

Alla G. Kravets
Alexander A. Bolshakov
Maxim Shcherbakov *Editors*

Society 5.0: Cyberspace for Advanced Human-Centered Society

Studies in Systems, Decision and Control

Volume 333

Series Editor

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

The series “Studies in Systems, Decision and Control” (SSDC) covers both new developments and advances, as well as the state of the art, in the various areas of broadly perceived systems, decision making and control—quickly, up to date and with a high quality. The intent is to cover the theory, applications, and perspectives on the state of the art and future developments relevant to systems, decision making, control, complex processes and related areas, as embedded in the fields of engineering, computer science, physics, economics, social and life sciences, as well as the paradigms and methodologies behind them. The series contains monographs, textbooks, lecture notes and edited volumes in systems, decision making and control 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.

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

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

More information about this series at <http://www.springer.com/series/13304>

Alla G. Kravets · Alexander A. Bolshakov ·
Maxim Shcherbakov
Editors

Society 5.0: Cyberspace for Advanced Human-Centered Society

Editors

Alla G. Kravets
Volgograd State Technical University
Volgograd, Russia

Alexander A. Bolshakov
Peter the Great St. Petersburg Polytechnic
University
St. Petersburg, Russia

Maxim Shcherbakov 
Volgograd State Technical University
Volgograd, Russia

ISSN 2198-4182 ISSN 2198-4190 (electronic)
Studies in Systems, Decision and Control
ISBN 978-3-030-63562-6 ISBN 978-3-030-63563-3 (eBook)
<https://doi.org/10.1007/978-3-030-63563-3>

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

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

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

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

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

Preface

Artificial intelligence and smart technologies change the landscape of common-used solutions. They have a great impact on human being and society in general. Despite on continents and countries, society is being changed, so there are new processes needs re-evaluation and discovering new approaches for understanding, modelling and management.

Society 5.0 put a human in a centre of economic and social activities via technologies based on highly integrated cyberspace and physical space. This new *Society 5.0* extends the information society paradigm. However, a simple extension and increasing complexity of information systems are not enough. That is why many questions remain open, the questions covering such domains as human-centred processes, ecology safety, the comfort of the environment in regions and cities, health care and medicine for life longevity. It required new ideas and solutions in computer science and artificial intelligence application for development society towards *Society 5.0*.

The book includes 21 chapters covering such application areas of cyber-physical systems in society as ecology, environmental issues, medicine and health care. All chapters are joined in four sections: human-centred society, smart cities and regions, smart technology for ecology and smart technology for health care.

This book is directed to researchers, practitioners, engineers, software developers, professor and students. We do hope the book will be useful for them.

Edition of the book is dedicated to the 130th Anniversary of Kazan National Research Technological University and technically supported by the Project Laboratory of Cyber-Physical Systems of Volgograd State Technical University. The book was prepared with the financial support of the Russian Foundation for Basic Research, project No. 20-08-20032.

Volgograd, Russia
St. Petersburg, Russia
Volgograd, Russia
September 2020

Alla G. Kravets
Alexander A. Bolshakov
Maxim Shcherbakov

Contents

Society 5.0: Human-Centred Society

Cyber-Social System as a Model of Narrative Management	3
Aleksandr Davtian, Olga Shabalina, Natalia Sadovnikova, and Danila Parygin	

On the Problem of Computability of Bounded Rationality Cognitive Solutions	15
Vladimir Zaborovskij, Vladimir Polyanskiy, and Sergey Popov	

Image Processing for Biometric Scanning of the Palm Vein Pattern . . .	25
Lina Kh. Safiullina and Rustem R. Maturov	

Automation of Demand Planning for IT Specialists Based on Ontological Modelling	35
Denis V. Yarullin, Rustam A. Faizrahmanov, and Polina Y. Fominykh	

Numerical Modeling of Business Processes Using the Apparatus of GERT Networks	47
Mikhail Dorrer, Alexandra Dorrer, and Anton Zyryanov	

Models and Methods of Forecasting and Tasks Distribution by Performers in Electronic Document Management Systems	57
Sofia S. Kildeeva, Alexey S. Katasev, and Nafis G. Talipov	

Society 5.0: Smart Cities and Regions

Smart City: Cyber-Physical Systems Modeling Features	75
Nikolai A. Fomin, Roman V. Meshcheryakov, Andrey Y. Iskhakov, and Yuri Y. Gromov	

Controlling Traffic Flows in Intelligent Transportation System	91
Alexander Galkin and Anton Sysoev	

Development of Scenarios for Modeling the Behavior of People in an Urban Environment	103
Alexander Anokhin, Sergey Burov, Danila Parygin, Vyacheslav Rent, Natalia Sadovnikova, and Alexey Finogeev	
Forecast of the Impact of Human Resources on the Effectiveness of the Petrochemical Cyber-Physical Cluster of the Samara Region ...	115
Natalya Baykina, Pavel Golovanov, Mikhail Livshits, and Elena Tuponosova	
Mathematical Model of Integration of Cyber-Physical Systems for Solving Problems of Increasing the Competitiveness of the Regions of the Russian Federation	129
Alexander Bolshakov, Irina Veshneva, and Dmitry Lushin	
Regional Competitiveness Research Based on Digital Models Using Kolmogorov-Chapman Equations	141
Irina Veshneva, Galina Chernyshova, and Alexander Bolshakov	
Society 5.0: Smart Technology for Ecology	
Transformation of Reduced Compounds of Carbon and Nitrogen Oxides in the Cramped Aerodynamic Conditions	157
Vadim A. Zaytsev	
Numerical Modeling of Pneumatic Conveying in the Mode of the Inhibited Dense Layer	169
Asia. G. Mukhametzyanova, Andrei O. Pankov, and Alina A. Abdrakhmanova	
Intelligent System for Determining the Presence of Falsification in Meat Products Based on Histological Methods	179
Alexander Bolshakov, Marina Nikitina, and Renata Kalimullina	
Society 5.0: Healthcare Smart Technology	
Probability-Entropy Model of Multidimensional Risk as a Tool for Population Health Research	205
Alexander N. Tyrsin, Dmitriy A. Yashin, and Alfiya A. Surina	
Generative Models Based on VAE and GAN for New Medical Data Synthesis	217
Vladislav V. Laptev, Olga M. Gerget, and Nataliia A. Markova	
Inhibitors Selection to Influenza Virus A by Method of Blocking Intermolecular Interaction	227
L. I. Zharkikh, Yu. A. Smirnova, I. M. Azhmukhamedov, E. V. Golubkina, and M. N. Trizno	

**The Stochastic and Singular Analysis of Fractal Signals
in Cyber-Physical Systems of Biomedicine 239**
Vladimir Kulikov, Alexander Kulikov, and Alexander Ignatyev

**Quality Research of the Interval Cetlin Method as a Component
of the Cyber-Physical System of Continuous Monitoring
of the Human-Operator State by ECG Signals 253**
Alexey Khalaydzhi

**Technology for Predictive Monitoring of the Performance
of Cyber-Physical System Operators Under Noise Conditions 269**
Igor Ushakov, Alexey Bogomolov, Sergey Dragan, and Sergey Soldatov

Society 5.0: Human-Centred Society

Cyber-Social System as a Model of Narrative Management



Aleksandr Davtian , Olga Shabalina , Natalia Sadovnikova ,
and Danila Parygin 

Abstract The chapter considers approaches to modeling goal setting and goal achievement as the main stages of the management process in socio-economic systems in the context of their natural evolution into cyber-social systems. A model of narrative management in a cyber-social system, in which the choice of management strategies is considered as a narrative practice, implemented through the duality of goal setting and goal achievement, is proposed. The goal model is represented as a dynamic structurally ordered space of goals, reflecting the conditions for the availability of goals, determined by the state of the managed system. The methods of universal algebra are used to model the space of goals, which makes it possible to form and dynamically modify the space, and to calculate the availability of goals and the dynamics of their achievement in space in a given structural connection. Responsibility for making managerial decisions to achieve goals remains with the person. The results of his actions, leading to the impossibility of achieving the goal, are compensated by modifying the goal space without losing its structural coherence. The architecture of the software complex for management support in cyber-social systems, which implements the proposed model, is presented. The mathematical methods used in the system to manage support are described.

Keywords Socio-economic system · Management in socio-economic system · Cyber-social system · Goal setting · Goal achievement · Goal model · Goal space · Dynamic goal setting · Narrative management

A. Davtian

Moscow Institute of Physics and Technology, Institutsky Ln. 9, 141701 Dolgoprudny, Russia

e-mail: agvs@mail.ru

O. Shabalina (✉) · N. Sadovnikova · D. Parygin

Volgograd State Technical University, Lenina Av. 28, 400005 Volgograd, Russia

e-mail: o.a.shabalina@gmail.com

N. Sadovnikova

e-mail: npsn1@ya.ru

D. Parygin

e-mail: dparygin@gmail.com

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_1

1 Introduction

The well-being of society to a large extent depends on the activities of socio-economic systems (SES), organized by the society itself to meet its needs. It is required to manage the processes of functioning of the system to achieve the results of the SES activities demanded by society.

The key stages in the process of complex socio-economic systems management are the establishment and structuring of management goals and the development of strategies to achieve them. Therefore, modern research in the management of socio-economic systems is associated with the development of effective models and methods of goal setting and goal achievement and mechanisms for their implementation in the conditions of modern society.

The technical, and after it the information explosion that took place in society, created a total world information space. Comprehensive changes in society caused by the influence of the information space, in which each SES is inevitably immersed, led to a qualitative change in the role of information in management tasks in such systems. It is necessary to rethink the processes of goal setting and goal achievement, as the main components in the SES management system, in a new context and the development of new approaches to modeling these processes under the new conditions of their functioning, in the context of the natural evolution of SES immersed in the information space.

2 Modeling Goal Setting and Goal Achievement in a Socio-economic System

It is customary to consider goal setting as a management function implemented at the first stage of organization the system's activities and aimed at achievement the general goal of management [1], under the classical principles of management theory. The process of goal-setting includes several stages associated with the analysis of the general goal of system management; detailing the goal according to the selected characteristics; the choice of qualitative and quantitative indicators to assess the achievement of goals and their relationships, etc.

The well-known models of goal setting (SMART (Specific, Measurable, Achievable, Relevant, Time-bound), SPIRO (Specificity, Performance, Observability)), etc. [2–4]. Are structured verbal descriptions of the principles of forming the goals of the system functioning, determined by the general goal. It is customary to present the results of goal setting in the form of goal models, reflecting the composition of goals, characteristics, and relationships, the form of which is determined by the type and purpose of the model [5]. The main structural model of goals is the goal tree [6, 7], which is a decomposition of the general goal set at the top level of the system into goals following the organizational structure of the system, taking into account

its purpose, scale, specifics and other characteristics. Graph models of goals are also used to synthesize goal models [8].

Strategies for achieving the goals determined by the goal model are developed at the next stage of the organization the system's activities. This implies that in the process of system development, the formed structure of goals does not change and the management task is formulated as the task of organizing activities to ensure the achievement of goals. At this stage, the state of the system is analyzed, its real capabilities are assessed, the tasks are formed, the solution of which should lead to the achievement of the corresponding goals, the ways of the most rational use of resources to improve the efficiency of the system are considered, etc.

Various quantitative and qualitative modeling methods are used to model goal achievement. Graph-theoretic models are widely used in the study of properties and analysis of complex systems [9, 10]. The nodes of the graph, depending on the purpose of modeling, can be system components; processes that determine the behavior of the system; various factors affecting the dynamics of system development, etc. Arcs of the graph can be informational, control, technological, and other connections between them.

The representation of a socio-economic system as an active system allows one to take into account the presence in such systems of a person who can conduct purposeful behavior following his preferences (interests) [11]. It is believed that the rational behavior of people corresponds to the choice of states (strategies) that would maximize their objective functions. In particular, game-theoretic models and the corresponding mathematical methods are used to describe the behavior of systems in conflict, possibly uncertain situations [12, 13].

Simulation methods and methods of qualitative modeling are also used to display the properties of dynamically changing systems [14–20]. The behavior of the system in such models is represented as a change in its state in the space of possible states. Management of the system is the choice of the trajectory of its development as an ordered set of states, leading the system from the initial to a given target state, unchanged during the development of the system. In general, the considered models of goal achievement are based on the assumption of the finiteness and fundamental achievability of the goal and the a priori existence of strategies for achieving it.

The considered goal achievement models are focused on the execution of projects completed in the time that determine the achievement of local goals. Goal achievement in this case is a time-repeating process of achieving goals within the framework of a developed model of goals, which is a general goal. The process of achieving the general goal itself is customarily correlated or even identified with the mission of the socio-economic system, determined by the needs of society [21].

Thus, goal setting and goal-achievement are viewed as two successive stages of managing a socio-economic system under the classical principles of the management theory. Accordingly, the models of the system management process include independent models of goal setting and goal achievement, designed to support the corresponding stages of management (see Fig. 1).

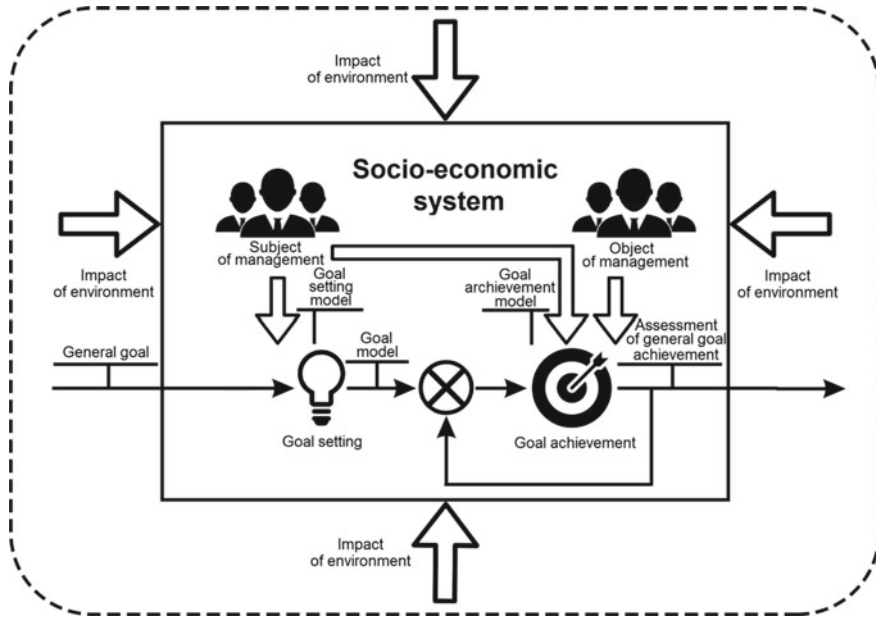


Fig. 1 Goal setting and goal achievement in a socio-economic system

3 Integration of Cybernetic Principles into Management Processes in SES

The organization of management in SES is always associated with the collection and processing of data on the current state of the system and its environment. The trade-off between the amount of data stored on paper and the ability of a person to process this data and make management decisions based on their analysis was maintained in the pre-information era. The possibilities of modern information technologies have led to a manifold increase in the amount of data available to humans. Today, the intensity of information flows significantly exceeds the ability of a person to select and analyze the data necessary to ensure the sustainable development of the system from the information noise surrounding him.

The successful organization of SES management in such conditions is only possible with the availability of computing resources and technologies that allow continuous collection, processing, and analysis of big data necessary to understand the system processes in SES. At the same time, computing resources should be integrated into the SES on the principles of cybernetics, which determine the methods of managing information resources in the system, and the management of information flows should be subordinate to the goals of the SES itself. Thus, an organized SES with a cybernetic “insert” into its structure can be defined as a cyber-social system (CSS).

The concept of a cyber-physical system has appeared relatively recently, and an unambiguous understanding of this term has not yet been achieved. Thus, a cyber-social system is modeled as a cyber-network of agents monitoring the state of people in a social network, following [24]. Moreover, each node of the cyber network represents an agent, and links represent the exchange of information between agents. The developed model is designed to solve NP-complex optimization problems in social networks. In [25], the category of cyber-social systems is considered as a result of the interaction of SES with cyber technologies, due to fundamental transformations in a wide range of human activities in the context of digital technologies. In [26, 27], cyber-social systems are allocated to a new class of SES as a result of the integration of human subjects into cyber-physical systems and simulates such systems as a combination of models of socio-technical systems and models of human behavior.

Thus, a new category of socio-economic systems (cyber-social systems) is designated, in which a continuous increase in spontaneous flows of information leads to the gradual replacement of the real system with its virtual representation in cyberspace. And the management of the system is replaced by the management of its information image based on the principles of cybernetics.

4 Goal Setting as a Component of Goal Achievement in a Cyber-Social System

Not only the initially set goals but also the very goal orientations of the system in its environment can change under the conditions of continuously growing information flows that determine the current state of the SES and the dynamics of its development in conditions of interaction with the external environment. In turn, this can significantly affect the a priori ideas about the goals of management; their fundamental achievability and strategies for achievement; and possibly the need to achieve the goal (or goals) as such. Management of a system without taking into account the continuous exchange of system data with the environment, which forms information flows that determine the target state of the system itself, can lead to the loss of its manageability and, ultimately, the loss of the meaning of its existence as a carrier of the mission.

Taking into account the influence of information flows on the development of the SES stipulates the need to additionally define or redefine goals in the process of system development. In this case, goal setting implies the construction of an initial model of the system management goals and its dynamic updating (modification) by the actions carried out both by the system managing staff and executors, taking into account the current state of the system and its environment.

In this context, goal setting should be considered as a continuous process of achieving a general goal, which forms the managed development of the system at all stages of its life cycle. At the same time, the general goal itself generates an increasing multiplicity of goals, the achievement of which determines the success of

the implementation of the mission of the socio-economic system in society, i.e. the general goal is not related to the time of its achievement, but to the time of existence of the system itself. Goal achievement in this case means the activities of subjects and objects of management in the execution of projects that determine the achievement of the corresponding goals in the context of constantly growing information flows that determine the dynamic formation of the goal space.

The fulfillment of the mission as a unity of goal setting and goal achievement in SES determines the need for constant monitoring and analysis of information flows, which are carriers of information necessary for making operational and informed decisions. The integration of cybernetic principles, as general principles of deduction of the subsequent state of the system from the previous one, into the management of the SES, allows making management decisions based on the analysis of information flows that determine the dynamics of the interaction of the system with the environment and, therefore, to consider the SES as a CSS (see Fig. 2).

Failure to achieve this or that set goal as a result of KSS management does not mean that cybernetic principles do not work. Cybernetics itself is not responsible for the setting of goals. The task of cybernetics is the formation and management of information flows to achieve the set goal of system management. Management in SES is carried out by a person, and the ultimate goal of management is to ensure the

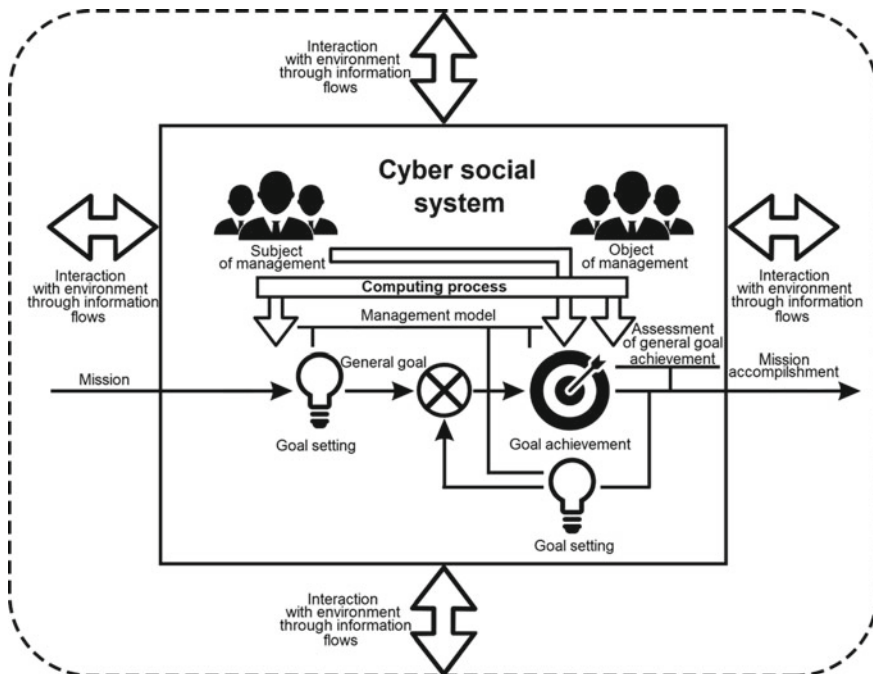


Fig. 2 Dynamic goal setting as a structural component of goal achievement in a cyber-social system

existence of the system and fulfill its mission, assigned by the environment of the system.

