results for a statistical experiment, according to [12], the number of observed time intervals $n_{\Delta t} \geq 500$. Under this condition, it can be argued that the statistical frequency of the occurrence of a random variable $W_\alpha(\Delta t)$ will be equal to the probability $P_\alpha(\Delta t)$, i.e. $W_\alpha(\Delta t) \rightarrow P_\alpha(\Delta t)$. Since there are only available statistics for 4 years $n_{\Delta t} = 209$, additional observations are required;
- during 2020, there were time deviations in the receipt of test requests related to the COVID-19 pandemic, so it is important to study the correspondence of the flow of requests to the Poisson exponential distribution:

$$P_\alpha(\Delta t) = \frac{(\lambda \Delta t)^x}{x!} e^{-\lambda \Delta t}; \quad x = 0, 1, \ldots; \tag{1}$$

- monitoring of the flow of requests should be carried out constantly due to possible deviations of intensity associated with advancements in the research, development and production areas, the number of government testing requests, the rise or decline of economic processes in the country, etc.

Modeling and Analysis of Results. When processing the general set of statistical data at $n_{\Delta t} = 209$ the following parameters of the input flow of test requests are obtained:

- statistical mathematical expectation of the number of testing requests per week $M^*[\alpha(\Delta t)] = 1.83$, statistical variance $D^*[\alpha(\Delta t)] = 1.95$ and statistical standard deviation $\sigma^*[\alpha(\Delta t)] = 1.39$;
- a study of the nature of the probability distribution of the random variable confirmed its exponential distribution. The graph of the statistical distribution function is shown in Fig. 2.

The calculation of the Poisson parameter was performed by the least squares method for the linearized function (1), resulting in $\lambda \Delta t \approx 1.94$. Given that $\Delta t = 1$, $\lambda \approx 1.94$ can be taken as the intensity of the testing requests. In comparison with previous results, there is a slight increase in the intensity of the test flow and an increase in variance.

The study of time intervals between related requests was carried out according to the calendar of receipt. If the requests were received in one day, the time interval between them is set to $\Delta \tau_{ij} = 0$. As a result of processing the general set of statistical
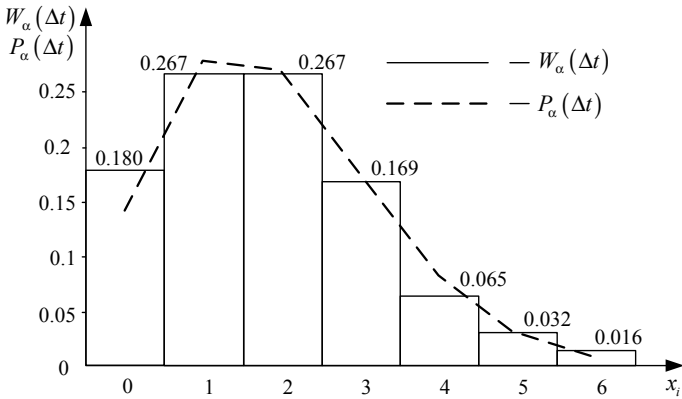
Fig. 2 Input flow parameters of testing requests: $W_\alpha(\Delta t)$—histogram; $P_\alpha(\Delta t)$—statistical distribution function

data $n_{\Delta\tau} = 1461$, the following results were obtained: mathematical expectation of the time interval between requirements $M[\Delta\tau] = 3.79$, the standard deviation $\sigma[\Delta\tau] = 4.16$, the coefficient of variation $\nu = 1.09$. These estimate were calculated with the general set of calendar dates for the entire observation period. Note that the time interval $\Delta t = 1$, for which 2 days are non-working days was chosen for our research, and which, in principle, cannot be included in the statistical estimate for the intervals of testing request receipts. Moreover, the receipt of test requests on non-working days is an impossible event. Therefore, it is justified to reduce the sample size by 2/7, and assume accordingly that $M[\Delta\tau] = 2.71$, the standard deviation will be $\sigma[\Delta\tau] = 2.97$.

3.3 Testing System

Formation of the List of Parameters of the Testing System. Let's define parameters of system of queuing which characterize its performance:

- numerical characteristics of the statistical distribution of the processing time of the request (conducting practical tests)—mathematical expectation $M^*[x]$, variance $D^*[x]$ and standard deviation $\sigma^*[x]$, where x—is the time of testing the sample of special equipment;
- the intensity of processing test requests $\mu = 1/M^*[x]$;
- the nature of the law for the request processing time probability distribution in the system;
- numerical characteristics of the statistical distribution of the waiting time of the request in the processing queue—mathematical expectation $M^*[\omega]$, variance $D^*[\omega]$ and standard deviation $\sigma^*[\omega]$;
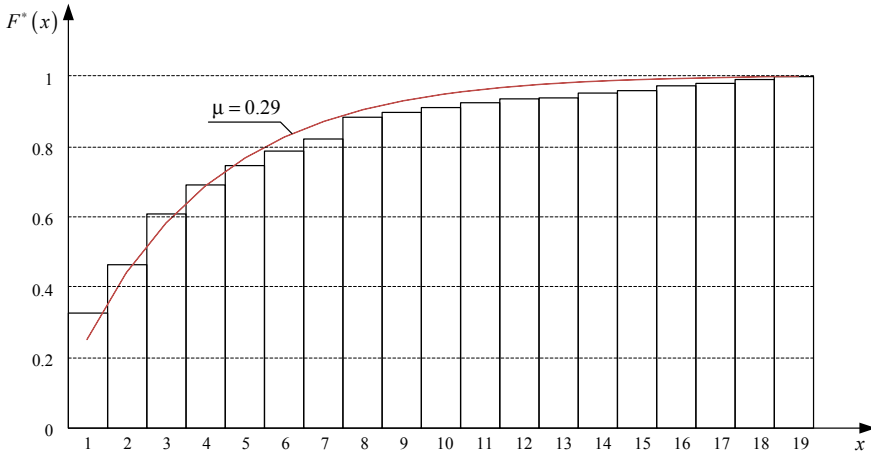
Fig. 3 Statistical distribution of the request service time probability

- statistical probability of successful tests $P^*(C_n)$.

Modeling. As a result of researching the general set of statistical data (on carrying out tests of specimens of special equipment), the following parameters of processing test requests were determined.