5 Model of Narrative Management in CSS

The key problem in organizing management in SES is the problem of forming strategies to achieve management goals. There is no way to determine whether or not the strategy chosen by the subject of management will lead to the achievement of the chosen goal in SES. A priori known “correct” strategies do not exist, but there is only a person’s belief in the achievability of the goal based on the strategy chosen by him, determined by his previous experience, a measure of responsibility and competence.

The management strategy chosen by a person determines not only the future state of the system but also the very existence of the system in the future. Thus, the future of the system is present, in the understanding of a person, in the present in the form of a narrative, i.e. “instructions for creating the future in the present” [12]. This understanding of the “structure” and behavior of the system allows applying a narrative approach to the organization of management of such systems, based on the subjectivity of any management strategy chosen by a person and his responsibility for its implementation.

The concept of narrative control in SES, proposed in [22], is based on the non-modeled competences and responsibilities of a person as a subject and object of management, and the organization of the management process itself on the principles of cybernetics. The model of goals is represented in the proposed concept as a dynamic structured-ordered space, reflecting the logic of goal achievement, determined by the dynamics of the state of the system in terms of its interaction with the environment through bidirectional information flows. The methods of universal algebra are used to model the space of goals, which allows forming and dynamically modify the space, and to calculate the availability of goals and the dynamics of their achievement in space in a given structural connection [23]. Goal achievement, in this case, is a process developing over time to achieve goals within the framework of the developed goal model.

The mission represents the inductive limit of the dynamically evolving goal space. The bearer of the mission is a person who is responsible for the development of a management strategy, while it is the person who has the “last word” in making management decisions. The business process of narrative management is shown in Fig. 3.

The proposed management model in the CSC, based on the duality of goal setting and goal achievement, implements the possibility of monitoring and maintaining the observability and controllability of the system during its operation, taking into account its possible states leading to changes in the goal space. In this case, the responsibility for the actions of the participants in the management process lies with the participants themselves. The results of their actions, leading to the impossibility

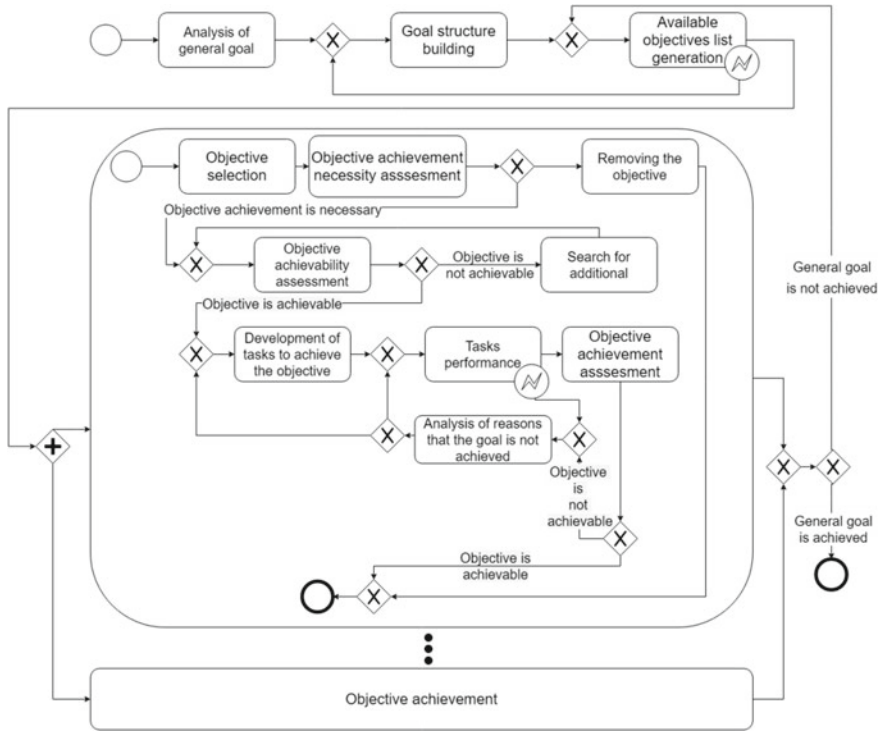


Fig. 3 Narrative-based management process (BPMN notation)

of achieving the goal, are compensated for by the goal space modification without losing its structural coherence.

6 Software Complex for Management Support in Cyber-Social Systems

The proposed models are implemented in a software complex for management support in cyber-social systems [28]. The behavior of people as subjects and objects of management cannot be modeled within the framework of the proposed models, since the responsibility for the implementation of management decisions remains with the person. Mathematical modeling is used to support those stages of management, within which models are applicable, that can be reasonably interpreted in terms of a real modeled system (see Table 1).

The architecture of the software complex is shown in Fig. 4.

Screen forms in the mode of forming a management strategy and determining the availability of current objectives are shown in Fig. 5.

Table 1 Mathematical methods of management support

Function of software complex	Mathematical methods
Development of management strategy	–
Building of goal space	Lattice theory methods
Goal space modification	Methods of universal algebras
Search of available objectives	Calculus on a topological lattice in idempotent algebra
Assessment the need to achieve the objective	Scenario forecasting methods
Assessment the attainability of the objective	Fuzzy hierarchical estimation methods
Matching tasks with objective	–
Execution of tasks	Network planning techniques
Objective achievement assessment	Methods for convolution of expert assessments Ranking methods for convolutions of numeric assessments

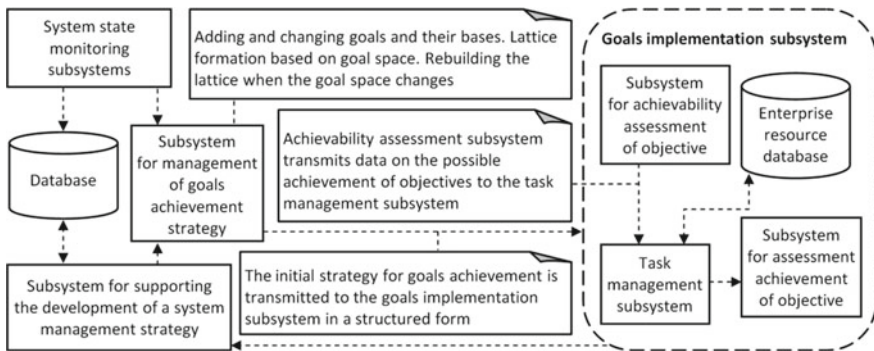


Fig. 4 Architecture of the software complex for supporting the dynamic goal-setting of SES

7 Conclusion

The model of narrative management, in which the bearer of the mission is a person who is responsible for management strategy development, is proposed in the work. At the same time, the person has the “last word” in making managerial decisions. Goal setting is a way of presenting the future in terms of the present, and goal achievement is actions performed in the present, determined by the future that has not yet taken place, in the context of the proposed model.

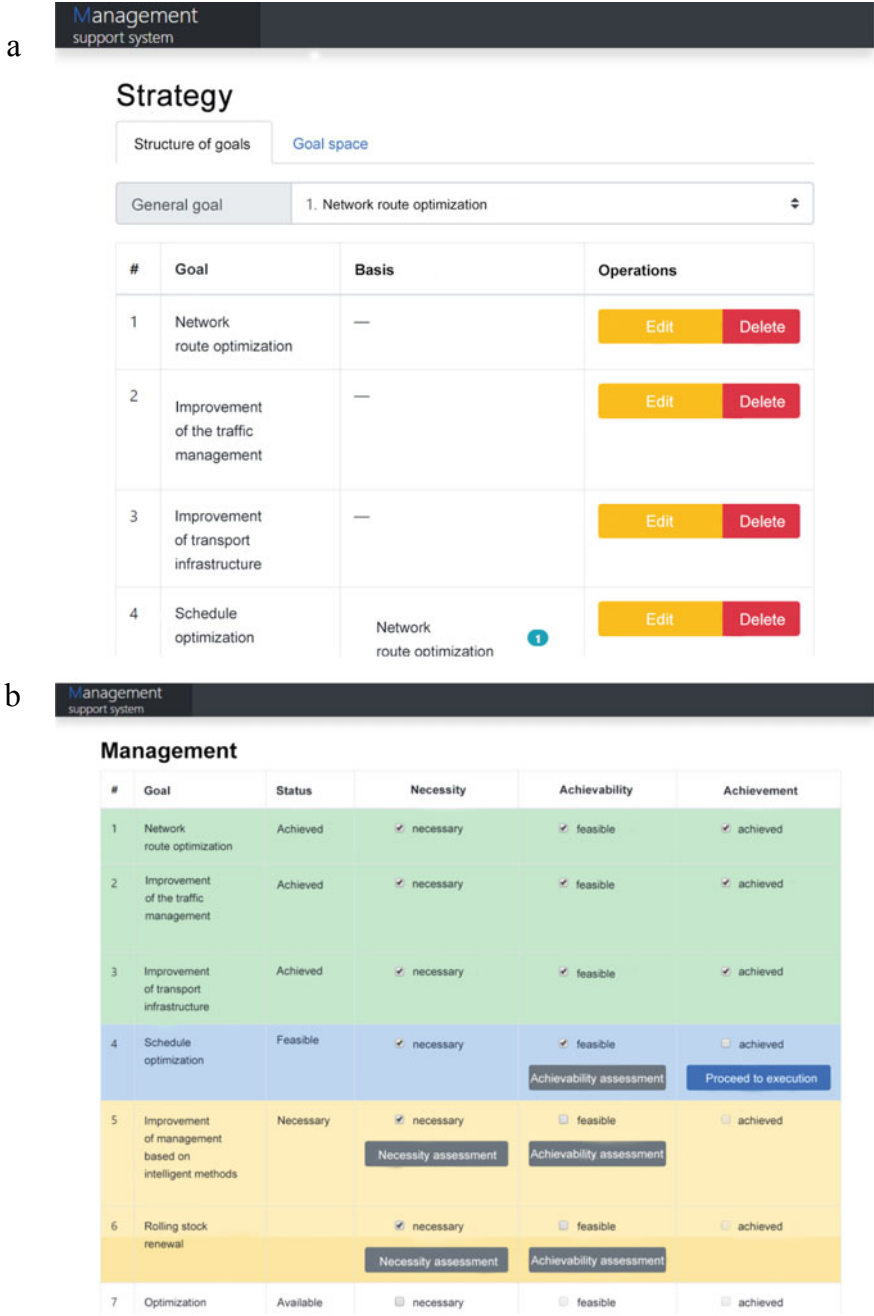


Fig. 5 Screenshots of the software complex for management support in cyber-social systems: a Formation of management strategy; b determination of available objectives

The integration of cybernetic principles into the management process in the SES makes it possible, based on goal setting, to determine the form of goal achievement, i.e. the order and type of actions determined by the respective goals. Narrativeness as a way of organizing management removes the problem of the inaccessibility of goals but ensures the need for the existence of SES through its integrability into the process of social development. Thus, the cyber-social system itself implements a model of narrative management, in which the behavior of people as subjects and objects of management is not subject to modeling, and the responsibility for the actions of participants in the management process lies with the participants themselves. The results of their actions, leading to the impossibility of achieving the goal, are compensated by goal space modification without the loss of its structural connectivity as a necessary condition for the system existence.

Acknowledgements The reported study was funded by the Russian Foundation for Basic Research (RFBR) according to the research project No. 20-07-00250_a.

References

1. Burkov, V.N., Irikov, V.A.: Control Methods of Organizational Systems. Moscow (1994) (in Russian)
2. Doran, G.T.: There's a S.M.A.R.T. way to write management's goals and objectives. *Manage Rev* **70**(11), 35–36 (1981)
3. SPIRO—A New (To You) Goal-Setting Model, <https://www.thecoachingtoolscompany.com/spiro-a-new-to-us-goal-setting-model/>. Last Accessed 08 April 2020
4. Odintsova, M.A.: Goal setting as one of the key problems of strategic management. *Econ. J.* **41**, 52–63 (2016)
5. Arkhipova, N.I., Kulba, V.V., Kosyachenko, S.A.: Organizational Management. Moscow (2007) (in Russian)
6. Chernyak, Y.I.: Information and Management. Moscow (1974) (in Russian)
7. Lukianova, L.M.: Systems analysis: the structure-and-purpose approach based on logic-linguistic formalization. *Int. J. Info. Theor. Appl.* **10**(4), 380–387 (2003)
8. Kushnikov, V.A., Murzin, S.I.: Development of goals models of management for complex social and economic systems based on sign digraphs. *Vestnik SGTU* **2**(43), 34–46 (2009)
9. Prazak, P., Bures, V.: Mathematical models of economic systems. *S. J.* **4**(20), 28–32 (2019)
10. Christensen, C., Albert, R.: Using graph concepts to understand the organization of complex systems. *Int. J. Bifurcat. Chaos Appl. Sci. Eng.* **17**(7), 2201–2214 (2007)
11. Burkov, V.N., Enaleev, A.K.: Stimulation and decision-making in the active system theory: review of problems and new results. *Mathe. Soc. Sci.* **27**(3), 271–291 (1994)
12. Neogy, S.K., Bapat, R.B., Das, A.K.: Optimization models with economic and game theoretic applications. *Ann. Oper. Res.* **243**, 1–3 (2016)
13. Farooqui, A.D., Niazi, M.A.: Game theory models for communication between agents: a review. *Complex Adapt. Syst. Model* **4**, 16–26 (2016)
14. Sabegh, M.H.Z., Mirzazadeh, A., Maass, E.C., Ozturkoglu, Y., Mohammadi, M., Moslemi, S.: A mathematical model and optimization of total production cost and quality for a deteriorating production process. *Cogent Mathe.* **3**(1), 12–23 (2016)
15. Subramanian, A.S.R., Gundersen, T., Adams, T.A.: Modeling and simulation of energy systems: a review. *Processes* **6**(238), 24–32 (2018)

16. Quang, L.A., Jung, N., Cho, E.S., Choi, J.H., Lee, J.W.: Agent-based models in social physics. *J. Korean Phys. Soc.* **72**(11), 1272–1280 (2018)
17. Clancey, W.J., Sierhuis, M., Damer, B., Brodsky, B.: Cognitive modeling of social behaviors, <https://cogprints.org/3966/1/CogSocialModelingClancey.pdf>. Last Accessed 28 April 2020
18. Papageorgiou, G., Hadjis, A.: Strategic management via system dynamics simulation models. *World Acad. Sci. Eng. Technol.* **7**, 227–232 (2011)
19. Parygin, D., Sadovnikova, N., Kravets, A., Gnedkova, E.: Cognitive and ontological modeling for decision support in the tasks of the urban transportation system development management. In: Proceedings of the Sixth International IEEE Conference on Information, Intelligence, Systems and Applications, pp. 1–5. IEEE (2015)
20. Sadovnikova, N., Parygin, D., Kalinkina, M., Sanzhapov, B.: Trieu Ni Ni: Models and methods for the urban transit system research. *Commun. Comput. Info. Sci.* **535**, 488–499 (2015)
21. Lynch, R.: *Corporate Strategy*. New York (2003)
22. Shabalina, O., Davtian, A., Sadovnikova, N., Parygin, D., Yerkin, D.: Narrative-based management in socio-economic systems. In: Proceedings of the International Conference ICT, Society and Human Beings 2017: Part of the Multi Conference on Computer Science and Information Systems 2017, pp. 73–79. IADIS (2017)
23. Shabalina, O., Yerkin, D., Davtian, A., Sadovnikova, N.: A Lattice-Theoretical Approach to Modeling Naturally Ordered Structures. In: Proceedings of the III International Scientific Conference on Information Technologies in Science, Management, Social Sphere and Medicine, pp.158–161. Atlantis Press (2016)
24. Doostmohammadian, M., Rabiee, H.R., Khan, U.A.: Cyber-Social Systems: Modeling, Inference, and Optimal Design. *IEEE Syst. J.* **14**(1), 73–83 (2020)
25. Stanford Cyber Initiative: Understanding “cyber-social systems”, https://cyber.stanford.edu/sites/default/files/stanford_cyber_initiative.pdf. Last Accessed 30 May 2020
26. Perno, J., Probst, C.W.: Behavioural Profiling in Cyber-Social Systems. *Hum. Aspects Info. Sec. Privacy Trust* **10292**, 507–517 (2017)
27. Hahanov, V., Soklakova, T., Hahanova, A., Chumachenko, S.: Cyber social computing. In: *Cyber Physical Computing for IoT-driven Services*. Springer, 233–250 (2018).
28. Shabalina, O.A., Sadovnikova, N.P., Parygin, D.S., Obratsov, E.A., Rubanyuk, V.N.: Support system for narrative management in socio-economic systems. *Electron. Sci. J. Eng. Bull. Don* **1** (2019), https://www.ivdon.ru/uploads/article/pdf/IVD_3_Shabalina_Sadovnikova.pdf_17937e01aa.pdf. Last Accessed 25 May 2020 (in Russian)

On the Problem of Computability of Bounded Rationality Cognitive Solutions



Vladimir Zaborovskij , Vladimir Polyanskiy , and Sergey Popov 

Let's not argue—let's do the math.
J. L. Lagrange

Abstract Modern science has not made the world less mysterious, and future discoveries may be even more surprising than the conclusions of relativity or quantum mechanics, confirming the thesis of A. Clarke that “advanced technology is indistinguishable from magic”. In other words, with a certain level of technology development, an adequate understanding of them may require intellectual resources that exceed the capabilities of a single human, primarily for handling large amounts of information. What is it, how are cognitive processes defined and physically implemented in nature, how effectively can their technical simulation be implemented, and, finally, whether it is possible to build not just a supercomputer, but a superbrain for solving super tasks. These issues are currently at the forefront of current research in both natural and human sciences, as well as computer technology. In this chapter, the problem is considered in the aspect of finding solutions of “bounded rationality” that is, based on the regularization of solutions obtained using computable functions that are defined on a set of data taking into account certain specific or even personal cognitive biases.

Keywords Technology development · Artificial intelligence · Machine learning · Exo-intelligence · Cognitive functions

V. Zaborovskij (✉) · S. Popov
Peter The Great St. Petersburg Polytechnic University, Saint-Petersburg, Russia
e-mail: zaborovskij_vs@spbstu.ru

S. Popov
e-mail: popovserge@spbstu.ru

V. Polyanskiy
Institute for Problems in Mechanical Engineering, Russian Academy of Sciences,
Saint-Petersburg, Russia
e-mail: vapol@mail.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_2

1 Introduction

Formally considered a problem is that the cognitive processes occurring in the human brain, being a part of objective reality,[1, 2] do not obey the common logic-mathematical principles of modern physics. To describe cognitive processes we need new models and applied methods that, on the one hand, would meet the principle of physical explanation of the phenomenon of intelligence, and with another—would consider their “informational entity” that in quantum physics, for example, connected with the effect of “retro-causality”, that is, the impact on the current state of the quantum object not only of past events but also events not yet held the future with the explanation of the phenomenon of quantum nonlocality or predictive modeling (calculations). The model that is constructive from the point of view of computer science should reflect aspects that are common to any type cognitive processes, related to their internal system, or more precisely, information organization, and not only with the properties of physical substance, which is a particular case can be considered as a carrier of cognitive processes.

Based on this, the research is based on the correspondence principle, which assumes that the formal description tools used at the system level reflect the properties of the studied objects that are available for measurement or experimental observation.[3, 4] According to this principle, the well-known properties of macroscopic physical substances in a 4-dimensional space–time continuum are characterized by the Archimedean metric and are accurately described by the properties of the real number field, but on the microscopic scale, the interaction processes of quantum particles have different nature and, therefore, are represented by operators in Hilbert space.

First of all, we will highlight the systemic essence of cognitive processes and try to understand by what kind of mathematical methods these processes can be adequately modeled, for example, to create artificial intelligence that combines the cognitive resources of people and classical Turing’ machines in a common global exo-intelligent digital space.

The research attempts to provide answers to the questions formulated above, based on the modern paradigm of computer science, the metaphorical formulation of which follows the ideas of R. Descartes, namely “compute ergo sum” (I calculate, so I exist). The implementation of this paradigm applied to the engineering education system allows preventing the situation when a hypothetical global failure of all computer systems no one will be able to spend vital for civilization calculations in manual, not because it is beyond human possibilities, but because this skill has ceased to learn, and as a result, the knowledge in the Sciences become incomplete and simplified to the extent that their use only to control “magical” devices, not having a full description and understanding of the principles of their work.

Solving this problem on the way to creating hybrid human–machine exo-intelligent systems will allow the designs a new space of digital knowledge that exists not only in a verbal-declarative but also in an operational-computational form available both for people and machines.

In our studies, we based a number of theses, namely: Thesis 1: intelligence as a phenomenon subject to natural physical laws; Thesis 2: the principle of “presumption of neutrality”, namely the need to find “natural” explanations for all the peculiarities of the manifestation of intelligence via formally defined cognitive functions; Thesis 3: the nature of intelligence is related to the organization of information connections between memory resources and data perception, and not to the properties of the substance in which the data obtained and the knowledge formed on their basis are stored.

2 Thesis 1: Why Bounded Rationality Solutions

Human behavior is based on the so-called. cognitive functions (attention, memory, perception, etc.), the carrier of which is the brain—an organ that combines the functions of controlling both biological and mental processes. From a formal point of view, such decisions are “fast”, although not strictly optimal, since they allow “cognitive distortions” of physical models of situations and the context of events, but they allow taking into account: (1) previous experience, (2) operational data, (3) cognitive forecast possible consequences of the decisions made. In many cases, the use of heuristics leads to the so-called decisions of bounded rationality, which may differ significantly from decisions based on conscious, controlled, and analytically sound decisions. However, heuristics can be very effective in both practical and specific situations. In practice, cognitive biases occur when people apply heuristics in previously unknown or unfamiliar conditions that do not fit the existing model of space–time patterns and current contexts of events. In the process of brain evolution, the realization of cognitive functions was associated with the development of controls for specific biological, perceptual, and motor operations, which inevitably deviate from the abstract laws of logic, the theory of probability, or mechanics.

The control processes of such operations using the biological neural network of the brain and the central nervous system were aimed at solving a poorly formalized problem—to deliver a cognitive subject to where he could survive. For this, cognitive functions must effectively implement universal mechanisms for coincidence detection, pattern recognition, and associative learning, which are important for maintaining the physical integrity of the subject in the natural environment. Poorly formalized decisions of “bounded rationality” differ significantly from decisions obtained on the basis of “higher” cognitive functions, such as analytical and symbolic reasoning.

However, in the process of evolution, the brain has not been optimized for the implementation of the cognitive functions that provide the processes of analytical thinking (for example, calculation, statistics, analysis, reasoning, abstraction, conceptual thinking) and which have become important for the purposes of human “survival” only relatively recently. Therefore, the natural characteristics of the brain as a complex neural network adapted for the implementation of the functions of perception and motor skills require additional resources used in solving conceptual

or analytical problems, which are based not on taking into account relative differences or comparisons, but on processing absolute values that characterize the phenomena under consideration. Understanding the mechanisms of cognitive processes associated with solving tasks for which the brain's neural network was initially optimized explains why a person can easily and effortlessly perform very computationally complex tasks of movement and perception (usually with massively parallel data flows), then has great difficulties in solving logical or arithmetic tasks, which are computationally much simpler.

3 Thesis 2: Computational Aspects

Cognitive processes of obtaining solutions of limited rationality give priority to information that is compatible and consistent with previously obtained knowledge while ignoring information that is not directly available or not recognized by the senses. There are four groups of cognitive distortions that determine the computability properties of cognitive functions, namely: association (correlations), compatibility, conservation, and attention. Currently, the greatest practical successes in solving intellectual problems have been achieved using methods that, in principle, are based on "brute computing force" or, in other words, on the ability to quickly sort possible solutions, for example, in analogy in playing chess. In our case, the "intellectualization" problem can be formulated as a solution to the inverse problem.

Many different options for representing such solutions in the form of a finite sequence of machine operations can be implemented in both inductive and deductive manners. In the first case, artificial neuromorphic computational structures are used, which are "programmed" by machine learning methods based on the previously classified dataset, and in the second case, digital platforms use target software libraries assembled from previously validated fragments. To concretize the task of building exo-intelligent platforms, we will rely on data aggregation and classification methods based on similarity criteria for objects or processes based on functional or structural associations, using the search for solutions based on the heterogeneous symbiosis of bio and machine intelligence resources shows on Fig. 1.

The basis of the technology of combining the computing resources of computer systems and the capabilities of human intelligence to solve P or NP complex tasks is the concept of mathematics of "big data". The "computational field" of such new math consists of (1) a tuple of "data-algorithms", (2) their meta characterization, and (3) resources of distributed heterogeneous reconfigurable computer platforms, the nodes of which are connected into the consistent system by "smart" data pass. Using the heterogeneity aspect of calculations, i.e. the use of processor elements with different architectures and set of operation, namely CPU, GPU, TPU, and FPGA [5], allows to effectively scale, both in "width" and "vertical", various algorithms of "extracting knowledge" from the processed data, "adjusting" taking into account the

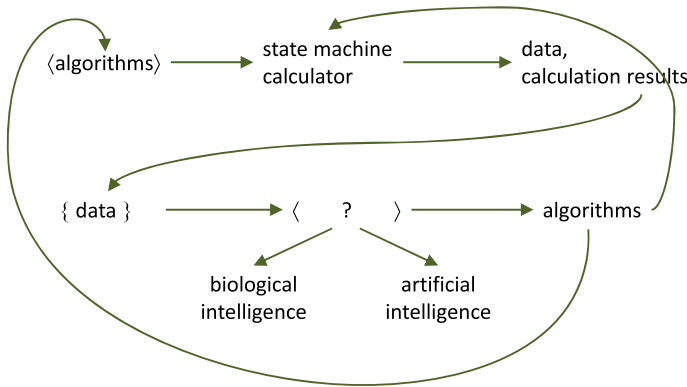


Fig. 1 The distribution between human and intelligence machine resources

situational context implemented a variety of algorithms that are used to assess potential threats and risks associated with the implementation of “calculated” solutions and the accuracy of computer models used [6].

It should be noted that due to heterogeneity and functional scaling, in exo-intelligent platforms, data processing algorithms oriented to the processor-centric approach can be effectively combined with memory-centric solutions, which allows simulating cognitive functions that can be characterized as “computational insight”, based on mechanisms to speed up the search algorithms by indexing the target data warehouse, potentially containing the whole spectrum of responses to correctly formulated queries. Similar mathematical processing technology is now widely used in modern search engines Yandex and Google, successfully modeling the functions of intellectual activity, which are usually associated with the concept of intuition. However, the technical implementation of this technology remains “flat,” that is, the adaptation of platform elements occurs only at the software level, when, as hardware components, they are formed from standard industrial systems.

This leads to a decrease in integrated energy-computational efficiency due to the fact that all modern processor elements at the macro level are built on the basis of logical gates AND-NOT or OR-NOT, therefore, any computational operations reduce the informational entropy of the processed data, contributing to the release of thermal energy no less than $Q = k * T * \ln 2$ J of energy, where k is the Boltzmann constant and T is the temperature of the system. This energy itself is small, so Q for $T = 300$ K is 0.017 eV per bit, but in terms of the number of logical elements (LE) which in the modern microprocessor (MP) is equal to $2-5 \times 10^{10}$, the total energy at a switching frequency of LEs of 5 GHz grows to values of the order of 1 J for each second of MP operation. Therefore, if modern microprocessors combined in a supercomputer cluster, and try to simulate the work of the humane brain, which including 1.5×10^{17} LE, then the energy costs will exceed the level of practical expediency of using computer technologies, can be represented only of purely scientific interest.

Although reconfigurable computers have a higher specific energy efficiency of digital elements compared to MIMD microprocessors and SIMD graphics accelerators, FPGAs, however, perform work at much lower frequencies, which, however, expands the synthesis capabilities, allowing, or to reduce unit costs energy when achieving comparable performance, or at the same energy costs to obtain greater productivity of data processing processes through the use of special architectural solutions [7–10]. Achieving the technical and economic efficiency of using exo-intelligent platforms requires the search for solutions balanced in the aspect of “standardization-specialization”, which, along with the effective implementation of standard computing procedures, have the resources necessary for using machine learning technologies and reconfiguring hardware accelerators taking into account the structural features of the implemented algorithms. In view of the foregoing, the transition to the use of technologies of hyper-converged clustering of heterogeneous processors, storage devices of a memory class, “smart” data channels endowed with intelligent processing functions, using specialized processors optimized for processing packet traffic, to create exo-intelligent platforms requires the development of universal heterogeneous computing modules that allow “vertical” and “horizontal” functional integration with the allocation of mutually agreed levels of “processing”, “aggregation” and “explanation” of calculation results. An ex-intelligent solution based on a hyper-convergent processor/storage platform differs significantly from well-known approaches implemented in the framework of the “one program—a lot of data” model (SPM model), Amdahl-Ware phenomenological laws for programs with an invariable proportion of serial and parallel computing or the law Gustavson-Bors for programs that may be complicated due to the increase in the volume of processed data because they are based on the use of hardware reconfiguration methods, which are supported by machine learning algorithms.

The use of computational acceleration nodes as a basic component of the heterogeneous software and hardware reconfigurable platform extends its functionality by quickly adapting the hardware components to the features of Fig. 2 solution method, algorithms, and corresponding source code that is implemented at a given time.

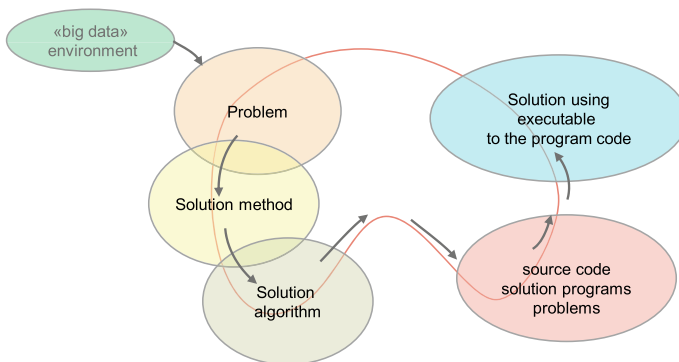


Fig. 2 Solution method, algorithms, and corresponding source code

4 Thesis 3: Discussion

Relentless digital transformation of modern knowledge, examples of which are supercomputer methods of predictive modeling, approximate optimization based on a random search, genetic algorithms based on meta-heuristics borrowed from nature, strictly speaking, form the components of a new exo-intellectual infrastructure for solving complex fundamental and applied tasks. This infrastructure is harmoniously supplemented by reinforcement machine learning methods, the formal prototypes of which are Robinson-Monroe and Kiefer-Wolfowitz stochastic approximation methods, as well as other well-known random search numerical optimization methods. Although all these methods were actively developed long before the advent of AI tasks, their implementation on modern hyperconverged computing platforms opens up new possibilities for integrating the resources of the natural intelligence of people and artificial intelligence of “smart” machines [11–15]. The main capabilities of hyper-convergent high-performance computing platforms depend significantly on the balanced loading of all hardware components and the correspondence of their architecture to the specific features of the application programs.

Therefore, the proposed solution-wide is based on “machine learning” methods to associate hardware platforms as well as software components of cognitive function in order to increase boundary rationality solution accuracy as well as energy consumption.

With regard to computing platforms that are used to carry out cognitive operations in order to determine the associative features characteristic of the set of reversible data, Fig. 3 presented attractors of states of individual layers of the Siamese neural network, which allow, on the basis of their analysis of their structure, to judge whether the sister has achieved an “acceptable” rational solution.

5 Conclusions

The idea to clarify the essence of bounded rationality cognitive solutions in terms of computable functions and corresponding neuromorphic computing platform which architecture was chosen and parameters tuning by “machine learning” process that taking into account the specific features of heuristic algorithms can be very attractive due to significantly simplify harmonic integration with the associative “exo-intellectual” infrastructure that can adapt humane and machine resources. For the effective use of such an infrastructure, deep interdisciplinary training of computer science specialists is of particular importance, which requires the careful development of new education programs that cover both fundamental and applied aspects of the natural and engendering sciences, including the use of fundamental methods for predictive analytics, big data mathematics, and machine learning.

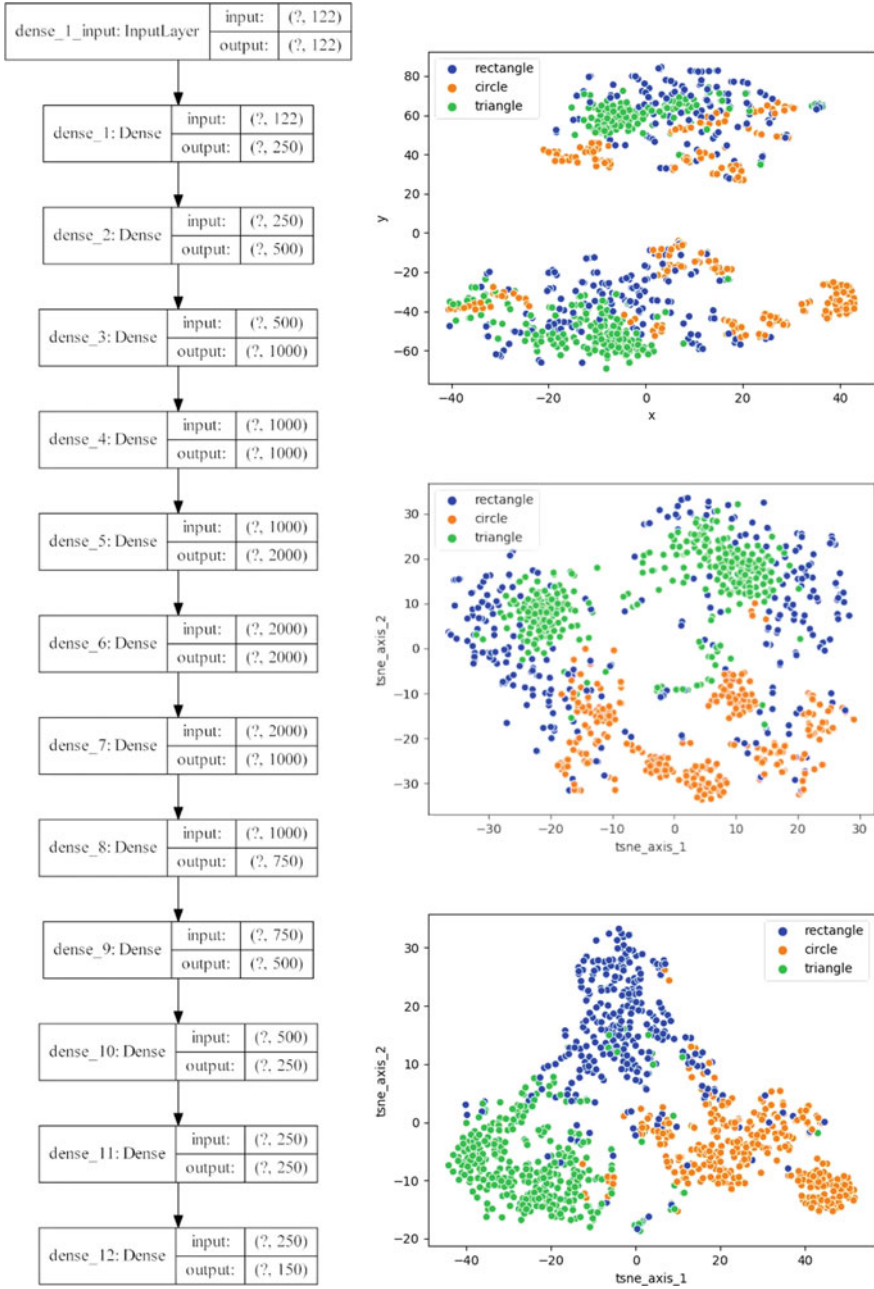


Fig. 3 Attractors of states of individual layers of the Siamese neural network

Acknowledgements The authors are grateful to the Supercomputer Center ‘Polytechnic’ for the help in gaining access to the resources of the supercomputer. The reported study was funded by RFBR according to the research project №18-2903250 MK “Robust methods of synthesis of intelligent transport systems of cyber-physical objects coalition based on the Bayesian concept of probability and the modal logic”.

References

1. Todd Oppenheimer.: The Flickering Mind: Saving Education from the False Promise of Technology, p. 528. Random House Publishing Group (2007)
2. Kalyaev, I., Antonov, A., Zaborovske, V.: The architecture of Reconfigurable Heterogeneous Distributed Supercomputer System for Solving Problems of Intelligent Data Processing in the Era of Digital Transformation of the Economy, pp. 2–11. *Voprosy kiberbezopasnosti*. 2–11. <https://doi.org/10.21681/2311-3456-2019-5-02-11>
3. Antonov, A., Zaborovskij, V., Kisilev, I.: Specialized reconfigurable computers in network-centric supercomputer systems. *High Availab. Syst.* **14**(3), 57–62 (2018). <https://doi.org/10.18127/j20729472-201803-09>
4. Usman Ashraf, M., Alburai Eassa, F., Ahmad Albeshri, A., Algarni, A.: Performance and power efficient massive parallel computational model for HPC heterogeneous exascale systems. *IEEE Access* **6**, 23095–23107 (2018). <https://doi.org/10.1109/ACCESS.2018.2823299>
5. Intel FPGA. [Online]. Available: <https://www.intel.com/content/www/us/en/products/programmable.html>. Last Accessed 19 April 2020
6. Dongarra, J., Gottlieb, S., Kramer, W.: Race to exascale. *Comput. Sci. Eng.* **21**(1), 4–5 (2019). <https://doi.org/10.1109/MCSE.2018.2882574>
7. Supercomputer Center ‘Polytechnic’. [Online]. Available: <https://www.top500.org/system/178469>. Last Accessed 19 April 2020
8. NVIDIA Tesla V100. [Online]. Available: <https://www.nvidia.com/en-us/data-center/tesla-v100/>. Last Accessed 19 April 2020
9. Xilinx FPGA. [Online]. Available: <https://www.xilinx.com/>. Last Accessed 01 Feb 2020
10. IDE Vivado HLS, <https://www.xilinx.com/video/hardware/vivado-hls-tool-overview.html>. Last Accessed 19 April 2020
11. UltraScale and UltraScale+ FPGA Product Table, <https://www.xilinx.com/products/silicon-devices/fpga/virtex-ultrascale.html#productTable>. Last Accessed 19 April 2020 (2019)
12. Sorting Methods, <https://www.mathworks.com/matlabcentral/fileexchange/45125-sorting-methods?focused=3805900&tab=function>. Last Accessed 19 April 2020
13. Vitis_Libraries, https://github.com/Xilinx/Vitis_Libraries. Last Accessed 19 April 2020
14. Antonov, A., Besedin, D., Filippov A.: Research of the efficiency of high-level synthesis tool for FPGA based hardware implementation of some basic algorithms for the big data analysis and management tasks. In *Proceedings of the FRUCT’26*, pp. 23–29 (2020)
15. Merge sort, https://rosettacode.org/wiki/Sorting_algorithms/Merge_sort. Last Accessed 19 April 2020

Image Processing for Biometric Scanning of the Palm Vein Pattern



Lina Kh. Safiullina  and Rustem R. Maturov

Abstract Due to increased requirements for access control systems, the use of biometric recognition technologies is becoming a reliable solution for the protection of critical information. One of the best ways of personal identification is to use the palm vein structure. The chapter deals with improving accuracy in the problem of recognizing the palm vein pattern when comparing biometric templates using the Canny edge detection algorithm and the Gabor filter. 2D Gabor filter improves the adaptability of recognition and is therefore proposed to solve the problem of image blurring and select a threshold when the traditional Canny algorithm smooths the edges. The results of experiments show that this filter can detect less pronounced edges and provides more complete information about the image, which has a positive effect on the result of biometric authentication. The similarity of two biometric templates is determined using the Minkowski metric. Experiments conducted on the original facility show high performance, as well as good results in false acceptance errors (FAR = 0%) and false rejection errors (FRR = 0.01%) based on processing 360 images captured from 26 people, which makes it possible to use the proposed method in the identification and authentication system at existing data security facilities.