Statistical estimates for the processing time of requests (test of the sample): mathematical expectation $M^*[x] = 5.85$, variance $D^*[x] = 85.24$, standard deviation $\sigma^*[x] = 9.23$. Estimates were performed using the assumed time unit of $\Delta t = 1$ week. The statistical value for the intensity of processing test requests $\mu = 0.17$.

To establish the nature of the law for the request service time probability in the system, the statistical histogram of the distribution of the request service time was built (see Fig. 3).

To verify the correspondence with the exponential law of the distribution of statistical results of the time of testing of a specimen of special equipment, we linearized the function $F(x) = 1 - e^{-\mu t}$ and calculated the value $\mu = 0.29$ with the method of the least squares. The theoretical curve, which is built on the results of calculations, is shown in Fig. 3. The level of correspondence of experimental and theoretical results is checked using the Pearson consistency criterion. For the accuracy parameter $\alpha = 0.05$ and degree of freedom $k = 17$ the result $\chi^2 = 11.99$ is obtained, which confirms the hypothesis of the Poisson character of the distribution of the request service time probability.

Statistical estimates for the time the request stayed in the queue were determined for planned ω and actual ω' parameters. The time spent in the queue was determined from the date of the request for testing the specimen to the beginning of the test phase. The mathematical expectation is then $M^*[\omega] = 20.27$ and $M^*[\omega'] = 21.28$; the variance $D^*[\omega] = 734.20$ and $D^*[\omega'] = 777.02$; the standard deviation $\sigma^*[\omega] =$

27.09 and $\sigma^*[\omega'] = 27.87$. Estimates were performed using the assumed time unit $\Delta t = 1$ week. Probability of test delays is $P(\omega \neq \omega') = 0.12$. The mathematical expectation for the delay is $M^*[\omega \neq \omega'] = 0.59$, the variance is $D^*[\omega \neq \omega'] = 11.97$ and the standard deviation is $\sigma^*[\omega \neq \omega'] = 3.46$.

The statistical probability of successful tests is $P^*(C_n) = 0.85$.

Analysis of results. Based on the defined statistical estimations of parameters for the functioning of the testing institution, it is possible to calculate average parameters of efficiency and formulate requirements for the bandwidth of the testing system.

Average time spent on service requests is $\overline{x} = M^*[x] = 5.85$ week (as before, units of time are counted in weeks). According to Little's formula, the average amount of requests which are in the practical testing phase is $\overline{N}_x = \lambda \overline{x} = 11.34$. Ideally, for ensuring a maximum performance of the queuing system the utilization factor should follow: $\rho \to 1; 0 \leq \rho < 1$, and according to [1] $\rho = \frac{\lambda \overline{x}}{m}$, when m—number of service channels, then $m > \frac{\lambda \overline{x}}{\rho}$. Hence, the minimum number of channels of the testing system $m_{\min} = 12$.

Accordingly [2], the total time spent by all requests in the testing system is $\gamma_x(t) = \overline{N}_x t$, then for a year $t \approx 52\Delta t$ and $\gamma_x(t) = 590.14$ (requests-weeks)/For a 40-h working week $\gamma_x(t) = 23605.92$ (requests-hours). The minimum need for personnel directly involved in testing is then $N_{tester} \geq \frac{\gamma_x(t)}{W_{norm}} \overline{n}_{team}$, when W_{norm}—the general norm of working hours, $\overline{n}_{team}$—the average number of personnel in a testing team. Taking for example that $W_{norm} \approx 2000$ (hours), $\overline{n}_{team} = 5$, then the minimum requirement for nominal operating modes $N_{tester} \geq 60$, that is, these personnel are exclusively engaged in practical tests.

The average waiting time for a request in the queue according to the actual data is $\overline{\omega}' = M^*[\omega'] = 21.28$. Respectively, the number of requests that are in the queue at the same time is $\overline{N}_q = \lambda \overline{\omega} = 41.28$. Taking into account the statistical probability of successful tests, the average number of requests that require preparation is $\overline{N}'_q = \overline{N}_q P^*(C_n) = 35.08$. The existing requests are distributed in queues according to the types of special equipment and the provisioning of specialized personnel (Fig. 1). Ideally, at the uniform distribution of specialized personnel for the needs of testing and personnel sufficient availability, the one service channel has $\overline{N}'_q/m$ request. Average time required to prepare one test request is $\overline{\tau}_q = \frac{\overline{\omega} m_{\min}}{\overline{N}_q}$, the intensity of preparation for the test for one request (speed of progress of the request in the queue to one channel) $\mu_q = 1/\overline{\tau}_q$. With the minimum number of channels allowed $m_{\min}$ the one channel has $\overline{N}'_q/m_{\min} = 2.92$ requests, then $\mu_q \geq 0.14$.

Using a similar approach for calculating the minimum required amount of personnel per year for preparing requests for the practical testing phase and taking into account the number of personnel $\overline{n}_{team} = 3$, required for such work, it can be concluded that $N_{researcher} \geq 110$.

4   Discussion

Investigations of the input flow parameters and their comparison with previous results did not reveal significant deviations. Even when we increased the sample size by a third, we have not yet reached the ideal general data setsize, for which it is safe to say that probability parameters are obtained. However, the similarity of the results obtained in different studies with an interval of 1 year is already a positive result.

Also, at this stage, we did not see a significant impact of COVID-19 quarantine measures on the nature of the law of distribution of the input flow of requests for testing specimens of special equipment. We confirmed the exponential law of distribution of the amount of requests received within an interval of 1 week. The impact of quarantine measures can be seen in the slightly increased variance in the number of test requests. There is still a further need to monitor the intensity of the input flow of requests for the timely management of the processes that accompany the testing of specimens of special equipment.