Keywords Canny algorithm · Gabor filter · Biometrics · Vein pattern · Identification · Authentication

1 Introduction

This chapter discusses the method of recognizing a person by the palm vein pattern captured in the IR spectrum. Initially, infrared scanning was used in healthcare to form a vein map and diagnose venous disorders. However, after Fujitsu became

L. Kh. Safiullina (✉) · R. R. Maturov
Kazan National Research Technological University, Kazan, Russia
e-mail: lina.kh.safiullina@mail.ru

R. R. Maturov
e-mail: maturov18@mail.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_3

interested in the scanning method, the first laboratory vein recognition system for human identification was installed in 2006 in the public library of the Japanese city of Naka. The widespread use of the palm vein pattern identification/authentication method began in Japan to prevent bank fraud around the mid-2000s, making it one of the youngest identification/authentication methods [1]. This technology is based on the fact that hemoglobin in the blood absorbs infrared (IR) radiation, i.e., when the palm is illuminated with IR light, veins, unlike other parts of the hand, do not reflect light and looks like dark parts. Thus, the unique pattern of the palm venous network can be captured using infrared radiation.

The technology of biometric user authentication based on the vein pattern has many significant advantages:

- possibility of contactless identification (which is more hygienic);
- high level of uniqueness;
- static biometric characteristic over a long time;
- high resistance to falsification due to the complexity of creating an artificial model of the user's vein pattern;
- the impossibility of stealing a biometric characteristic;
- low cost of hardware and the entire system,

which makes it applicable to many areas [2].

The main difficulty in implementing the vein pattern recognition method is the high-quality processing of images that highlight the structure of veins without foreign objects, not related to them: skin creases, the borders of IR illumination, the background behind the palm, the borders of the palm, etc. At the same time, the performance of a biometric human recognition system, represented by the following key indicators [3]:

- false acceptance rate (FAR);
- false rejection rate (FRR);
- failure to enroll (FTE);

should not exceed 0–0.05% for each of them.

2 Problem Statement

The purpose of this research is to improve the quality and performance of human recognition based on palm vein patterns, using the Canny algorithm and the Gabor filter.

2.1 Recognition Device Operating Principle

The designed hardware and software system (HSS) consists of a scanner, an application handler, and a program comparing biometric samples. The scanner creates an image of the palm of the user to be identified and then sends it to an application handler written in the built-in Matlab language, which, in turn, generates a binary pattern with an image of palm veins (biometric template). The resulting biometric template is processed by a program created in Microsoft Visual Studio 2018 in C++ that compares it with other templates contained in the database, and then the authentication results are sent to the ID scanner. If the palm image matches the existing templates, a corresponding message appears on the screen. The appearance of the device and its connection diagram is shown in Fig. 1.

The Arduino UNO microcontroller with an integrated development environment (Arduino IDE) was chosen as the programmer unit to control the whole set of device components. The Arduino UNO is connected to 6 IR diodes and a distance sensor. The diodes are positioned symmetrically on the card around the camera and point in the direction where the palm is brought. They are tilted towards the camera at a certain angle to focus their rays on the identification object, which is about 19 cm away from the lens. The programmer and camera are connected via a USB cable to a computer with all the necessary software installed.

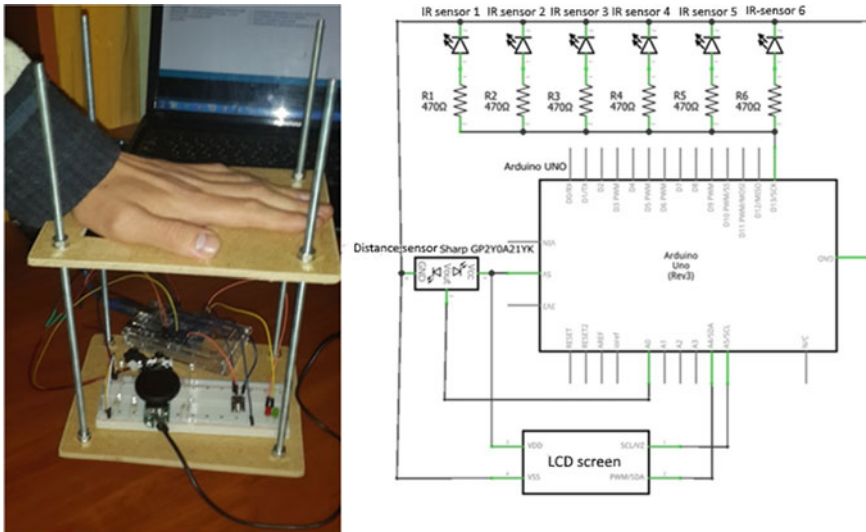


Fig. 1 Appearance (left) and diagram (right) for connecting components to the HSS programmer

2.2 Device Operation Algorithm

After switching on, the device goes into standby mode. In this operating mode, only the distance sensor is activated. It is directed towards the palm, and when the identifier (hand) approaches the device, the sensor sends information about the distance to the camera to the microcontroller. If the distance is the same as the value pre-programmed in the microcontroller, a disabling signal is sent to the sensor; the camera and IR illumination are switched on. At the same time, the distance sensor is disabled, because the light it emits to measure the distance to the object will create noise and interference to the camera, the lens of which will reflect its beam through an infrared filter along with an IR beam, which will create a “flash exposure” and adverse noise for the subsequent processing of the image.

It is enough to turn on IR illumination for two seconds after the camera is turned on and ready; this time suffices to focus the lens on the identification object and capture an image. After creating an image, it is better to disable the camera and the illumination; the image is sent for processing by one program, and the result of authentication comes from another program.

Next, the result is passed to the identification object: if positive, the signals are sent to the green LED; if negative, the signals are sent to the red LED. Then the device switches to the initial mode.

2.3 Image Processing

As mentioned above, one of the main problems in image processing is the separation of the edges of the vein map and the removal of papillary patterns, shadows, and reflections. The heterogeneity of the image and the angle of the light source create a sharp change in the color shade in the image. As a result of this difference, there is a probability of erroneous allocation of an area that is out of all relation to the vein pattern [4]. It might often be the case that the algorithm cannot distinguish the edges of small objects. As a rule, the borders of such “islands” are poorly defined and are absorbed by the objects against which they are located.

The Gabor filter and the Canny algorithm were used to filter the image and improve the quality of recognition. The Gabor filter, which belongs to the bandpass filter family, is a linear filter used for detecting borders. Frequency and orientation representations of Gabor filters are similar to those of the human visual system in terms of textural representation and discrimination [5]. In the spatial domain, the two-dimensional Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. The formula for constructing a two-dimensional Gabor filter (1–3):

$$G(x, y) = \exp\left(-\frac{1}{2}\left[\frac{x_\phi^2}{\sigma_x^2} + \frac{y_\phi^2}{\sigma_y^2}\right]\right) \cos(2\pi\theta x_\phi) \quad (1)$$

$$x_\phi = x \cos(\phi) + y \sin(\phi) \quad (2)$$

$$y_\phi = -x \sin(\phi) + y \cos(\phi) \quad (3)$$

where σ_x, σ_y are the deviations of the Gaussian kernel along the x and y axes, which determine the length of the filter, θ is the frequency modulation of the filter, ϕ is the spatial orientation of the filter that determines its position relative to the main axes.

Image processing by the Gabor filter is achieved by averaging the values of the processed image over a certain area at each point. Accordingly, the imposition of the Gabor filter on the image has the form:

$$I'(x, y) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n I\left(x - \frac{n}{2} + i, y - \frac{n}{2} + j\right) \cdot G(i, j), \quad (4)$$

where $I(x, y)$ is the intensity of the source image at the point (x, y) ,

$I'(x, y)$ is the intensity of the new image at the point (x, y) ,

$G(i, j)$ is the Gabor function value, $i \in [0, n], j \in [0, n]$.

Before applying the filter, it is necessary to construct the mid-frequency and orientation fields using four gradations for the current image and then make a convolution in all points of each direction. The result of applying this filter is shown in Fig. 2.

The Canny algorithm generates an optimal smoothing filter based on the criteria for detecting, localizing, and minimizing multiple responses to a single edge [6]. The edge of the image can have different directions, so the Canny algorithm uses four filters to detect horizontal, vertical, and diagonal edges of the blurred image. The edge detection operator returns the value for the first derivative in the horizontal (G_x) and vertical (G_y) directions. From here, the gradient and direction of the rib can be determined:

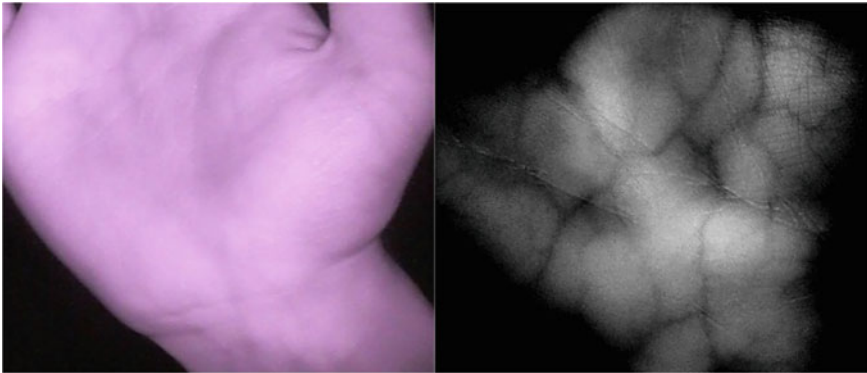


Fig. 2 Image before (left) and after (right) processing

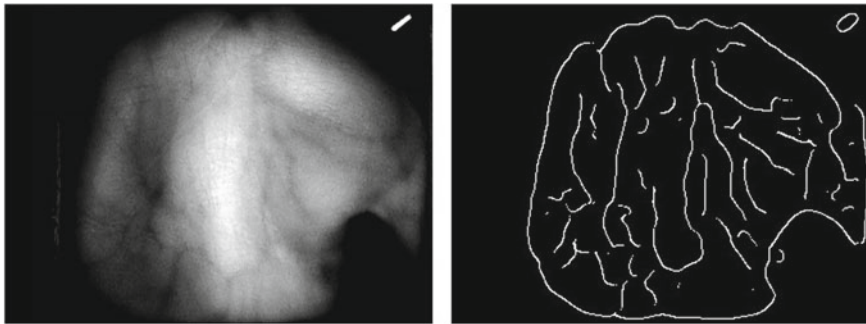


Fig. 3 HSS image before (left) and after (right) applying the Canny algorithm

$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}, \quad \Theta = \text{atan2}(\mathbf{G}_y, \mathbf{G}_x), \quad (5)$$

where G is the length of the hypotenuse,

atan2 is an arc tangent function with two arguments.

The edge direction angle is rounded to one of the four angles that represent the vertical, horizontal, and two diagonals (for example, 0° , 45° , 90° , and 135°). The result of applying the Canny filter in this chapter is shown in Fig. 3.

The final binary pattern depends on the sensitivity of the *threshold value* and the standard deviation of the Gaussian filter (*sigma*) [7, 8]. The optimal values were assumed to be $\text{threshold} = 0.25$ and $\text{sigma} = 10$.

2.4 Comparison of Biometric Templates

As a result of applying the Gabor filter and the Canny operator, we get a binary pattern that is converted into a perceptual hash, which is used for further comparison with reference hash patterns. To obtain a perceptual hash, the following transformations over the images were sequentially performed:

1. Size reduction to 8×8 pixels.
2. Discoloration (i.e., converting all colors to grayscale gradients).
3. Binarization (i.e. converting each pixel to either black or white). In this case, the threshold of the binarization process is the obtained average color value.
4. Conversion of the resulting 64-bit pixel value into a hash.

After these transformations, two hashes are compared with each other using the Minkowski metric [9–12]:

$$\rho(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p} \quad (6)$$

Table 1 Comparison of FAR, FRR and FTE indicators

Authors	Objects	Number of images to analyze	FAR (%)	FRR (%)	FTE (%)
Kulkarni et al. [17]	50	500	5.66	5.66	0
Chunyi et al. [18]	40	588	5.19	0.02	0
Qui et al. [19]	106	3816	0.1	0.09	0
Joseph [20] (matching without fuzzy retrieval)	156	6264	0.25	0	0
Joseph [20] (matching after fuzzy retrieval)	156	6264	0.17	0	0
This research	26	360	0	0.01	0

3 Results and Discussion

The results of recognizing and processing 360 images of 26 users were analyzed using three key indicators of biometric systems [15, 16]:

- False Access Rate (FAR)—the percentage threshold that determines the probability that a user will be incorrectly accepted;
- False Rejection Rate (FRR)—the probability that the biometric security system will incorrectly reject an access attempt by an authorized user.
- Failure to Enroll (FTE)—the probability that the user will not be able to be enrolled in a biometric system due to an insufficiently distinctive biometric sampling.
- A comparison of these indicators with data published in the scientific literature is presented in Table 1.

Despite the small number of images analyzed in this chapter, the method described above shows rather well values of performance indicators. In Kulkarni et al. [17], Chunyi et al. [18] error rates are too high. The FAR value is especially dangerous for the access control system because it shows the percentage of users who can access it without having the right to do so.

Results of Qui et al. [19], Joseph, Ezhilmaran [20] have good performance indicators, and the analysis was performed using a very large number of images, which reduces the probability of statistical error.

The authors of the chapter will continue to increase the biometric data bank and clarifying the performance indicators of the biometric system.

4 Conclusion

The results of a statistical analysis based on 26 users and 360 images for testing the HSS and the proposed recognition algorithm showed the following:

- false acceptance rate FAR = 0%;
- false rejection rate FRR = 0.01%;
- failure to enroll FTE = 0%,

which is within normal limits [3]. Therefore, the combination of the Canny operator and Gabor filter can be considered appropriate to improve the algorithm for recognizing a person by the palm vein pattern.

References

1. Wu, W., Elliott, S. J., Lin, S., Sun, S., Tang, Y.: Review of palm vein recognition. *IET Biom.* **9**(1), 1–10 (2019)
2. Miura, N.: Feature extraction of finger-vein patterns based on repeated line tracking and its applications to personal identification. In: Miura, N., Nagasaka, A., Miyatake, T. (eds.) *Machine Vision and Applications*, pp. 194–203 (2004)
3. Antipov, R.S., Martynenko, T.V.: Automated access control and management system based on the analysis of human biometric parameters. *Comput. Sci. Cybern.* **1**(15), 21–26 (2019). DonNTU Publ., Donetsk
4. Grizhebovskaia, A.G., Mikhalev, A.V.: A biometric method of identification of a person by the vascular pattern of the finger. *Cybersecurity* **5**(33), 51–56 (2019)
5. Sakharova, M.A.: Fingerprint image processing using the Gabor filter. *Act. Probl. Aviat. Astronaut* **2**, 167–169 (2018)
6. Canny, F.J.A.: computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698 (1986)
7. Kim, Y.W., Oh, A.R., Krishna, A.V.: Analyzing the performance of canny edge detection on interpolated Images. In: *International Conference on Information and Communication Technology Convergence (ICTC)*. – June, 2018. <https://www.cnki.net/kcms/doi/10.14132/j.cnki.1673-5439.2018.03.011.htm>. Accessed 12 Feb 2020
8. Fu, F., Wang, C., Li, Y., Fan, H.: An improved adaptive edge detection algorithm based on canny. In: *Sixth International Conference on Optical and Photonic Engineering*. – July, 2018. <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10827/2500361/An-improved-adaptive-edge-detection-algorithm-based-on-Canny/10.1117/12.2500361.short>. Accessed 14 Feb 2020
9. Lepsky, A.E., Bronevich, A.G.: *Mathematical Methods For Pattern Recognition: Course of Lectures*. Taganrog: TTI SFU Publ., 155 p (2009)
10. Suyatinov, S.: Bernstein’s theory of levels and its application for assessing the human operator state. In: Dolinina, O., et al. (eds.) *Springer Nature Switzerland AG*, pp. 298–312 (2019). ICIT 2019, SSSC 199. https://doi.org/10.1007/978-3-030-12072-6_25
11. Matokhina, A.: Method of the exoskeleton assembly synthesis on the base of anthropometric characteristics analysis. *Stud. Syst. Decis. Control* **259**, 361–393
12. Agafonov, V.: Super-resolution approach to increasing the resolution of image. In: Kravets A., Shcherbakov M., Kultsova M., Iijima T. (eds.) *Knowledge-Based Software Engineering, JCKBSE 2014. Communications in Computer and Information Science*, vol. 466. Springer, Cham (2014)
13. Xin, M., Xiaojun, J.: Palm vein recognition method based on fusion of local Gabor histograms. *J. China Univ. Posts Telecommun.* **24**(6), 55–66 (2017). [https://doi.org/10.1016/s1005-8885\(17\)60242-5](https://doi.org/10.1016/s1005-8885(17)60242-5)
14. Liu, J., Jing, X.J., Sun, S.L., et al.: Local Gabor dominant direction pattern for face recognition. *Chin. J. Electron.* **24**(2), 245–250 (2015)

15. Wang, J.G., Yau, W.Y., Suwandy, A., et al.: Fusion of palmprint and palm vein images for person recognition based on “Laplacianpalm” feature. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR’07), Jun 17 – 22, 2007, 8 p. IEEE, Minneapolis, MN, USA. Piscataway, NJ, USA (2007)
16. Wang, L.Y., Leedham, G., Cho, D.S.Y.: Minutiae feature analysis for infrared hand vein pattern biometrics. *Pattern Recogn.* **41**(3), 920–929 (2008)
17. Kulkarni, S., Raut, R.D., Dakhole, P.K.: A Novel authentication system based on hidden biometric trait. *Procedia Comput. Sci.* **85**, 255–262 (2016). <https://doi.org/10.1016/j.procs.2016.05.229>
18. Chunyi, L., Mingzhong, L., Xiao, S.: A finger vein recognition algorithm based on gradient correlation. *AASRI Procedia* **1**, 40–45 (2012). <https://doi.org/10.1016/j.aasri.2012.06.008>
19. Qiu, S., Liu, Y., Zhou, Y., Huang, J., Nie, Y.: Finger-vein recognition based on dual-sliding window localization and pseudo-elliptical transformer. *Expert Syst. Appl.* **64**, 618–632 (2016). <https://doi.org/10.1016/j.eswa.2016.08.031>
20. Joseph, R.B., Ezhilmaran, D.: A smart computing algorithm for finger vein matching with affine invariant features using fuzzy image retrieval. *Procedia Comput. Sci.* **125**, 172–178 (2018). <https://doi.org/10.1016/j.procs.2017.12.024>

Automation of Demand Planning for IT Specialists Based on Ontological Modelling



Denis V. Yarullin , Rustam A. Faizrakhmanov, and Polina Y. Fominykh

Abstract The chapter addresses the issue of correspondence between the skills required by employers and the professional competencies of specialists in the IT field. The discrepancy between the said sets has been highlighted. An approach based on extracting skills from the natural language vacancies texts published on the job aggregators sites is proposed. The method allows for analyzing the required professional competences from the employers' point of view to eliminate the identified differences. Possible ways of structuring the selected skills including ontological modeling and cluster analysis are described. An ontological model has been created to proceed with the hierarchical structuring of the professional competencies set. Skill groups have been formed based on domain knowledge, and cluster analysis has been applied to form workload sets. The method of dynamic cluster formation and skill attribution to a particular group within the domain is described. The applied aspects of the approaches are examined using data of Russian regions and federal states of Germany. The differences between a set of workloads and a skill set are determined. The strengths and weaknesses of the highlighted approaches are described. The automation method of demand planning for IT specialists based on an integrated model combining the described approaches above is suggested. Prospects for its further application are outlined.