The study of the behavior of the requests in the queue showed the need to detail a model in the overall scheme of the testing institution as a queuing system, which is presented in Fig. 1. This representation does not reflect the procedures for the initial processing of requests while they are in the queue. Then the queuing system can be represented by the scheme shown in Fig. 4.
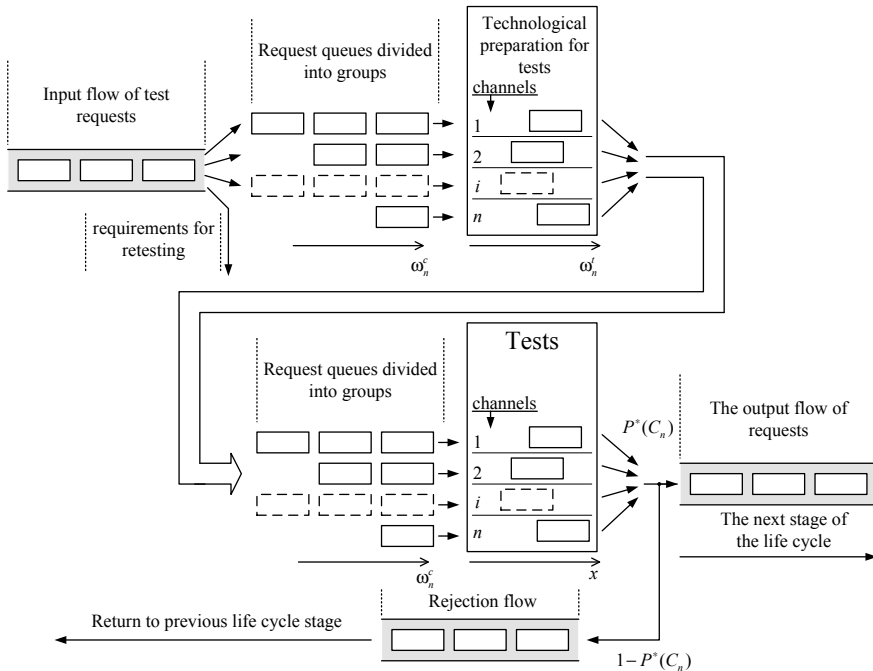
Fig. 4 Diagram of the testing institution queuing system with technological preparation phase

In Fig. 4, the processing of incoming requests requirethe mandatory initial preparation of the requests, if such a request entered the system for the first time. If the request has already been in the system before and was rejected as a result of the tests, it skips the preparation phase and is sent to the queue of the "Tests" service system.

In theory, we can assume that the queue has a limit on the request waiting time, but during our observations such cases were not detected, so in the simulation we can assume that the waiting time of the request in the queue is unlimited. Similarly, the queue is unlimited in the number of requests that can expect processing. We also found the statistical frequency of deviation of the starting time of the tests from the planned one. Such time deviations are the result of insurmountable factors independent of the operation of the testing institution. The need to study such factors was pointed out in [2] and statistics are currently being gathered on this issue.

It should be noted that the performed and presented approximate calculations of the structural parameters of queuing nodes were performed for the worst-case scenario of extreme occupation of test personnel, despite the average load parameters. In addition, the structure of the flow of requests has certain inhomogeneity, estimates of which can be seen in [2]. This means that certain requests may be in the queue, despite the availability of free service channels (for example, due to lack of personnel of the required specialization, which is busy testing another model of the same type of special equipment). In order to prevent this, the presence of a reserve (personnel, equipment i.e.) in the testing system is required, which minimizes the risk of temporary deviation in the servicing of the request. The need for reserve personnel is also shown by the presence of a peak in incoming requests. Despite the stationary nature of the input flow, a significant increase in the intensity of request receipts is periodically observed. The periodicity of the peak load, which can be called "slow oscillations", is shown in Fig. 5.

For these reasons, and taking into account the wide range of special purpose equipment and, often, its uniqueness, in the future it is necessary to separately investigate the issue of the required reserve, and possibly develop a mechanism for the optimal formation of test teams.

Studies of statistics on the processes associated with the preparation and conducting of tests have identified the main statistical characteristics of processing requests. Analysis of the nature of the statistical law of distribution of request processing time (carrying out tests) showed its proximity to a stationary Poisson flow. This hypothesis was tested using the Pearson consistency criterion, which resulted in a positive result. This allows us to study the processes in the testing institution, which is represented by a queuing system of the kind $M/M/n$, which is the best case scenario because of the convenience of the mathematical description of Markov processes for queuing systems of this type. It is also known [7] that service nodes that are designed for Poisson flows will serve other kinds of flows with the same efficiency.
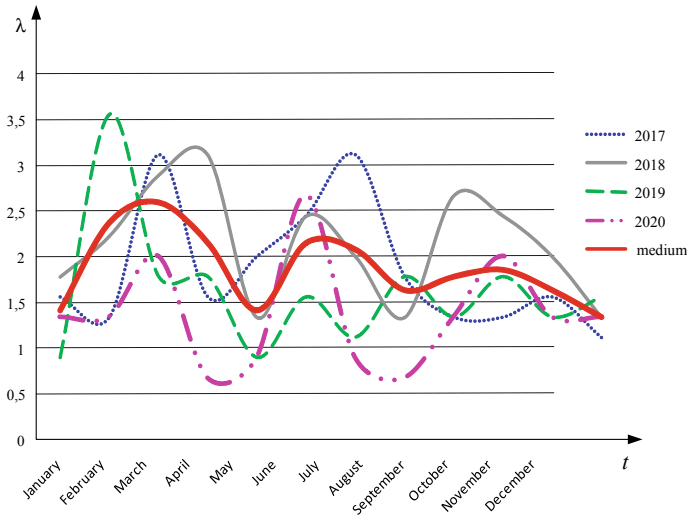
Fig. 5 Smoothed graphs of the annual dynamics of the intensity of the total input flow of test requests

5 Conclusions and Future Work

5.1 Conclusions

The obtained numerical statistical characteristics of the processes accompanying the testing of specimens of special equipment allow for a close approximation to modeling the activities of the testing institution. The established exponential laws of distribution of random variables—intensity of receipt of requests for testing and intensity of requests processing, allow to use the most convenient mathematical approach and considerably simplify the process of modeling.

Based on the analysis of statistical data, we determined the magnitude and nature of the time deviation of the beginning of the phase of practical tests that took place under the influence of various internal and external factors. Despite the presence of such deviations and the need to predict them, their share in the overall flow of input requests is not critical and does not pose significant threats to the test process as a whole. The determined numerical characteristics of the input flow of requests also allowed to model in detail the general scheme of test processes.

The analysis of the established processing parameters allowed to determine the limit parameters of the queuing node and to formulate the minimum requirements for the number of service nodes and, accordingly, the number of personnel. We also substantiated the needs for the formation of reserve in the queuing node, which will ensure adequate capacity and prevent disruption of government tasks.

5.2 Future Work

Investigate and detail the processes of queue allocation and develop an appropriate mathematical study for modeling the behavior of test request sub-flows for different types of special equipment.

Investigate the process of forming test teams (channels in the queuing node) and develop an approach to the optimal distribution of personnel in test teams in accordance with the structure of the input flow of testing requests.

References

1. Kleinrock L (1975) Queueing systems—volume I: theory. John Wiley & Sons, New York
2. Korniienko I, Korniienko S, Dmytriiev V, Pavlenko A, Kamak D (2021) Investigation of the model of testing for weapons and military equipment. In: Shkarlet S, Morozov A, Palagin A (eds) Mathematical modeling and simulation of systems (MODS'2020). MODS 2020. Advances in intelligent systems and computing, vol 1265. Springer, Cham. https://doi.org/10.1007/978-3-030-58124-4_30
3. Kleinrock L, Gail R (1996) Queueing systems: problems and solutions. A Wiley-Interscience publication. New York, NY [u.a.]: Wiley
4. Allen A (1990) Probability, statistics and queuing theory with compute science applications (2nd ed). Academic Press Inc.
5. Matveev VF, Ushakov VG (1984) Queuing systems: manual on a special course "applied mathematics", p 239. Moscow State University (in Russian)
6. Bocharov PP, Pechinkin AV (1995) Queueing theory: the textbook, p 529. University under P.Lumumba (in Russian)
7. Kazachinsky VZ, Levitsky GE (1980) Mathematical methods for solving military special tasks, p 292. Kiev Higher Air Force Academy (in Russian)
8. Nikolić N (2008) Monte Carlo modeling of military Queueing systems—challenge of the initial transience. Belgrade: Andrejevic Endowment/StrategicResearch Institute
9. Jenkins P (2017) Using Markov decision process with heterogeneous Queueing system to examine military MEDEWAC dispatching policies. In: MSc thesis in operations research. Ohio, USA: Air Force Institute of Technology—Wright-Patterson Air Foce Base
10. Li L, Liu F, Long G, Zhao H, Mei Y (2016) Performance analysis and optimal allocation of layered defense M/M/N Queueing systems. Math Probl Eng 2016:1–21. https://doi.org/10.1155/2016/5915918,lastaccessed2021/03/14
11. Mao Z, Yu-juan W, Chao W, Sheng H (2013) 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
12. Ventsel ES (1969) The probability theory, 3rd edn. Nauka, Moscow (in Russian)
13. Cramer H (1946) Mathematical methods of statistics. Princeton University Press, Princeton
14. Gnedenko BV, Kovalenko IN (2012) Introduction to Queuing theory. LKT
15. Saaty TL (1964) Stochastic network flows: advances in networks of queues. In: Smith WL, Wilkinson WE (eds) Congestion theory. University of North Carolina, Chapel Hill, pp 86–107
16. Lakatos L, Szeidl L, Telek M (2012) Introduction to Queueing systems with telecommunication applications. Springer Science+Business Media, New York, NY, USA
17. Kumar R, Lewis M, Topaloglu H (2013) Dynamic service rate control for a single-server queue with Markov-modulated arrivals. Naval Res Logist 60:661–677

Mathematical Modeling and Simulation of Systems in Project Management

Modeling of Project Portfolio Management Process in Banking

Nataliia Yehorchenkova [ID]**, Oleksii Yehorchenkov** [ID]**, and Anton Sazonov** [ID]

Abstract In the paper, the authors propose to use a process approach and machine learning methods, namely Logical Regression, Random Forest, and XGboost algorithm, for project selection in the portfolio in banking. The paper's objective is an improvement of the PDCA model using of machine learning method for the project portfolio management in banking. Tasks of the paper: to develop a model of the PDCA approach for project portfolio management in banking; to check the different machine learning algorithms for the project portfolio dataset of a real bank; to evaluate and select the best algorithm. As a result, there was concluded, the construction of the cyclical PDCA process using project and process management technologies and machine learning methods in symbiosis can ensure effective project and program portfolio management, as the process will be controlled with a constant cycle of improvement and minimal human impact.

Keywords Logistic regression · Random forest · XGBoost · Machine learning · Project portfolio management · Banking · Process approach · PDCA

1 Introduction

1.1 Problem Statement

In the conditions of the developing economy of Ukraine, business entities are in a tough competitive environment and are comprehensively focused on the struggle for the consumer and long-term development. Such factors necessitate the adoption of quality and rapid decisions in the management, adoption, and implementation of necessary changes, targeting new consumers and markets, creating new offers to

N. Yehorchenkova (✉) · A. Sazonov
Kyiv National University of Construction and Architecture, Povitroflotsky Ave. 31, Kyiv 03037, Ukraine
e-mail: realnata@ukr.net

O. Yehorchenkov
Taras Shevchenko National University of Kyiv, Volodymyrska Street, 64/13, Kyiv 01601, Ukraine

meet existing and create new demand for goods and services. Today's companies understand that the speed of effective change is the key to success and long-term sustainability. It is important to note that project management covers almost all sectors of the economy, as project management tools have a very wide range of applications. One of such sectors is the financial sector of the economy in which, in particular, the implementation of projects plays an extremely important role. CEOs of financial institutions are well aware that the implementation of projects is not a whim but a forced affair. The high level of competition, the developing market, and the increase in the demand for lending among individuals and legal entities intensify the competition for consumers of financial services. Depending on the Strategy, projects are implemented to optimize resources (implementation of Lean Production) or increase competitive advantages (i.e. those that allow to meet the requirements of the existing market). The National Bank of Ukraine, as a regulator of the financial sector, implements a culture of project management, setting an example to other state control and private banks in Ukraine. Looking at public information from open sources about the achievements of the banking system of Ukraine on the way to the implementation of project management, you can find different approaches, methodologies, and concepts.

In practice, project management, at different levels of the banking sector solves different problems. It might be a one-time solution to existing problems, or the way of existence (implementation of projects in construction, marketing, etc.), or it is a system function, the task of which is achieving the strategic prices of the bank or reducing costs. Of course, if the bank implements projects in its activities (both external and internal), project managers and functional managers are faced with some issues, such as:

- unavailability of resources (material or labor);
- excess of projects number in the portfolio;
- delays in project implementation;
- inability to achieve the desired results, goals, objectives;
- failure at the linear level of problem-solving, and others.

In the paper for solving defined problems, the authors propose to develop a PDCA process approach model using machine learning methods (MLM) for project portfolio management in banking.