Keywords IT-specialist · Skill · Structuring · Ontology · Cluster analysis · Workload sets · Dynamic modeling

D. V. Yarullin (✉) · R. A. Faizrakhmanov · P. Y. Fominykh
Perm National Research Polytechnic University, 29 Komsomolsky prospect, Perm 614990, Russia
e-mail: d.v.yarullin@gmail.com

R. A. Faizrakhmanov
e-mail: Fayzrakhmanov@gmail.com

P. Y. Fominykh
e-mail: phominykh1997@gmail.com

1 Introduction

The study of the competencies required for a specialist in a particular field could be essential for both an employer and a potential employee. Currently, the skills employers expect from specialists often do not correspond to the skills specialists acquire during their training [1].

By 2013, 15% of Russian companies reported a lack of employees' qualifications. Moreover, the qualification gap in the Russian labor market is not a temporary issue, but a constant and worrying trend [2]. Such skills discrepancies are also observed in European countries. Data analysis revealed a gap between the skills of young professionals and the employers' requirements in different activity sectors. It is important to note that such skills discrepancies are more prevalent in intensively developing fields [3].

In the present study, the issue of skills discrepancy is examined on the example of software engineering specialists and their professional competencies. This field is rapidly developing in the labor market, and the demand for specialists is constantly growing. *HH.ru* job aggregator data analysis shows a vacancy rate 5.5% increase for IT specialists during the period from 2016 to 2018 [4]. At the same time, it is necessary to note that the necessity in IT specialists arises in other areas where intensive processes of automation, digitalization, analytics take place. The demand for IT specialists is observed in all the developed countries, and the staff shortage issue is rather acute [5, 6]. Therefore, closing the gap between the skills required by employers and real specialists' competences is especially important in this field.

We suggest that in order to successfully identify discrepancies between the employers' requirements and the actual skills of candidates, an approach allowing analysis and formalization of the said requirements is necessary. The skills systematization will provide an opportunity to take into account the most valuable competencies during the programs planning for future specialist training.

2 Approaches to Solving the Problem

The proposed approach is based on the method of extracting the employers' requirements from the vacancies texts published on the job aggregator sites. The vacancies descriptions published on the *HH.ru* aggregator were selected as the primary data source [7]. The data were retrieved via API using the "Programmer" query for each Russian region individually. After the vacancies' text retrieval, the skills were collected using Natural Language Processing (NLP) techniques. The method includes the following steps:

1. Splitting the text into sentences;
2. Splitting the sentences into words;
3. Words normalization;
4. Stop-words deletion and filtering;

5. Converting words into vectors [8, 9].

This algorithm was applied to vacancies text. The skills were identified during the filtering step. After the skills were extracted, it was necessary to create a model that would allow structuring the data. Two approaches were considered as possible methods: ontological model and cluster analysis. The ontological model does not require words to be converted into vectors, so the skills set was formed based on the ontology before the words to vectors' conversion step.

The ontological model is a hierarchical structure of the domain concepts. The formation of ontology includes a description of the studied issue using concepts, their attributes, and specific objects. This could be compared to the object-oriented programming paradigm, where concepts are presented as classes, properties are class attributes, and objects are class instances. The advantage of this structure of knowledge organization is the machine processing capability as well as the flexibility and scalability. At the same time, ontology is presented as a holistic model of knowledge [10, 11]. Retrieved skills do not initially have any system organization, so there are no logical links between certain skills. It should also be noted that often vacancies do not explicitly specify a full hierarchy of required skills. For example, a vacancy may mention the Flask framework proficiency, but it does not highlight that Flask is a Python programming language framework. Thus, the vacancy implicitly specifies the following requirements: proficiency in the Python programming language, proficiency in the Flask framework. The integrity of ontology allows us to restore missing dependencies and build a complete list of necessary skills.

A skill was chosen as a key concept forming the structure of the ontology, defined as a fragment of the domain knowledge, which allows performing specific tasks within the domain [12]. Using the automatic construction model of the hypertext denotation graph, the following groups of skills were defined [13]:

1. Programming language;
2. Development environment;
3. Library;
4. Framework;
5. Programming technologies;
6. Operating system;
7. Software;
8. Information transfer protocols/Server;
9. Non-specialized skills.

Figures 1, 2 and 3 show the ontology excerpts based on the vacancies localized in Moscow.

The presented graph has a rather complicated structure and a multilevel dependency. Moreover, an instance of one class can be an instance of another class. It is also worth mentioning that the list of classes in the method is constant.

For the second method of skills structuration, it is necessary to convert words to their vector representations as this method functions in vector space. The "Bag of Words" approach was used in the study. The given algorithm allows us to define

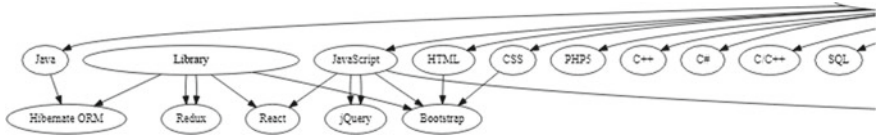


Fig. 1 Fragment of the “Programming languages and libraries” graph for Moscow

Fig. 2 Fragment of the “Operating system” graph for Moscow

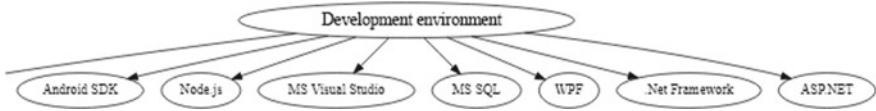
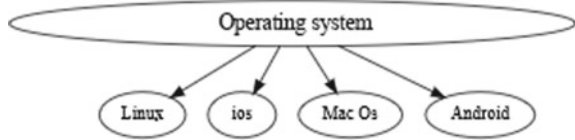


Fig. 3 Fragment of the “Development environment” graph for Moscow

the word usage frequency in the overall text scope. For the attribute calculation, the metric TF-IDF was used, where TF is term frequency, and IDF is inverse document frequency. This approach does not take into account the word order in the text, which can lead to data loss. This drawback is eliminated to some extent by using the N-gram algorithm, which makes it possible to consider not only words but also phrases. A combination of these methods reduces the number of errors in semantic understanding of words with the same spelling but different meanings [14, 15]. The algorithms were implemented using Python programming language, and NLTK library for natural language symbolic and statistical processing [16].

For further processing, cluster analysis was performed to create workload sets. Affinity propagation was chosen as the clustering algorithm. The algorithm automatically determines the structure and number of clusters by passing messages between vector representations of words. When passing the information about the points’ location relative to each other, matrices are formed that define the “leader” of the cluster and the points that fall into the cluster with the said leader. Recalculation of the matrices occurs until the system is settled [17, 18]. Cosine similarity had been chosen as a metric determining the elements affinity [19]. The similarity level between vectors A and B is determined by scalar product and vectors normalization using the Formula (1).

$$\text{similarity} = \frac{A \times B}{AB} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \tag{1}$$

Cluster analysis grouped the skills defining workload sets. As a group designation, the cluster central element was chosen (cluster centers are marked in red). In Fig. 4 there is an example of clustering for Moscow.

In the region, 79 skills were identified and grouped into 11 clusters by the algorithm. It is noteworthy that the clusters overlap as a relatively low Silhouette coefficient (0.363) shows. Therefore, it is advisable to conduct an additional study using fuzzy algorithms such as fuzzy c-means.

The algorithm has identified the following groups by their central skills:

1. PHP	7. Android
2. C + +	8. JavaScript
3. PostgreSQL	9. ORACLE
4. iOS	10. JSON API
5. IC programming	11. Spring Framework

(continued)

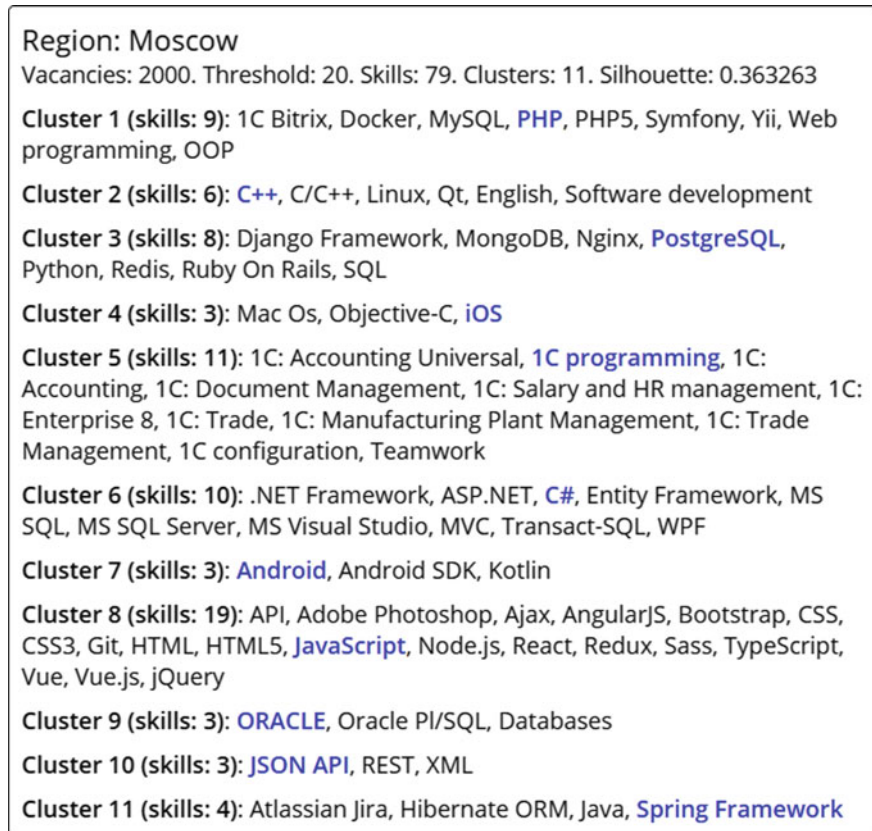


Fig. 4 Cluster analysis result for Moscow

(continued)

6. C#	
-------	--

This approach involves a simple hierarchy with two levels, and an instance can belong to only one class at a time. It is worth pointing out that the centers and the number of clusters vary depending on the region.

Comparing both approaches to the skill structuring for Moscow, we can see that the ontological model provides a multilevel hierarchy of classes with a strict organization system. Another advantage of structuration via an ontology is the possibility of assigning one element to multiple classes while clustering is intended to create non-crossing sets and therefore does not provide such an opportunity. Nevertheless, developing an ontological model for each region requires significant resources.

Cluster analysis does not allow to identify the vertical hierarchy of the skills within the domain and to highlight the deep implicit dependencies between the skills, but the identification of the key element of the workload sets yields sufficiently good results regarding horizontal skill integration. For instance, the group in Moscow clusters determined by the “1C Programming” skill includes all the skills used for the 1C programming workload. The advantage of this approach is the automation of the group selection process, which reduces the data structuring labor intensity. The approaches differ in the principles that define a skill belonging to a class: in cluster analysis, the group includes a skill that is more similar to all the elements of a given cluster; in ontology, the class is an abstract concept to which the selected skills assigned.

Furthermore, it has to be emphasized that the ontological model is structurally more complex and the structure itself is stricter. Grouping in ontology occurs by category, with a specific category not being a set of skills that are required for a particular field of software engineering. On the contrary, the cluster approach groups skills according to a specific field of IT specialist expertise, the structure of such a model is more dynamic and flexible. A cluster may contain a set of skills that a specialist requires for a particular position.

The heterogeneous nature of the IT field leads us to emphasize the necessity of specifying a region during clusters and ontologies creation. To prove this, let us give an example of clustering results for the Sverdlovsk Oblast region (Fig. 5).

There were 73 skills identified in the Sverdlovsk Oblast region, and the algorithm formed 13 clusters with the following centers:

1. 1C programming	8. iOS
2. JavaScript	9. Spring Framework
3. Qt	10. Java SE
4. PHP	11. Android
5..NET Framework	12. Django Framework
6. SAP	13. MongoDB
7. Project management	

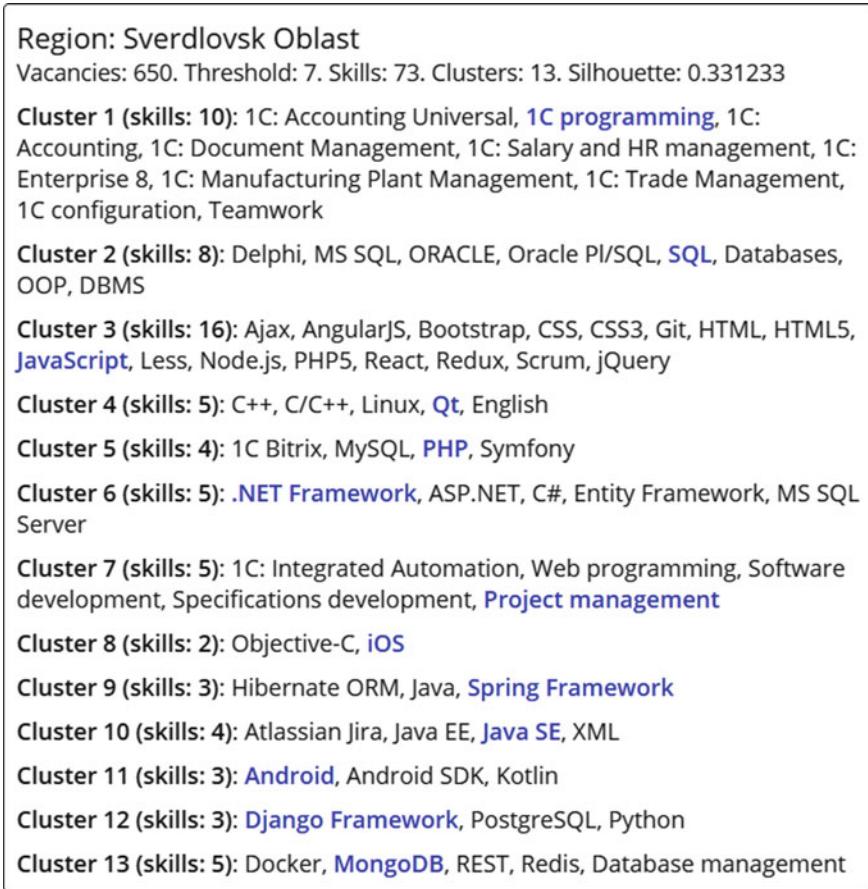


Fig. 5 Cluster analysis for the Sverdlovsk oblast region

Several clusters and their centers in the Sverdlovsk Oblast overlap with those in Moscow. There are even fully matching clusters, for instance, cluster 7 in Moscow and cluster 11 in the Sverdlovsk Oblast have both the same central elements and the same list of skills.

But there are also differences, for example, clusters with the “PHP” central element have different sizes: the set of skills of the Sverdlovsk Oblast region is a subset of Moscow skills. Also, there is no “JSON API” skill in the Sverdlovsk Oblast region, and in Moscow, this skill is the center of cluster 10.

According to this, one can assume that the workloads set of the Sverdlovsk Oblast are a subset of the workloads set of Moscow. However, a more detailed analysis of the cluster structure reveals non-overlapping skills and clusters, for instance, in the Sverdlovsk Oblast there is the “Development of technical tasks” skill, which is not

present in Moscow clusters. Thus, it is confirmed that it is necessary to analyze a set of workloads for each region individually to provide the most complete representation of the local companies demand IT specialists possessing certain professional competences.

To evaluate the quality of the identification of workloads sets for different countries, vacancies in the German federal states were also analyzed. The primary data source was the “Monster.de” job aggregator [20]. The query “Programmierer” was used for data retrieval for each federal state individually. In Fig. 6 the workloads set identified for the federal state of Saxony is presented.

In Saxony, 98 skills were divided into 18 clusters. Noteworthy that in Moscow, 2000 vacancies were analyzed and 79 skills were identified, while in Saxony 98

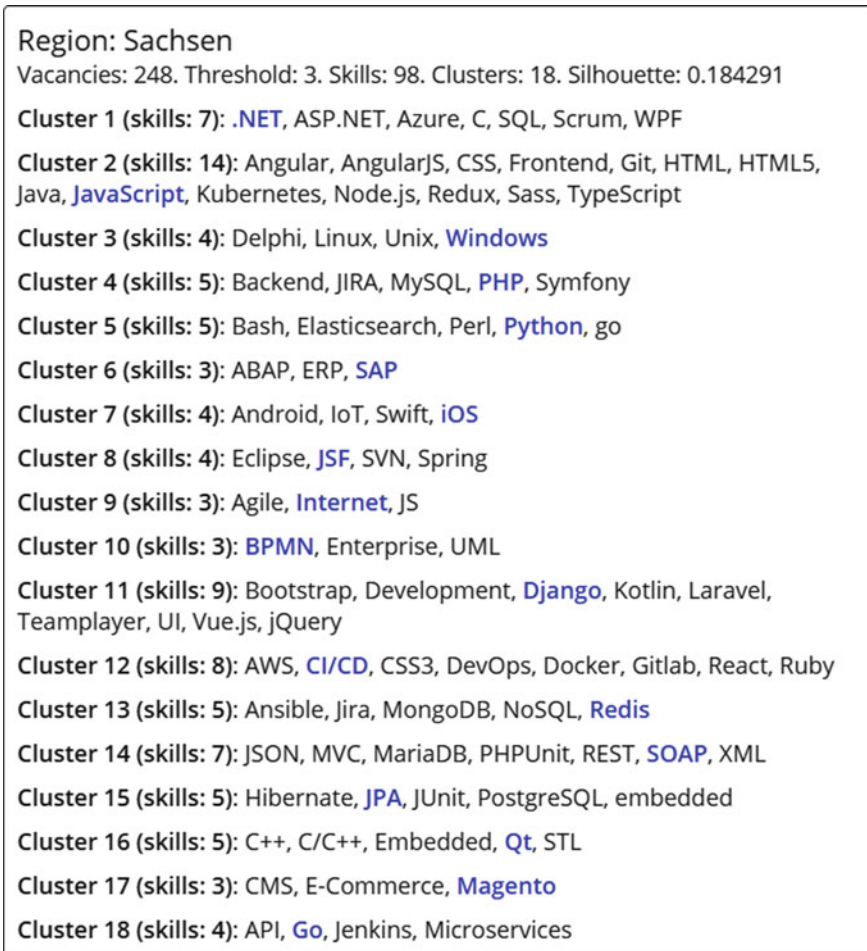


Fig. 6 Cluster analysis for the federal state of Saxony

unique skills were extracted out of 248 vacancies. Hence, there are more vacancies similar to each other in Moscow. It should also be noted that there is more fragmentation than in Moscow. The average cluster size for Moscow is 7.18, and for the federal state of Saxony, it is 5.44. The Silhouette coefficient is even lower than in Moscow (0.184) which leads us to the conclusion that there is indeed a need for further studies using fuzzy methods.