1.2   Process Approach for Banking

Theoretical and practical issues of the process approach to management have been studied by such foreign scientists as E. Deming, M. Hammer, D. Champy, D. Harrington, R. Susa, M. Robson, F. Ullah, and others, among domestic researchers— L. Balabanova, V. Baranovsky, O. Belarus, O. Vasyurenko, O. Vovchak, I. Bushuyeva, A. Kozachenko, and others. It is worth noting that process management involves continuous improvement of processes.

Among the methods of process improvement are three main conceptual approaches to improving the efficiency of the organization in the process management:

1. insignificant improvements to the process—improvement/change of the sequence of participants and executors of individual operations, procedures, or stages;
2. gradual improvement of processes (optimization)—requires certain re-source costs and constant improvement of the process, the founder of such a model is E. Deming, which he described in his work "Modern Time Management" (PDCA cycle) [1];
3. redesign (reengineering) of processes, which leads to significant changes in established management functions, comprehensive and radical transformation of the entire business. M. Hammer and D. Champy. in his work "Faster, better, cheaper: Nine methods of business process reengineering" in 2016 [2] describe reengineering as follows: "Reengineering is a fundamental rethinking and radical redesign of business processes to achieve significant improvements in such key performance indicators for modern business as cost, quality, and efficiency".

For project portfolio management in banking, it is better to use the second approach—PDCA cycle, because processes of this approach allow to continuously improve the portfolio at the expense of regular revision of executing projects, initiating new and suspending the worst projects.

1.3 PDCA Approach

The PDCA cycle of continual improvement should answer four key questions: What should it be like? What should we do and how should we do it? What was achieved? What else is there to be done? (Fig. 1).

PLAN (P)—What should it be like?

Each process cycle starts with defining objectives and action planning. The success of the whole cycle mostly depends on this stage. The reason for that is clear. If objectives are not properly defined, and particularly, if the plan was not set properly, it is difficult to expect that the complete cycle will function properly. The plan needs to have clearly defined operational rules and, to be comprehensive and besides planned activities, it has to be able to predict unexpected activities and problems.

DO (D)—What should we do and how should we do it?

Implementation of actions is crucial for this stage. The set plan is developed and worked out during this stage. In production, this phase means production and supply, and in-services—it means realization of services. According to the current standard, the implementation of services is defined as the implementation of products.
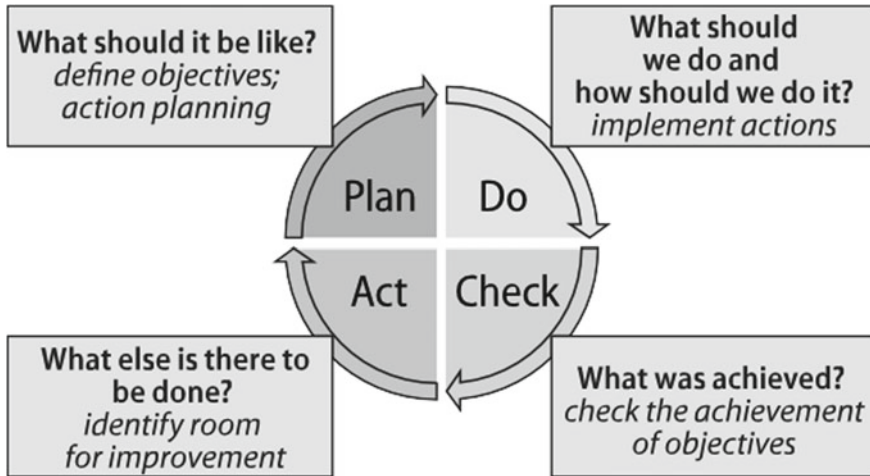
Fig. 1 PDCA cycle of continual improvement of processes

CHECK (C)—What was achieved?

The primary task of this stage is to check the achievement of objectives. It studies the contributions of changes and examines all factors affecting the understanding of the problem. For comprehensiveness and validity of initial assumptions and the comprehensiveness of the problem study, this stage applies specific methods and procedures such as statistical methods. In the production conditions, this stage cycle has the role of studying customers' reactions to the new product or service.

ACT (A)—What else is there to be done?

Identifying room for improvement is embedded in the last but very important stage. If previous stages have given good results, the implementation that has been planned, developed, and studied is implemented.

One complete rotation of the circle means the complete cycle of improvement and/or quality improvement. If the performed cycle has not given expected results, the news cycle should be started with the new plan, new information, and applying knowledge gained during the previous cycles.

The PDCA is a universal approach—suitable for all domains (production, education, health, financial institutions, etc.). It has the main attributes of the quality management philosophy and the scientific approach to understanding and improving the process. PDCA cycles are continual processes of all working groups and they refer to all processes in the organization. If applied, this approach indicated the management's commitment to its main role:

- in continual improvements. The PDCA cycle means continual training and distribution of knowledge. It means that one completed cycle encourages starting a new

one with a new plan, new information, and applying knowledge gained during the previous cycles. Every next cycle has improved plan and process;

- activities, improved checks and analyses, and other new improvements. This methodology is a dynamic process with the role of catalyst in process improvement.

The application of this principle and the PDCA methodology enables central banks to continually re-examine the manner of functioning and improving their performances. The only seeming limitation in terms of central bank objectives may be that the central bank does not define them individually since they are primarily defined by the law. However, in no way does this limit the central bank to continually re-examine the implementation of set objectives and to improve its processes and complete performances. The application of this principle records the following benefits: determining processes and their improvement, better budgeting, removing deficiencies from previous period, enhanced credibility, more quality inputs for deciding, and the like (ISO 2006) [3].

Thus, the use of PDCA ensures the constant development of processes and meets the needs of a rapidly changing market. And any activity that has a set of regular, interconnected operations, procedures, actions performed by the owner and/or participants of the process to create a product or service that has value for the consumer (customer) both external and internal can be defined as a process. Thus, it can be described, controlled, and constantly improved. The PDCA principle can be used as the main method that needs improvement to develop a cycle of project portfolio management process for its effective management.

As far as the most important goal of project portfolio management for the banking is implementing the strategy, the project portfolio management in the PDCA model has to meet the following conditions:

1. to provide the implementation of the strategy;
2. to provide effective interaction of the bank departments when updating the project portfolio;
3. to provide information and report to stakeholders on the status of the project portfolio implementation.

1.4 Paper's Purpose and Objectives

The paper's purpose is the improvement of the PDCA model using of machine learning method for the project portfolio management in banking.

Objectives of the paper:

- to develop a model of the PDCA approach for project portfolio management in banking;
- to check the different machine learning algorithms for project portfolio dataset of a real bank;
- to evaluate and select the best algorithm.
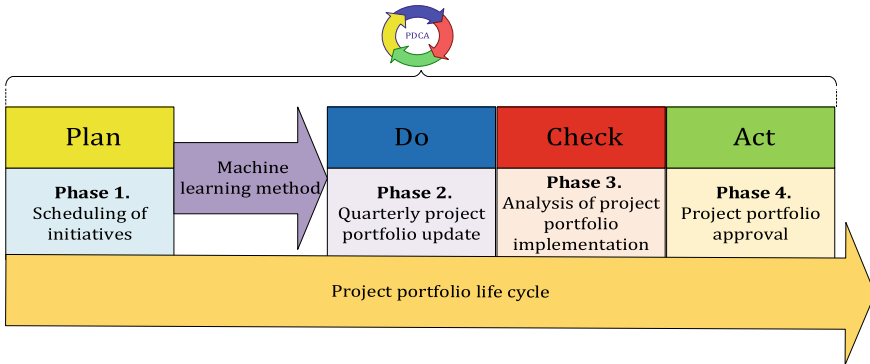
Fig. 2 PDCA model for achieving strategic goals and initiatives

2 A PDCA Model for Project Portfolio Management in Banking

2.1 Model of a PDCA Approach for Project Portfolio Management in Banking

The PDCA can be applied to a specific set of project portfolio components. In the paper, it is proposed an updated model of the project portfolio life cycle, in which the phases will be clearly defined in time constraints and consistent in their implementation, and will last 1 calendar year in the total. The general model that will be used in the research on achieving strategic goals and benefits is presented in Fig. 2 "PDCA model for achieving strategic goals and initiatives".

1. **Phase 1 "Scheduling of initiatives"**—defines and establishes the sequence of works at which a strategic plan of project initiatives implementation is created.
2. **Machine learning method**—methods of decision support for the project included in the portfolio based on the developed models and algorithms.
3. **Phase 2. "Quarterly project portfolio update"**—provides interaction and sequence of works at which the initiatives identified in the strategic plan are initiated and selected for further authorization in the project portfolio. This phase has the following sequence of works, with a total duration of 3 calendar months:
4. **Phase 3. "Analysis of project portfolio implementation"**—provides an analysis of the annual performance of the portfolio and provides answers to questions about the extent to which the bank was able to achieve strategic goals through the implementation of projects.
5. **Phase 4. "Project portfolio approval"** is responsible for the strategic adjustment of the portfolio, the applying of the necessary administrative measures to improve/correct deviations and approval (Fig. 3).

3.1. Planning of potential components
provides identification, categorization, evaluation, selection, prioritization and balancing of the project portfolio potential components.

3.2. Authorization of potential components
provides coordination and authorization of project portfolio components

3.3. Monitoring and reporting
provides monitoring and reporting during the project portfolio implementation

3.4. Update and approval
provides updating and approval of the portfolio as of the new quarter.

Fig. 3 Phases of PDCA model

For now, the decision-making process is carried out by the management of a bank. But criteria of these decisions often are subject to the influence of the human factor. So not every decision will be loyal and effective enough. So there is necessary to use the machine learning method for the formalization of the decision-making process.

3 Machine Learning Method

To develop a model of decision-making processes in project initiating in portfolio an experiment was conducted. For this experiment, a dataset with real bank projects was chosen. The dataset has 17 columns (features) and 155 raws. The issue with creating the models was that we have a small dataset. Therefore, to avoid overfitting, we used relatively simple models. In logistic regression, we controlled the regularization using the C parameters (inverse of regularization strength—smaller values specify stronger regularization) to deal with the overfitting. For tree-based models like XGBoost, we controlled the overfitting by tuning a series of parameters:

- Restricting the maximum depth of trees via max_depth (low values).
- Making the model more conservative via gamma and eta (high values).

- L1 and L2 regularization via reg_alpha and reg_lambda (high values) [4].

For the experiment was used the programming language Python and environment Jupiter Notebook.

Figure 4 is shown the logical structure of the experiment.

1. **Data reading and evaluating**

Table 1 is presented a fragment of the row dataset.
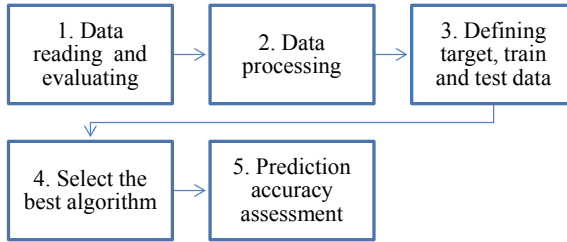
Fig. 4 The logical structure of the experiment

Table 1 Fragment of the dataset (transformed table)

No	1	2	3	…
Group	UB	SP	OB	…
Initiative code	UB_0_008	SP_1_004	OB_1_009	…
Category	Strat	Usual	Usual	…
Budget <5 Mln	1	0	0	…
Budget Mln	7,18	2,55	3,68	…
Date start plan	17.06.2016	01.07.2015	28.07.2015	…
Date end plan	28.12.2018	30.08.2016	30.04.2019	…
Duration plan	30	13	45	…
Duration >30	1	0	1	…
Compliance strategy	1	0	1	…
Authorization	1	1	1	…
Finish fact	1	1	1	…
Requirements law	1	0	1	…
PM name	William T	Thomas M	Scarlett Y	…
Sponsor name	Bill E	Jerry H	Gordon D	…

2. Data clearing and processing

Preparing dataset to learning and development of machine learning model dataset was processed. There were removed unnecessary features: '№', 'Initiative code', 'Authorization. For categorical data were used One Hot Encoding.

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 "Authorization". This feature shows the status of initiating projects: 0—a project was rejected (18 projects), 1—a project was approved (137 projects) (Fig. 5).