The algorithm has identified the following groups by their central skills:

1. NET	10. BPMN
2. JavaScript	11. Django
3. Windows	12. CI/CD
4. PHP	13. Redis
5. Python	14. SOAP
6. SAP	15. JPA
7. IOS	16. Qt
8. JSF	17. Magento
9. Internet	18. Go

The group formation principle is the same. Let us also remark that Moscow and Saxony both have “PHP”, “iOS” and “JavaScript” cluster centers, with “JavaScript” clusters being relatively similar. However, the differences between the clusters as a whole are significant, which once again demonstrates that a skills structure should be created taking into consideration country and region specifics.

3 Conclusion

Both of the outlined approaches quite effectively structure the required skills. For strict formalization and identification of top-level skills, the ontological model is more appropriate. At the same time, the cluster analysis allows to identify a group of skills needed in a particular programming area and is more suitable for creating a list of a particular specialist professional competencies. A combination of these methods will allow, on the one hand, to form a workload-based list of the required skills, and on the other hand, to form a complete list of skills based on the professional competencies’ hierarchy within the domain.

Therefore, we can suggest a possible approach to an integrated model combining the advantages of both ontology and cluster analysis. We assume that it can be an automated interpretation of the cluster based on its center and the skills closest to it and its further integration into a broader hierarchy of the domain ontology.

It is easy enough to interpret cluster 7 for Moscow (Fig. 4) as a group of skills required for an Android OS application developer. Describing the cluster as “an Android developer skills”, we at the same time integrate it into the ontological model: the “Android” skill is linked to the “Operating systems” class, “Kotlin” is part of

the “Programming languages” class, and “Android SDK” belongs to development environments. Consequently, the requirements for general competencies related to operating systems, programming languages, as well as familiarity with development environments are implicitly imposed on an Android developer. Based on the given interpretation, it is possible to provide an individualized training program aimed at improving the professional level in these particular areas.

The clusters formed for the German federal states can be interpreted similarly. For example, for Saxony, cluster 16 (Fig. 6) can be interpreted as “a C/C++ developer skills”. This cluster includes the programming languages themselves (“C” and “C++”), a standard template library for the C++ programming language (“STL”), a cross-platform framework for software development in C++ (“Qt”), a specialized microprocessor-based monitoring and control system that is compatible with Qt (“Embedded”). Thus, this cluster defines C/C++ developer skills as proficiency in programming languages and libraries, the ability to use the frameworks, and the ability to work with embedded systems.

The integrated model implies regional localization of workloads sets. In this case, in order to simplify the model, it is possible not to fragment the ontology by region. It is sufficient to create a unified ontology for all the skills. The skills required for a particular IT area in the region can be drawn from clusters, with the missing skills identified using the different hierarchy levels of the general ontological model.

The proposed approach may become the basis for the decision support system both in the field of human resources management and specialists training.

References

1. Popova, T.N.: (2011) Structural imbalance of the employment system in the region. *Modern Econ. Probl. Trends Persp.* **5**, 1–6 (2011)
2. Bondarenko N.V.: The nature of the current and expected shortage of workers’ professional skills and qualities on the Russian labor market. *Public opinion bulletin. Data. Anal. Dis.* **3–4**(116), 34–46 (2013)
3. Cedefop: Insights into skill shortages and skill mismatch: learning from Cedefop’s European skills and jobs survey, p. 106, 107. Publications Office, Cedefop Reference Series, Luxembourg (2018)
4. IT: Job Market Overview and Top 15 Professions; <https://perm.hh.ru/article/24562>
5. Zemnukhova, L.V.: (2013) IT workers on the labor market. *Soc. Sci. Technol.* **4**(2), 77–90 (2013)
6. Eurostat Statistic Explained: ICT Specialists in Employment. Eurostat Statistic Explained; https://ec.europa.eu/eurostat/statistics-explained/index.php/ICT_specialists_in_employment#Number_of_ICT_specialists (2019)
7. HeadHunter API; <https://github.com/hhru/api>.
8. Jurafsky D., Martin J.H.: *Speech and Language Processing* 3rd ed. draft, 613 p (2019)
9. Manning, C.D., Raghavan, Schütze, H.: *Introduction to Information Retrieval*; <https://nlp.stanford.edu/IR-book/html/htmledition/irbook.html> (2018)
10. Gruber, T.R.: (1993) A translation approach to portable ontology specifications. *Knowl. Acquis.* **5**(2), 199–220 (1993)
11. Uschold, M., Gruninger, M.: (1996) *Ontologies: principles, methods and applications*. *Knowl. Eng. Rev.* **11**(2), 93–136 (1996)

12. Faizrakhmanov R.A., Yarullin D.V.: Web-data driven ontological approach to modelling IT specialists recruitment needs. In: Proceedings of 2019 20th IEEE International Conference on Soft Computing and Measurements (SCM), pp. 252–255 (2019). <https://doi.org/10.1109/SCM.2019.8903715>
13. Kurushin D.S., Leonov E.R., Soboleva O.V.: A possible approach to automatic construction of the hypertext denotation graph. Information Structure of the Text, pp. 113–118. RAS.INION, Moscow (2018)
14. Kim S., Gil J.: Research paper classification systems based on TF-IDF and LDA schemes. Hum.-Cent. Comput. Info. Sci. **9**(30) (2019)
15. Zhang, Y., Jin, R., Zhou, Z.: (2010) Understanding bag-of-words model: a statistical framework. Int. J. Mach. Learn. Cybern. **1**, 43–52 (2010)
16. Bird, S., Loper, E., Klein, E.: Natural Language Processing With Python. O’Reilly Media Inc., 502 p (2009)
17. Thavikulwat, P.: (2008) Affinity propagation: a clustering algorithm for computer-assisted business simulations and experiential exercises. Develop. Bus. Simul. Experient. Lear. **35**, 220–224 (2008)
18. Frey, B.J., Dueck, D.: (2007) Clustering by passing messages between data points. Science **315**, 972–976 (2007)
19. Han J., Kamber M., Pei J.: Data Mining: Concepts and Techniques 3rd ed. 703 p. Elsevier (2012)
20. Monster Job Search API; <https://partner.monster.com/job-search>

Numerical Modeling of Business Processes Using the Apparatus of GERT Networks



Mikhail Dorrer , Alexandra Dorrer , and Anton Zyryanov 

Abstract The chapter solves the problem of numerically predicting the parameters of a business process model (using the cost example) without using multi-pass simulation models. To solve this problem, an apparatus of stochastic GERT networks is used. The edge of the GERT network is associated with the operation of the business process, and the node of the GERT network is associated with the event or branching of the business process. An algorithm for translating a business process model into an equivalent GERT network is given, as well as calculating the parameters of the business process cost distribution law based on the resulting GERT network. The proposed approach allows us to solve the problems of predicting the dynamics of discrete-event models described in standard model notations (IDEF3, ARIS EPC) without using multi-pass simulation models.

Keywords Business process · Simulation model · ARIS eEPC · GERT

1 Introduction

Markets of goods and services in the modern economy are highly competitive. The condition for the survival of enterprises in the competition is their effectiveness and ability to quickly change production and management processes. The answer to these requirements in management technologies is the application of a process approach. The process approach is the basis of such standards as ISO 9001: 2015 [1], ISO 12207: 2017 [2]. Best practices in business process management are outlined in BPM CBOK [3].

M. Dorrer (✉) · A. Dorrer

Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, prospect
“Krasnoyarsky rabochoy” 31, Krasnoyarsk, Russia
e-mail: mdorrer@mail.ru

A. Zyryanov

Individual Entrepreneur, Krasnoyarsk, Parashyutnaya street 14, 58., Krasnoyarsk, Russia
e-mail: ZyryanovAntonA@gmail.com

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced
Human-Centered Society*, Studies in Systems, Decision and Control 333,
https://doi.org/10.1007/978-3-030-63563-3_5

Most business process management methodologies include one version or another of a continuous improvement cycle. An important role is played by the continuous improvement cycle in Kaizen philosophy [4, 5]. The classic quality manager W. Deming developed and described the PDCA continuous improvement cycle [6, 7]. Continuous improvement cycles are also used in more general tasks. So in Ackoff's works, a cyclic procedure for resolving problem situations in applied system analysis is proposed [8, 9]. According to [10], the business process management cycle includes the stage of "Simulation and Analysis". The high-level model obtained at this stage is used to predict the behavior of the system when performing various scenarios to detect critical sections and bottlenecks. The results of the analysis are used to configure the process before its implementation. However, this stage of the analysis is characterized by both great laboriousness and high computational complexity. In this work, the authors set themselves the goal of demonstrating a technology that provides a solution to the problems of analysis and forecasting of socio-economic objects, but without this drawback, using a substantial example.

The problem of the complexity of simulation in this chapter is proposed to be solved by using the apparatus of GERT-networks.

Computer simulation models make it possible to evaluate in measurable terms the consequences of changing business processes, to predict with the help of a computational experiment how the "image of the future", the "to-be" model, will behave. The simulation model allows you to identify potential problems associated with the proposed improvement, to build a forecast of the dynamics of the system.

Simulation, therefore, is a powerful tool for studying the behavior of real systems. Moreover, the simulation itself does not solve optimization problems, but rather is a technique for assessing the values of the functional characteristics of the simulated system, allowing you to identify problem areas in the system [11].

However, imitation is inherently a random process. Therefore, any result obtained by simulation is subject to random fluctuations and, therefore, as in any statistical experiment, should be based on the results of relevant statistical checks [11]. To eliminate these negative features of the approach, the urgent task is to replace simulation experiments with analytical models.

With the accumulation of experience and statistics in the field of complex systems research, alternative stochastic networks, in particular PERT and GERT, are becoming more widespread. Model PERT (Project evaluation and review technique) [12] is used to model projects and programs, and GERT (Graphical evaluation and review technique) [13] is used to model technological and business processes.

A detailed description of GERT networks is presented by Phillips [14], Neumann [15], Pritsker [13]. A significant contribution to the development of the apparatus of GERT networks was made by Alexander Shibanov [16].

Attempts to study business processes based on GERT networks were made by Barjis [17], Aytulun [18]. Barjis and Dietz [17] use the DEMO methodology developed by them based on the BPM standard to model business processes. Neither Barjis and Dietz [17] nor Aytulun [18] apply any of the most widely used business process description methodologies in the world (IDEF, ARIS, BPMN). In these works, there is no description of formal methods and algorithms for converting business process

models to GERT-network models; the features of modeling business processes by GERT-networks are not investigated. One of these features is the study of financial and resource flows of business processes, since the study of the probability-time characteristics of processes is not basic, although it provides a lot of useful information.

The article [19] proposes the use of GERT networks to construct an input–output table for a carbon fiber production chain in accordance with the input–output theory.

Article [20] is devoted to an overview of the application of GERT networks to management problems. The authors demonstrate the results of using the GERT network to analyze a hypothetical R&D project. The chapter provides an assessment of improving the efficiency of project planning, workload, resources, and equipment of the project.

Thus, the modeling of business processes based on GERT networks is a poorly studied topic, therefore, additional research is relevant. The authors have already attempted to work with the use of GERT-networks for the analysis of business processes in [21], but the idea to show the application of the same approach to assessing the cost parameters of a business process seemed interesting.

Representation of business processes in the form of a GERT-network will allow for research related to the forecast of business process dynamics. In particular, it is possible to determine the probability density function of runoff performance over time and resources, as well as the required central distribution moments – expectation, variance.

2 Methods

GERT networks are a variant of semi-Markov models, but the random variables in them are characterized not only by dispersion but also by the distribution law. GERT-networks allows you to include random deviations and uncertainty that occur directly during the execution of each individual work [13]. The execution of work (operation) in the system is associated with the branches (arcs) of the GERT-network, which are characterized by additive random variables. To calculate the output characteristics of GERT networks, the generating functions of the moments of random variables are used. Activation of each subsequent branch is generally probabilistic.

A GERT network can be described by a directional weighted graph.

$$G = (V, E),$$

where V is the set of vertices (nodes); E many directed edges (arcs).

GERT network nodes are interpreted as system states, and arcs as transitions from one state to another. Such transitions are associated with the implementation of generalized operations characterized by the density of distribution and probability of completion.

Thus, a GERT network is a network with sources R and sinks S of the “work on an arc” type, in which each node belongs to one of six types of nodes [13], for each arc $\langle i, j \rangle$ a weight of the form $[p_{ij}, F_{ij}]$ is defined, where p_{ij} is the conditional probability of the arc $\langle i, j \rangle$ execution subject to activation of the node i , F_{ij} is the conditional distribution function of some random variable.

The calculation of the parameters of the GERT network represents the finding of the first central moments of the distribution of the random value of the network. In particular, the first and second central moments of a random variable are found—the mean and variance, respectively. In addition, for some tasks, it is important to find the probability of the network flow and the distribution function of the random variable of the entire network.

Methods for calculating the parameters of a GERT network are described in [14].

3 Results

Consider the ARIS eEPC model of the “Product Manufacturing” business process (Fig. 1) and the GERT network that corresponds to this business process. The task of translating a business process model into a GERT network model was considered in [22]. We also used assumptions about the laws of the probability distribution of the parameters of business process operations given in [23].

Table 1 and Fig. 1 present a comparison of elements of a business process model and a GERT network.

Note that in the business process under consideration there are two initial events. Each of the events initiates the start of the process with some probability: “Order for a standard product”—60%, “Order for a non-standard product”—40%. Therefore, for the GERT-network model, we introduce the network source V_0 and the arcs connecting the source with the vertices V_1 and V_2 .

In Table 2 are presented the parameters characterizing the arcs of the GERT-network, according to the additive parameter—the financial costs of the function. To simplify the notation, we introduced the index k , which replaces the indices ij for the characteristics, corresponding to the arc.

Thus, a GERT-network model is obtained that fully displays the model of the system under study, and for each arc of the network, the conditional probability and the generating function of the moments are determined.

Next, you need to close the GERT network with an arc W_A , leading from node V_{15} to node V_1 .

Replacing $W_A(s)$ by $\frac{1}{W_E(s)}$, we obtain the following transmittances for the network loops.

Loops of the first order: $W_8(W_9W_{11} + W_{10}W_{12})W_{13}W_{14}W_{15}, (W_1W_3 + W_2W_4W_5W_6)W_7W_8(W_9W_{11} + W_{10}W_{12})W_{13}W_{16}W_{17}W_{18}\left(\frac{1}{W_E}\right).$

There are no loops larger than the first order in this GERT network.

Using the topological Mason equation [7], we obtain:

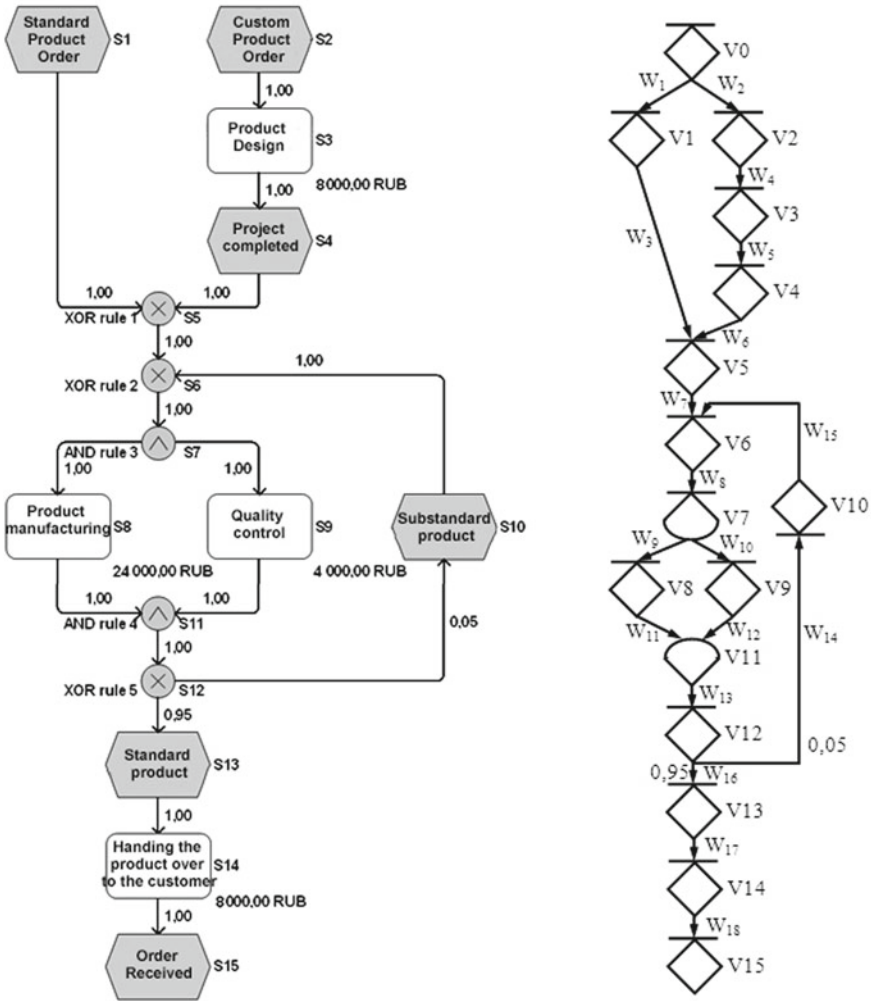


Fig. 1 ARIS eEPC model and its corresponding GERT network

$$\begin{aligned}
 H = & 1 - W_8(W_9 W_{11} + W_{10} W_{12}) W_{13} W_{14} W_{15} - (W_1 W_3 + W_2 W_4 W_5 W_6) \times \\
 & \times W_7 W_8 (W_9 W_{11} + W_{10} W_{12}) W_{13} W_{16} W_{17} W_{18} \left(\frac{1}{W_E} \right) = 0
 \end{aligned}$$

Table 1 Comparison of model objects

Model eEPC object	Model eEPC object type	GERT network node	GERT network node type
S ₁	Event	V ₁	STEOR
S ₂	Event	V ₂	STEOR
S ₃	Process	V ₃	STEOR
S ₄	Event	V ₄	STEOR
S ₅	Logical Connector	V ₅	STEOR
S ₆	Logical Connector	V ₆	STEOR
S ₇	Logical Connector	V ₇	[EOR, DT]
S ₈	Process	V ₈	STEOR
S ₉	Process	V ₉	STEOR
S ₁₀	Event	V ₁₀	STEOR
S ₁₁	Logical Connector	V ₁₁	[AND, ST]
S ₁₂	Logical Connector	V ₁₂	STEOR
S ₁₃	Event	V ₁₃	STEOR
S ₁₄	Process	V ₁₄	STEOR
S ₁₅	Event	V ₁₅	STEOR

Transforming this expression, we get:

$$W_E(s) = (W_1 W_3 + W_2 W_4 W_5 W_6) W_7 W_8 (W_9 W_{11} + W_{10} W_{12}) \\ \times \frac{W_{13} W_{16} W_{17} W_{18}}{1 - W_8 (W_9 W_{11} + W_{10} W_{12}) W_{13} W_{14} W_{15}}$$

which $W_E(s)$ is the equivalent W -function for a GERT network.

Substituting the values of the probabilities and generating functions of the moments from Table 2, we find the value $W_E(0)$ and then calculate the first central moment of the distribution relative to the origin.

The mathematical expectation and the variance of the drain of the GERT network are calculated in this way:

$$\mu = 12.879$$

$$\sigma^2 = 9.211$$

Table 2 Parameters of the GERT network

Arc (i, j)	k	p_k	$M_k(s)$	$W_k(s)$
$\langle V_0, V_1 \rangle$	1	0.6	1	0,6
$\langle V_0, V_2 \rangle$	2	0.4	1	0,4
$\langle V_1, V_5 \rangle$	3	1	1	1
$\langle V_2, V_3 \rangle$	4	1	1	1
$\langle V_3, V_4 \rangle$	5	1	$(1 - 2s)^{-1}$	$(1 - 2s)^{-1}$
$\langle V_4, V_5 \rangle$	6	1	1	1
$\langle V_5, V_6 \rangle$	7	1	1	1
$\langle V_6, V_7 \rangle$	8	1	1	1
$\langle V_7, V_8 \rangle$	9	1	1	1
$\langle V_7, V_9 \rangle$	10	1	1	1
$\langle V_8, V_{11} \rangle$	11	1	$\exp(10s + 0.0003125s^2)$	$\exp(10s + 0.0003125s^2)$
$\langle V_9, V_{11} \rangle$	12	1	$\exp(s + 0.00005s^2)$	$\exp(s + 0.00005s^2)$
$\langle V_{11}, V_{12} \rangle$	13	1	1	1
$\langle V_{12}, V_{10} \rangle$	14	0.05	1	0.05
$\langle V_{10}, V_6 \rangle$	15	1	1	1
$\langle V_{12}, V_{13} \rangle$	16	0.95	1	0,95
$\langle V_{13}, V_{14} \rangle$	17	1	1	1
$\langle V_{14}, V_{15} \rangle$	18	1	$\exp(0.5s + 0.0001125s^2)$	$\exp(0.5s + 0.0001125s^2)$
$\langle V_{15}, V_1 \rangle$	W_A	1	1	1

Figure 2 shows the probability density, the mathematical expectation, and the distribution function of the calculated random variable (financial costs for the business process) of the GERT network. Figure 2 shows the probabilistic forecast of the dynamics of the event model of the business process “Product manufacturing” (Fig. 1) based on the GERT network. Thus, we can say that with a probability of

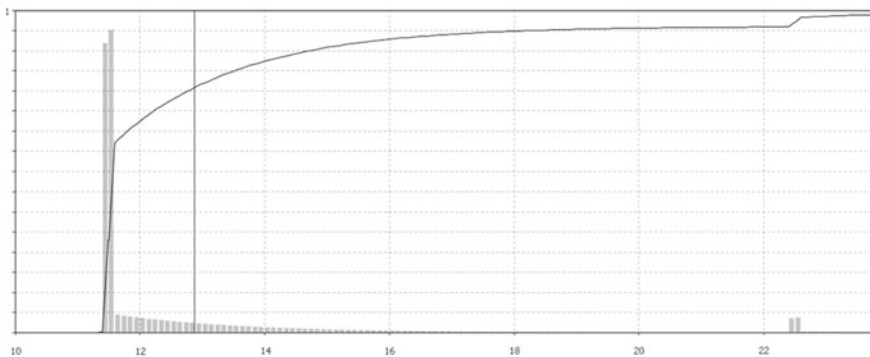


Fig. 2 Probability density and distribution function of the GERT-network (in thousand rubles)

90%, the costs of the business process will be in the range of 11.500 thousand rubles up to 16.000 thousand rubles.

4 Conclusion

Thus, an approach is proposed that allows solving the problems of predicting the dynamics of discrete-event models described in standard model notations (IDEF3, ARIS EPC). The study of business processes based on GERT-networks can provide a lot of useful information about the behavior of the studied system. Unlike simulation experiments, this method is analytical, and therefore has several advantages compared to a simulation experiment—less computational complexity, obtaining a result that does not depend on the randomness factor.

The next steps in the development of this work may be to check the sensitivity of central moments in time and cost to various model parameters (probabilities, laws, and parameters of the distribution of time and cost for individual operations) with access to recommendations for optimizing the investment of resources in improving the business process as a whole.

References

1. International Organization for Standardization: ISO 9001:2015 Quality management systems—Requirements. 0310070 Manag Syst (2015)
2. ISO/IEC 12207: Systems and Software Engineering— Software Life Cycle Processes (2008)
3. Tony, B., Nancy, B., Phil, V., Emmett, P., Dan, M., Marc, S., Denis, L., Gabrielle, F., Todd, L., Raju, S., Michael Fuller, J.F.: ABPMP CBOK Guide (2013)
4. Ishikawa, K.: What is Total Quality Control : The Japanese Way. New York (1981)
5. Imai, M.: Kaizen: The Key to Japan’s Competitive Success (1986)
6. Deming, W.E.: Out of the Crisis: Quality, Productivity and Competitive Position. Massachusetts Institute of Technology. Center for Advanced Engineering Study (1986). <https://doi.org/10.1002/qre.4680020421>
7. Propst, A.L., Deming, W.E.: The New Economics: For Industry. Technometrics, Government, Education (1996) <https://doi.org/10.2307/1270625>
8. Greenfield, T., Ackoff, R.L.: The Art of Problem Solving (Accompanied by Ackoff’s Fables). Stat (1979). <https://doi.org/10.2307/2987880>
9. Ackoff, R.L.: Re-creating the corporation (1999)
10. Nainani, B.: Closed loop BPM using standards based tools (2004)
11. Ord, K., Taha, H.A.: Operations research: an introduction. Oper Res Q (1972). <https://doi.org/10.2307/3008276>
12. Kerzner, H.: A systems approach to planning scheduling and controlling (20170)
13. Pritsker, A.A.B.: GERT: Graphical Evaluation and Review Technique. RAND Corporation, Santa Monica, CA (1966)
14. (1982) Fundamentals of network analysis, by Don T. Phillips and Alberto Garcia-Diaz, Prentice Hall, Englewood Cliffs, NJ, 1981, 474 pp. Price: \$26.90. Networks 12:209–210. <https://doi.org/10.1002/net.3230120210>
15. Neumann, K.: Stochastic Project Networks. Springer, Berlin Heidelberg, Berlin, Heidelberg (1990)

16. Koryachko, V., Shibanov, A., Shibanov, V., et al.: Hierarchic GERT networks for simulating systems with checkpoints. In: 2017 6th Mediterranean Conference on Embedded Computing (MECO), pp. 1–4. IEEE (2017)
17. Barjis, J., Dietz, J.L.G.: Business Process Modeling and Analysis Using Gert Networks. In: Enterprise Information Systems, pp. 71–80. Springer, Netherlands, Dordrecht (2000)
18. Aytulun, S.K., Guneri, A.F.: Business process modelling with stochastic networks. *Int J Prod Res* **46**, 2743–2764 (2008). <https://doi.org/10.1080/00207540701543601>
19. Liu, X., Fang, Z., Zhang, N.: A value transfer GERT network model for carbon fiber industry chain based on input–output table. *Cluster Comput* **20**, 2993–3001 (2017). <https://doi.org/10.1007/s10586-017-0960-y>
20. Ambika, S., Indhumathi, R., Mytherae, R., Spandana, R.: Application of Gert Analysis in Management. *Int J Latest Eng Manag Res* **03**, 01–04 (2018)
21. Dorrer, M., Dorrer, A.: Forecasting e-Learning Processes Using GERT Models and Process Mining Tools. In: Solovev D., Savaley V., Bekker A. P V. (eds) Proceeding of the International Science and Technology Conference “FarEastC on 2019”. Smart Innovation, Systems and Technologies, pp. 857–866. Springer, Singapore (2020)
22. Zyryanov, A.A., Dorrer, M.G.: The algorithm of business process model translation into the GERT-network model. *Bull KrasGAU* 13–18 (2012)
23. Golenko-Ginzburg, D.: Stochastic network models in innovative projecting. Science Book Publishing House, Yelm, WA, USA (2014)

Models and Methods of Forecasting and Tasks Distribution by Performers in Electronic Document Management Systems



Sofia S. Kildeeva, Alexey S. Katasev, and Nafis G. Talipov

Abstract This paper describes the problem of task distribution received through the electronic document management system. The fuzzy-production model underlying the solution to this problem is described. Based on the proposed model, a software package was developed for decision-making support of task performers selection, its structure is presented. A model of task distribution is considered taking into account its forecast values. The basic steps in forecasting model construction are described. The effectiveness of this approach for task distribution, based on workload indicators of specialists with different levels of working capacity and qualifications, is shown.

Keywords Electronic document management system · Task distribution · Fuzzy production model · Forecasting · Data mining

1 Introduction

Currently, electronic document management systems (EDMS) are widely used in many fields of human activity [1–3]. The use of such systems can improve the efficiency of working with documents by reducing the time for making managerial decisions and ensuring quality control of performance discipline. However, due to a large number of incoming tasks of various difficulty levels, the problem of their rational distribution among performers arises [4]. Often for solving this problem, an expert approach is used, which is effective in terms of the quality of managerial decision-making. However, in the absence of an expert, tasks distribution difficulties

S. S. Kildeeva (✉) · A. S. Katasev · N. G. Talipov
Kazan National Research Technical University named after A.N. Tupolev-KAI, str., 10, 420111
Kazan, K.Marx, Russia
e-mail: Sofi-pi@mail.ru

A. S. Katasev
e-mail: Kat_726@mail.ru

N. G. Talipov
e-mail: Nafis.Talipov@mail.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_6

arise. To eliminate this drawback, the development of effective models, methods, and technologies for assigning tasks to performers is relevant [5]. Such technology should not only allow distributing the tasks received for execution but solve this problem taking into account the forecasting of the possible number of tasks with various categories of complexity.

As an example of the EDMS, consider the existing electronic document management system of the territorial office of The Federal Service for Supervision of Communications, Information Technology, And Mass Media (Roskomnadzor). In the field of protecting the rights of personal data subjects, the most time-consuming and urgent task, for automated decision support, is maintaining a register of personal data operators (PDO), that is, timely updating information about PDO previously included in the register, and monitoring the provision of the corresponding notification by unregistered PDO, which is the implementation of the legislation of the Russian Federation in the field of personal data in general.

2 Statement of the Tasks Distribution Objective for Maintaining the Register of Personal Data Operators

Consider the formal statement of the tasks distribution objective between performers when maintaining the register of PDO [6]. Let $Z = \{z_1, z_2, \dots, z_N\}$ be a set of tasks of volume N . Each incoming assignment can be classified according to a specific level of difficulty. Let's highlight the following levels of complexity of incoming tasks corresponding to different categories of PDO:

- (1) S_1 —"Low" (this level of complexity includes tasks coming from such PDOs, such as "physical person" and "private entrepreneur");
- (2) S_2 —"Medium" (this level of complexity includes tasks coming from such PDOs, such as "juridical entity");
- (3) S_3 —"High" (this level of complexity includes assignments from such PDOs such as "government agencies and municipalities").

Let $A = \{a_1, a_2, \dots, a_n\}$ is a set of performers who are processed incoming tasks. It should be noted that the number and composition of performers who are included in a given set may be subject to changes over time. For each potential executor of the received task, the following characteristics can be determined: the level of workload ($C1$), efficiency ($C2$), and the level of qualification ($C3$). It is necessary to carry out a rational distribution of all N tasks included in the set of received tasks Z between potential performers from set A . In this case, the individual characteristics of each performer should be taken into account in terms of the entered characteristics.

In the course of the analysis, it was found that the most rational method for solving the problem is the method of fuzzy inference [7–9], which is based on fuzzy rules [10, 11]. To implement it, it was necessary to solve the following tasks:

- (1) selection of the type of fuzzy production rules for making decisions on the distribution of tasks by the performer;
- (2) development of a methodology for constructing a system of fuzzy rules;
- (3) development of an inference algorithm on a system of fuzzy rules;
- (4) development of a method for constructing membership functions in fuzzy rules;
- (5) development of a method for determining the reliability of fuzzy rules;
- (6) development of a model for accounting for predicted values of the number of incoming tasks when they are distributed among performers;
- (7) development of a software package for decision-making support for the distribution of tasks.

Let's consider the solution of these tasks in more detail.

2.1 The Type of Fuzzy Production Rules for Making Decisions on the Distribution of Tasks

Fuzzy-production rules underlie the model of the knowledge representation of an expert on the distribution of tasks between performers, taking into account their special aspects and characteristics. Within the framework of solving the problem under consideration, the following type of fuzzy rules was chosen [12]:

$$IF \wedge (x_1 \text{ is } \tilde{A}_1, \dots, x_n \text{ is } \tilde{A}_n, x_{n+1} \text{ is } A_{n+1}) \Rightarrow y = a_i [CF_i], \quad (1)$$

where $x_i, i = \overline{1, n}$ —workload of the i -th performer; x_{n+1} —task difficulty; $\tilde{A}_i = \{x_i, \mu_{\tilde{A}_i}(x_i)\}, i = \overline{1, n}$ —fuzzy gradations of the workload of performers; $\mu_{\tilde{A}_i}(x_i) \in [0; 1]$ —the degree of x_i belonging to \tilde{A}_i ; A_{n+1} —the value of the complexity of the task from the set $\{S_1, S_2, S_3\}$; y —output variable that defines the executor of the task; $a_i, i = \overline{1, n}$ —a specific performer from $\{a_1, a_2, \dots, a_n\}$; CF_i —the utility of choosing the i -th performer.

The workload of the performers and the complexity of the task are the input parameters of the fuzzy rule. In this case, specific task executors act as output parameters. A feature of this type of rule is the use of the utility parameter CF_i of the choice of the i -th performer. Thus, the rules of the form (1) reflect the logic of an expert when making a managerial decision on the distribution of tasks between performers [13–15].

2.2 Development of a Method for Constructing a System of Fuzzy Production Rules

When deciding on the selection of a performer for a specific task, the total number, and composition of potential performers are taken into account. Let consider the developed technique for constructing a system of fuzzy production rules for a specific number and composition of task performers. This technique includes the following steps [6]:

- (1) assignment of many task performers $A = \{a_1, a_2, \dots, a_n\}$;
- (2) setting the number m and names of gradations that determine the workload indicator of potential task performers (for example, for $m = 3$, the gradations can be designated as $\tilde{A}_1 = \text{“low workload”}$, $\tilde{A}_2 = \text{“medium workload”}$, $\tilde{A}_3 = \text{“high workload”}$);
- (3) construction of all possible combinations of the values of the input parameters (x_i, x_{i+j}) , which are responsible for the workload of potential performers and the complexity of the assigned task, and the output parameter (y) , which determines who of the potential performers is selected to perform it. Given that the number of task difficulty values is three, the number of possible combinations is calculated using the following formula:

$$N = 3 mn^2 \quad (2)$$

where m —the number of grades of the complexity of incoming tasks, n —number of tasks.

Thus, each fuzzy-production rule corresponds to a combination of input conditions that determine the rate of the workload of potential performers and the complexity of a specific task, and an output value that determines which of the potential performers is selected to complete it.

Using the developed technique, it is possible to draw up a system of fuzzy production rules for a specific number and composition of performers. This system of rules has the following form (3):

$$\left\{ \begin{array}{l} \text{If } \wedge (x_1 \text{ is } \tilde{A}_1^j, \dots, x_n \text{ is } \tilde{A}_n^j, x_{n+1} \text{ is } A_{n+1}^k) \Rightarrow y = a_1 \text{ [CF}_1\text{]} \\ \text{If } \wedge (x_1 \text{ is } \tilde{A}_1^j, \dots, x_n \text{ is } \tilde{A}_n^j, x_{n+1} \text{ is } A_{n+1}^k) \Rightarrow y = a_2 \text{ [CF}_2\text{]} \\ \dots \\ \text{If } \wedge (x_1 \text{ is } \tilde{A}_1^j, \dots, x_n \text{ is } \tilde{A}_n^j, x_{n+1} \text{ is } A_{n+1}^k) \Rightarrow y = a_n \text{ [CF}_n\text{]} \end{array} \right. \quad (3)$$

where $j = \overline{1, m}$ determines the value of the performer's workload, $k = \overline{1, 3}$ —the difficulty of the task.

It should be noted that for a different number and composition of performers, it is necessary to form an individual system of rules. To solve the task of choosing

a task performer, an inference algorithm has been developed on a system of fuzzy production rules.

2.3 Development of an Inference Algorithm Based on a System of Fuzzy Production Rules

To determine the specific executor of the received task, an inference algorithm was developed based on the rules of a fuzzy production model. The following indicators are calculated for each rule [6]:

- (1) a confidence level of the antecedent of rule $V \in [0;1]$ (*veracity*):

$$V = \min\left(\mu_{\bar{A}_1^j}(x_1^*), \dots, \mu_{\bar{A}_i^j}(x_i^*), \dots, \mu_{\bar{A}_n^j}(x_n^*), \mu_{A_{n+1}^k}(x_{n+1}^*)\right) \quad (4)$$

where x_i^* , $i = \overline{1, n}$ —number of tasks of the i -th performer, x_{n+1}^* —the difficulty of the task, moreover $\mu_{\bar{A}_i^j}(x_i^*) \in [0; 1]$, $\mu_{A_{n+1}^k}(x_{n+1}^*) = \begin{cases} 1, & \text{if } x_{n+1}^* = A_{n+1}^k, \\ 0, & \text{if } x_{n+1}^* \neq A_{n+1}^k \end{cases}$,

- (2) complex assessment of the reliability of the rule solution $C \in [0;1]$ (*complex*):

$$C = V * CF, \quad (5)$$

where CF —the usefulness of choosing a performer in a rule.

Consider the stages of the developed algorithm for assigning tasks to performers, taking into account the introduced indicators [6]:

- (1) determination of the level of complexity x_{n+1}^* of the requirements of the task entering the EDMS;
- (2) determination of the number of tasks x_i^* , that are simultaneously performed by the i -th performer;
- (3) calculation of the degrees of operation $\mu_{\bar{A}_i^j}(x_i^*)$ and $\mu_{A_{n+1}^k}(x_{n+1}^*)$ conditions for each r -th rule of the system $Rule_r$, $r = \overline{1, N}$ of the S_R system;
- (4) for each rule, the calculation of values V_r by the formula (4);
- (5) formation of a set of rules with a non-zero degree of confidence: $S_{conf} = \{Rule_r | V_r \neq 0\}$, $r = \overline{1, N}$;
- (6) calculation of the estimate C_r by formula (5) for all rules from the set $Rule_r \in S_{conf}$;
- (7) selection of the rule with the maximum complex assessment $Rule_r^* : \max_{r: Rule_r \in S_{conf}} C_r$;
- (8) getting the value a_i^* of a rule $Rule_r^*$ as a solution to a problem.

Thus, the fuzzy-production model of task distribution is a system of fuzzy-production rules of the form (3), which are determined by a combination of input conditions with task executors, as well as an inference algorithm based on rules. The membership functions (MF) and the reliability of the rules are used as model parameters. For the practical use of the model, it is necessary to identify the values of these parameters [16–18].

2.4 Development of a Method for Constructing Membership Functions in Model Rules

For the formation of MF in the rules of the model, a method for approximating the subjective assessments of performers (SAP) has been developed [6]. Let there be n potential executors of tasks $\{a_1, a_2, \dots, a_n\}$. Each of them sets the value of its workload level based on the number of tasks that it performs simultaneously. The method is based on the processing of the subjective assessment of the level of the performer’s workload using the scale, which is presented in Table 1.