Table 2 Fragment of processed dataset (transformed table)

Group	UB	SP	OB	…
Category	Strat	Usual	Usual	…
Budget <5 Mln	1	0	0	…
Budget Mln	7,18	2,55	3,68	…
Date start plan	17.06.2016	01.07.2015	28.07.2015	…
Date end plan	28.12.2018	30.08.2016	30.04.2019	…
Duration plan	30	13	45	…
Duration >30	1	0	1	…
Compliance strategy	1	0	1	…
Finish fact	1	1	1	…
Requirements law	1	0	1	…
PM name	William T	Thomas M	Scarlett Y	…
Sponsor name	Bill E	Jerry H	Gordon D	…

Note full table consists of 13 rows × 124 columns
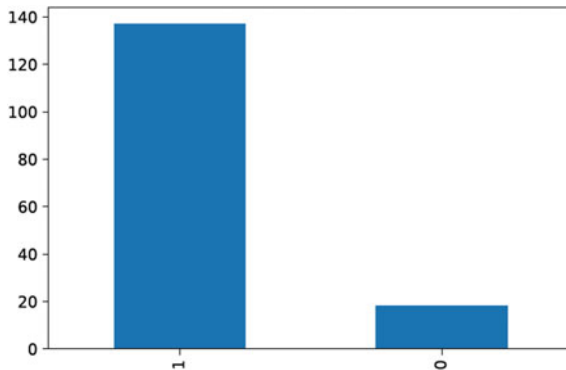
Fig. 5 Approved and reject projects

The learning models and evaluation of its accuracy dataset were split in the ratio of 80% train to 20% test data.

4.   Select the best algorithm

For development machine learning model for selecting project in portfolio in banking, there are necessary to check several algorithms on the dataset and choose the most suitable one. Authors considered the next algorithms: Logistic Regression, Random Forest, and XGBoost. Let's consider every algorithm in more detail.

4.1   *Logistic Regression*

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist [5]. Logistic Regression is one of the most simple and commonly used Machine Learning algorithms for two-class classification. It is easy to implement and can be used as the baseline for any binary classification problem. Its basic fundamental concepts are also constructive in deep learning. Logistic regression describes and estimates the relationship between one dependent binary variable and independent variables. [6].

An algorithm of logistic regression learning is presented in Fig. 6.

There were received the next parameters:

- ROC AUC mean—0.9050505050505052;
- Standard deviation—0.115440656564388.

The scores parameter is shown in Fig. 7.

4.2   *Random Forest*
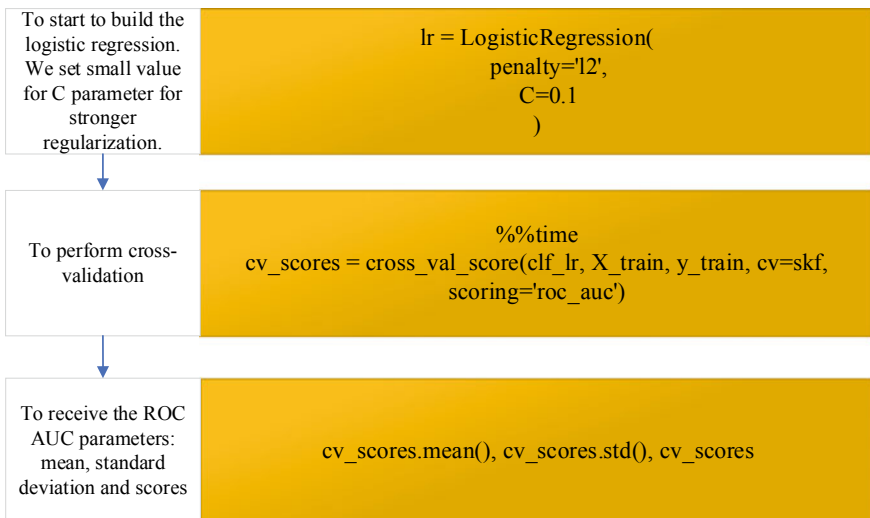
Fig. 6 An algorithm of logistic regression
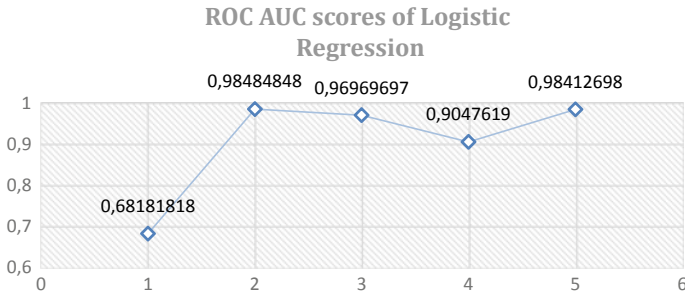
Fig. 7 Logistic regression ROC AUC scores

Random forests or random decision forests are an ensemble learning method for classification, regression, and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees [7].

An algorithm of Random Forest learning is presented in Fig. 8.

There were received the next parameters for Random Forest:

- ROC AUC mean—0.9054834054834056;
- Standard deviation—0.09226411754916562.

The scores parameter is shown in Fig. 9.

4.3 XGBoost

Fig. 8 An algorithm of random forest
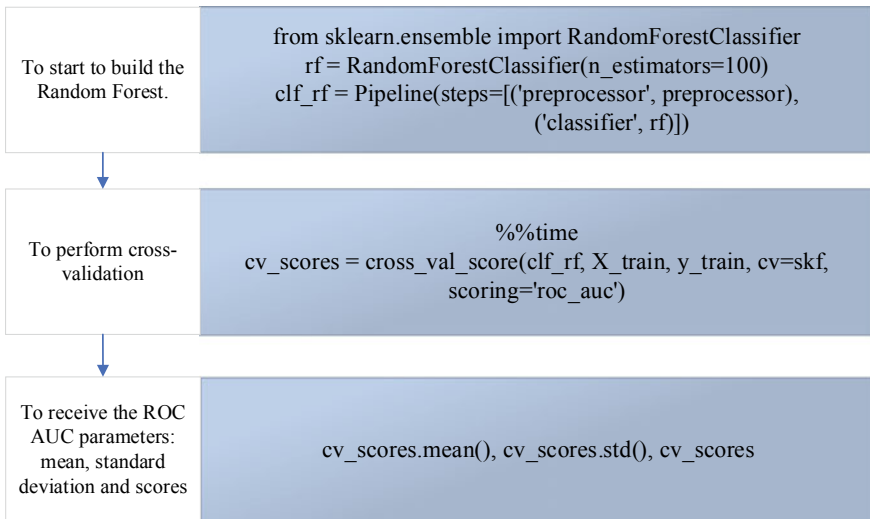
ROC AUC scores of Random Forest
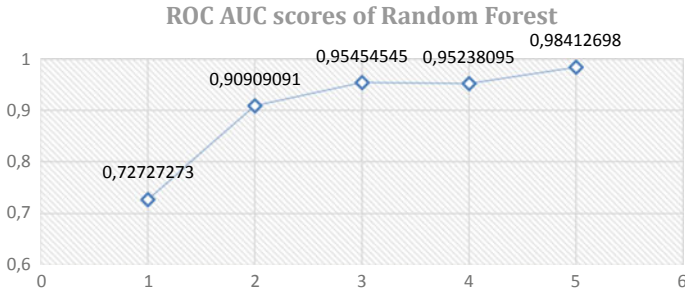
Fig. 9 Random forest ROC AUC scores

XGBoost is an open-source software library that provides a gradient boosting framework for C++, Java, Python, R, Julia, Perl, and Scala. It works on Linux, Windows, and macOS. From the project description, it aims to provide a "Scalable, Portable and Distributed Gradient Boosting (GBM, GBRT, GBDT) Library". It runs on a single machine, as well as the distributed processing frameworks Apache Hadoop, Apache Spark, and Apache Flink. It has gained much popularity and attention recently as the algorithm of choice for many winning teams of machine learning competitions [8].

An algorithm of XGBoost learning is presented in Fig. 10.

There were received the next parameters for XGBoost:
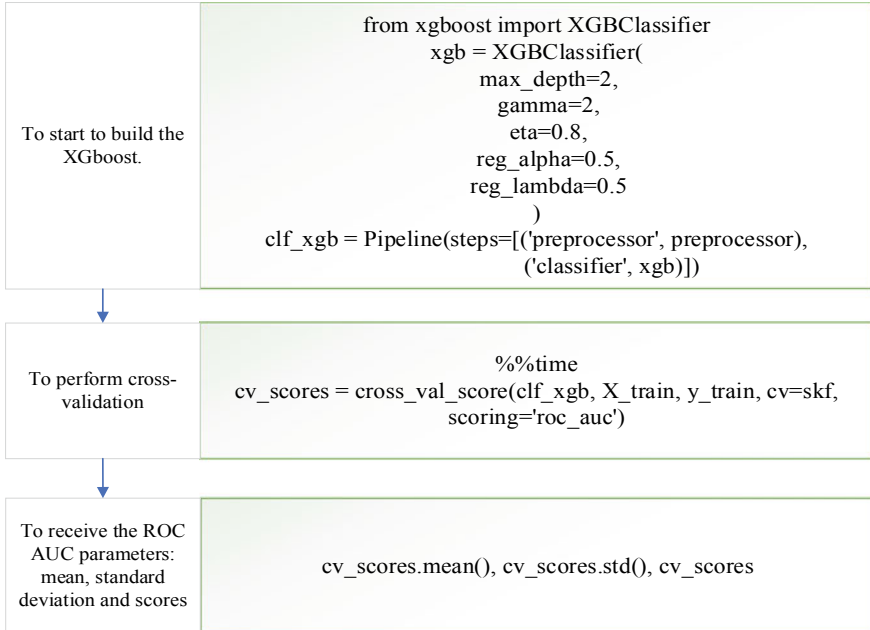
Fig. 10 An algorithm of XGBoost
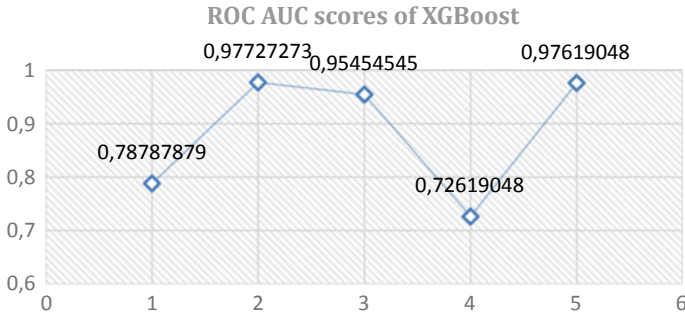
Fig. 11 XGBoost ROC AUC scores

- ROC AUC mean—0.8844155844155844;
- Standard deviation—0.10612993581714016.

The scores parameter is shown in Fig. 11.

5. **Prediction accuracy assessment**

As far as we have two suitable algorithms with good performance (mean accuracy and standard deviation) Random Forest and XGBoost, the authors developed two learning models and compared the results.

In the process of XGBoost construction was defined prediction accuracy assessment by ROC AUC metric (0.9655172413793103). According to this metric, the accuracy of the XGBoost is 96%. In the Random Forest model, there was received prediction accuracy assessment by ROC AUC metric (0.7327586206896552). So, the accuracy of the Random Forest is 73%. Therefore XGBoost algorithm is the best solution for the decision-making process of project initialization in the portfolio of banking.

4 Conclusion

In the paper, the authors propose to consider using of process approach and machine learning method for the task of project selection to a portfolio in banking. It was improved the PDCA model and developed models by algorithms Logical Regression, Random Forest, and XGBoost and evaluate these models through an experiment for a project portfolio dataset of a bank. In the result of the experiment, it was defined that the best algorithm is XGBoost, and the prediction accuracy for the project portfolio dataset is 98%. Such research allows us to consider using machine learning approaches for project portfolio management in banking.

Thus, the construction of a cyclical PDCA process using project and process management technologies and machine learning methods in symbiosis can ensure

effective project and program portfolio management, as the process will be controlled with a constant cycle of improvement and minimal human impact.

References

1. Deming E (2019) New age management: simple mechanisms leading to growth, innovation and market dominance. Alpina Publisher, Moskow
2. Hammer M, Hershman L (2016) Faster, better, cheaper: nine methods of business process reengineering, 2nd edn. Alpina Publisher, Moskow
3. Luburić R (2015) Quality management principles and benefits of their implementation in central banks. J Cent Bank Theory Pract 91–121
4. Dealing with very small dataset. https://www.kaggle.com/rafjaa/dealing-with-very-small-datasets. Last Accessed on 3 March 2021
5. Logistic regression. https://en.wikipedia.org/wiki/Logistic_regression. Last Accessed on 3 March 2021
6. Understanding logistic regression Python. https://www.datacamp.com/community/tutorials/understanding-logistic-regression-python. Last Accessed on 3 March 2021
7. Random Forest. https://en.wikipedia.org/wiki/Random_forest. Last Accessed on 3 March 2021
8. XGBoost. https://en.wikipedia.org/wiki/XGBoost. Last Accessed on 3 March 2021

Author Index