The SAP method includes the following main stages:

- (1) assignment by the expert of the carrier S of a fuzzy set \tilde{A} , which corresponds to the MF for the level of the workload of performers;
- (2) survey of performers and the formation of their subjective assessments of the correspondence of the left $L_i(\alpha^*)$ and right $R_i(\alpha^*)$ boundaries of the selected workload level to a specific value of α^* from the set $\{1, 0.8, 0.6, 0.4, 0.2\}$ in accordance with Table 1 (moreover $[L_i(\alpha^*); R_i(\alpha^*)] = A_{\alpha^*} \subset S$, where A_{α^*} is an α^* - slice of a fuzzy set \tilde{A});
- (3) calculation of the average values of the left $L_{cp}(\alpha^*)$ and right $R_{cp}(\alpha^*)$ boundaries of the α^* - slice $A_{\alpha^*}^{cp} = [L_{cp}(\alpha^*); R_{cp}(\alpha^*)]$ for all α^* from the set of values of the performer’s confidence $\{1, 0.8, 0.6, 0.4, 0.2\}$ according to the following formulas:

$$L_{cp}(\alpha^*) = \sum_{i=1}^n \frac{L_i(\alpha^*)}{n}; \quad R_{cp}(\alpha^*) = \sum_{i=1}^n \frac{R_i(\alpha^*)}{n} \tag{6}$$

Table 1 Performer confidence rating scale

The numerical value of confidence, α	1	0.8	0.6	0.4	0.2
Interpretation	Absolutely sure	Substantially sure	Very sure	More or less sure	Poorly sure

- (4) construction of the MF of a fuzzy set \tilde{A} by combining the obtained α^* -slices $\tilde{A} = \bigcup_{\alpha^*} \alpha^* A_{\alpha^*}^{cp}$ and approximating their vertices by the method of least squares [19].

The proposed method allows you to build membership functions that determine the workload of performers in a fuzzy production model of assignment distribution.

2.5 Development of a Method for Determining the Reliability Values of Fuzzy Production Rules

To determine the values of the reliability of fuzzy rules, the *CF*-expert method was developed. Let *CF*—the reliability of the fuzzy rule *Rule*, which expresses the degree of the expert's confidence in the correctness and optimality of his decision on the distribution of the received task to a specific performer (the usefulness of the performer's choice). This parameter depends on the workload, performance, and qualifications of potential performers.

Within the framework of the method under consideration, the following concepts of the utility of choosing an executor are used, which take into account various factors for calculating the values of the *CF* parameter:

- (1) $\mu_{\tilde{c}_1}(a_i) \in [0;1]$ —the usefulness of choosing the *i*-th performer according to his current workload;
- (2) $\mu_{\tilde{c}_2}(a_i) \in [0;1]$ —the usefulness of choosing the *i*-th performer according to his performance;
- (3) $\mu_{\tilde{c}_{3k}}(a_i) \in [0;1]$ —the usefulness of choosing the *i*-th performer according to his qualifications for performing tasks of the *k*-th level of complexity.

The method for determining the reliability of fuzzy rules includes the following main stages:

- (1) calculating the utility of choosing the *i*-th performer based on his current workload $\mu_{\tilde{c}_1}(a_i)$ based on the following formula:

$$\mu_{\tilde{c}_1}(a_i) = \begin{cases} 1 - \frac{n_i}{N}, & \text{if } N \neq 0; \\ 1, & \text{if } N = 0 \end{cases} \quad (7)$$

where n_i —the number of simultaneously performed tasks by the *i*-th performer, $N = \sum_{i=1}^n n_i$ —the total number of tasks for all performers;

- (2) determination of utility $\mu_{\tilde{c}_2}(a_i)$ and $\mu_{\tilde{c}_{3k}}(a_i)$ based on the method of paired comparisons [20];
- (3) calculation of the reliability of fuzzy production rules by the formula:

$$CF_k^i = \mu_{\tilde{c}_1}(a_i) * \mu_{\tilde{c}_2}(a_i) * \mu_{\tilde{c}_{3k}}(a_i) \quad (8)$$

Thus, the *CF*-expert method is based on numerical utility estimates calculated by the formula (7), and an assessment of the utility of choosing a particular contractor based on his performance and qualifications.

The use of the *SAP* and *CF*-expert methods allows determining the values of the parameters of the membership functions $\mu_{\bar{A}_i}(x_i)$ and the reliability *CF* of each fuzzy production rule. As a result of the application of these methods, parametric identification of the fuzzy production model of the distribution of tasks occurs [6]. Thus, we can conclude that the construction of a set of rule systems for a different number and composition of task executors, identification of model parameter values, as well as the use of an inference algorithm on a rule system allows us to form a fuzzy production model of task distribution in the EDMS.

2.6 Decision Support Software Package

Based on the developed mathematical support, a software package has been implemented that allows decision-making support [21] on the choice of an executor for a specific task. The structure of the software package is shown in Fig. 1.

The module for forming the composition of performers is designed to add or exclude, if necessary (vacation, business trip, sick leave, etc.) performers of tasks and indicate their characteristics. The module for distributing tasks among performers includes a block for constructing a fuzzy-production model of assigning tasks, as well as a block for fuzzy inference. The experimental research module is designed to generate tasks, assess the accuracy of the task distribution model, and visualize the results obtained.

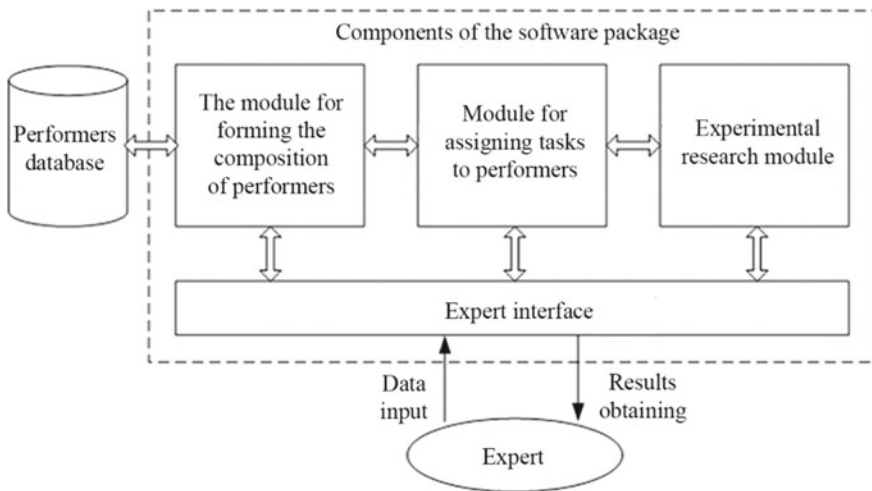


Fig. 1 Block diagram of the software package

Table 2 Fragment of the expert task distribution scheme

№ п/п	Difficulty level of the task	The current number of tasks for the performers				Executor of the task
		a_1	a_2	a_3	a_4	
1	High	14	16	14	6	a_1
2	High	15	16	14	6	a_3
3	Low	15	16	15	6	a_4
4	Middle	15	16	15	7	a_4

3 Formation of Reference Schemes for the Distribution of Tasks

To check the adequacy of the developed fuzzy-production model for the distribution of tasks, the data accumulated in the EDMS of the Roskommnadzor Office for the Republic of Tatarstan were used. A comparison of the results of the model with the reference (expert) schemes of task distribution was done. Table 2 shows a fragment of one of the schemes used.

In total, 10 reference schemes for the distribution of tasks were formed. The average number of tasks included in each scheme is 166. Table 3 shows the characteristics of the generated reference circuits.

The table for each scheme indicates the total number of tasks, their distribution by difficulty levels, the total number, and composition of task performers, as well as the structure of the expert distribution of tasks by performers.

4 Models of Assignment Distribution by Performers

The module for assigning tasks to performers has two modes of operation: direct assignments distribution (see Fig. 2) and assignments distribution, taking into account their predicted value (see Fig. 3).

It can be seen from the figure that the tasks to be distributed go to the tasks distribution module. This takes into account the composition of their potential performers. At the output, tasks are formed, distributed by performers.

In this case, the tasks to be distributed are a set of actually received and predicted values of the number of tasks from the PDOs. All tasks are submitted to the input of the task distribution module and are distributed among performers, taking into account the complexity of each task, as well as the qualifications, workload, and performance of the performers. After the distribution of the entire set of the current and predicted number of tasks, those tasks that have not been received yet, but were only predicted are excluded from the resulting distribution. Thus, the final set of tasks is formed, rationally distributed among the performers.

Table 3 Characteristics of reference job distribution schemes
Table 3 Characteristics of reference job distribution schemes

№ of schemes	Total tasks	Number of tasks by difficulty levels			Total performers	The structure of the expert distribution of tasks						
		Low (L)	Middle (M)	High (H)		a_1 (L, M, H)	a_2 (L, M, H)	a_3 (L, M, H)	a_4 (L, M, H)	a_5 (L, M, H)	a_6 (L, M, H)	
1	148	2	36	110	4	0,0,40	0,0,40	0,0,36	0,0,36	0,6,34	0,6,34	2,30,0
2	156	21	109	26	4	0,37,11	0,36,12	13,15,3	13,15,3	8,21,0	8,21,0	8,21,0
3	180	15	82	83	4	0,18,38	0,17,35	0,17,35	0,17,35	9,25,0	9,25,0	9,25,0
4	174	18	78	78	4	0,21,27	0,21,27	0,19,29	0,19,29	6,18,22	6,18,22	6,18,22
5	190	14	116	60	5	0,16,21	0,20,19	0,23,15	0,23,15	4,28,5	4,28,5	10,29,0
6	163	12	89	62	5	0,18,22	0,20,19	0,20,19	0,20,19	1,23,13	1,23,13	1,23,13
7	127	14	65	48	5	0,5,18	0,8,14	0,11,12	0,11,12	1,18,4	1,18,4	1,18,4
8	135	17	65	53	6	0,4,22	0,7,16	0,11,9	0,11,9	0,14,6	0,14,6	13,14,0
9	207	20	119	68	6	0,10,23	0,19,17	0,22,16	0,22,16	8,23,1	8,23,1	12,21,0
10	179	21	111	47	6	0,8,18	0,12,16	0,24,9	0,24,9	0,26,4	0,26,4	14,16,0

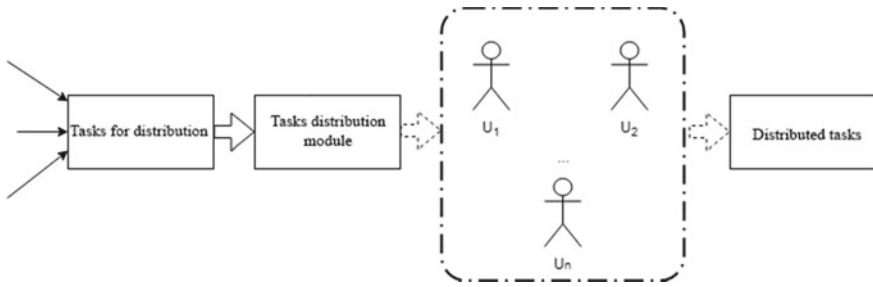


Fig. 2 Scheme of the model of the direct distribution of tasks without taking into account their predicted values

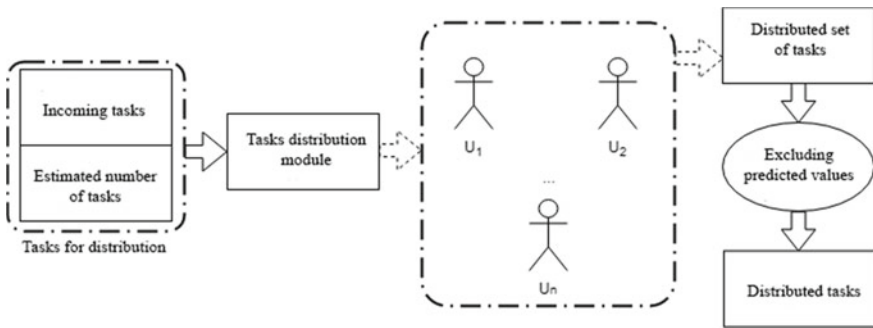


Fig. 3 Scheme of the task distribution model taking into account their predicted values

5 Solving the Problem of Predicting the Number of Tasks Received Through the EDMS

To solve the problem of predicting the number of tasks received through the EDMS, we used real data in the form of time series [22–24], describing the actual number of received jobs of varying complexity with a step of one month, starting from April 2015 to the present. The tasks are classified according to three levels of difficulty, respectively, three-time series were used for the analysis. Each such series was a sample containing the following data: date (month, year) and the number of tasks.

A multilayer feedforward neural network was used as a model for predicting the number of tasks [25–29]. For its construction, the analytical platform Deductor was used [30], on the basis of which the following stages of modeling were performed:

- (1) loading and preparing initial data for analysis;
- (2) building neural network models of various structures with a change in the number of hidden layers and neurons in each layer;
- (3) testing the constructed models and choosing the best one in terms of the accuracy of the predicted values obtained;

- (4) application of the selected neural network model to predict the number of tasks arriving for distribution with different levels of complexity.

In accordance with the forecasts obtained, in the further distribution of tasks between the performers, the following number of tasks was used: 1 task of high difficulty, 15—of medium difficulty, and 22—of low difficulty.

6 Solution of the Task Distribution Problem Taking into Account Their Predicted Values

Let us consider the distribution of tasks between the performers, taking into account the predicted values obtained as a result of applying the neural network predictive model. Tasks were distributed among four specialists of the personal data protection department of the territorial office of Roskomnadzor:

- (1) chief specialist (performance level—1, qualification for performing tasks of high complexity—1, medium complexity—1, low complexity—1);
- (2) leading specialist (performance level—0.8, qualification for performing tasks of high complexity—1, medium complexity—1, low complexity—1);
- (3) a specialist-expert (performance level—0.6, qualification for performing tasks of high complexity—0.5, medium complexity—0.9, low complexity—1);
- (4) specialist of the 1st category (performance level—0.7, qualification for performing tasks of high complexity—0.2, medium complexity—0.8, low complexity—0.9).

During the setting parameters for performers, the initial number of tasks was set to zero. In the next step, the current number of tasks was generated:

- (1) high level of complexity—2;
- (2) medium level of complexity—14;
- (3) low level of complexity—21.

After the successful generation of tasks, they were distributed among four performers by their levels of performance and qualifications for performing tasks of a specific level of complexity.

Consider the distribution of tasks taking into account their predicted values. As described earlier, to distribute tasks taking into account the predictive model, it is necessary to form a set consisting of the current number of tasks and their predicted number. In this case, the current number of tasks is represented by 2 tasks of high complexity, 14 tasks of medium complexity, and 21 tasks of low complexity. Per the constructed predictive models, the predicted number of tasks of high complexity is 1, tasks of medium complexity—15, tasks of low complexity—22. Accordingly, the following number of tasks will participate in the current distribution:

- (1) 3 tasks of high complexity;

Table 4 Workload of performers with different models of assignment distribution

№	Performers	The workload in the distribution of tasks, without taking into account the predicted values, %	The workload in the distribution of tasks, taking into account the predicted values before their exclusion, %	The workload in the distribution of tasks, taking into account the predicted values after their exclusion, %
1	Chief specialist	65	95	40
2	Leading specialist	50	95	45
3	Specialist-expert	35	90	55
4	Specialist of the 1st category	35	95	45

- (2) 29 tasks of medium complexity;
- (3) 43 tasks of low complexity.

As a result of the distribution for execution, the chief specialist received 19 tasks, the leading specialist received 19 tasks, the specialist-expert received 18 tasks, and the specialist of the 1st category received 19 tasks.

In accordance with the described model, after the distribution of all tasks among the performers, it is necessary to exclude the forecast tasks from the resulting distribution. In this case, this is 1 task of high complexity, 15 tasks of medium complexity, and 22 tasks of low complexity.

Distributing tasks without taking into account and taking into account their predicted values, the results of the workload of performers were obtained, shown in Table 4.

As can be seen from the presented table, the use of the proposed technology for forecasting and distribution of tasks allows reducing the burden of maintaining the register of PDOs on the chief and leading specialists. This is a positive effect since the main job responsibilities of these specialists include the implementation of inspections of the activities of the PDOs for compliance with the legislation of the Russian Federation in the field of personal data protection and responding to complaints from citizens in this area.

The main duties of a specialist-expert and a specialist of the 1st category include maintaining a register of PDOs, which explains the increased load on them to perform the type of tasks under consideration.

7 Conclusion

The proposed model for the distribution of tasks, taking into account their predicted values in comparison with the original model, allows achieving the following results:

- (1) reducing the workload of the chief and leading specialists;

- (2) more even distribution of workload between performers;
- (3) the possibility of the optimal distribution of tasks between performers, taking into account such features of time series as trend and seasonality;
- (4) the possibility of the optimal distribution of tasks between performers, taking into account the vacation schedule and other situations involving the replacement of any position by another specialist.

Thus, the results of the research have shown the effectiveness of the proposed approach and the possibility of its practical use for forecasting and assigning tasks to performers in EDMS.

References

1. Paramonova, I.E.: Electronic document-management systems: a classification and new opportunities for a scientific technical library. *Sci. Tech. Info. Process.* **43**(3), 136–143 (2016)
2. Obukhov, A.D., Krasnyanskiy, M.N., Dedov, D.L.: Automation of structural and parametric synthesis of electronic document management systems based on neural network architecture. *Int. Rev. Aut. Control* **12**(3), 115–122 (2019)
3. Rosa, A.T.R., Pustokhina, I.V., Lydia, E.L., Shankar, K., Huda, M.: Concept of electronic document management system (EDMS) as an efficient tool for storing document. *J. Critical Rev.* **6**(5), 85–90 (2019)
4. Zhang, C.-J., Zhang, W.-H.: Study on quick response distribution task management in agricultural products logistics based on e-commerce. In: *Proceedings of the International Asia Conference on Industrial Engineering and Management Innovation: Core Areas of Industrial Engineering*, pp. 1535–1542 (2012)
5. Khalid, S., Ullah, S., Alam, A., Ur Rahman, S.: A new task distribution model to increase user performance in collaborative virtual environment. *Jurnal Teknologi.* **78**(4–3), 23–29 (2016)
6. Ismagilov, I.I., Nugae, F.S., Katasev, A.S., Talipov, N.G., Kataseva, D.V.: Decision-making support system for tasks distribution in personal data operators register maintaining based on a fuzzy-production model. *Dilemas contemporaneos-educacion politica y valores* **6** (SI), 64 (2019)
7. Jain, D.K., Kumar, A., Sharma, V.: Tweet recommender model using adaptive neuro-fuzzy inference system. *Fut. Gener. Comput. Syst.* **112**, 996–1009 (2020)
8. Katasev, A.S.: Neuro-fuzzy model of fuzzy rules formation for objects state evaluation in conditions of uncertainty. *Comput. Res. Model.* **11**(3), 477–492 (2019)
9. Radhi, A.M.: Risk assessment optimization for decision support using intelligent model based on fuzzy inference renewable rules. *Indonesian J. Electri. Eng. Comput. Sci.* **19**(2), 1028–1035 (2020)
10. Chupin, M.M., Katasev, A.S., Akhmetvaleev, A.M., Kataseva, D.V.: Neuro-fuzzy model in supply chain management for objects state assessing. *Int. J. Supply Chain Manage.* **8**(5), 201–208 (2019)
11. Bounabi, M., El Moutaouakil, K., Satori, K.: Association models to select the best rules for fuzzy inference system. *Adv. Intell. Syst. Comput.* **1076**, 349–357 (2020)
12. Katasev, A.S., Kataseva, D.V., Emaletdinova, Yu, L.: Neuro-fuzzy model of complex objects approximation with discrete output. In: *Proceedings of the 2nd International Conference on Industrial Engineering, Applications and Manufacturing*, 7911653 (2016)
13. Yusoff, B., Merigo, J.M., Hornero, D.C.: Analysis on extensions of multi-expert decision making model with respect to OWA-based aggregation processes. *Adv. Intell. Syst. Comput.* **730**, 179–196 (2018)

14. Kizim, A.V., Kravets, A.G.: On systemological approach to intelligent decision-making support in industrial cyber-physical systems. *Stud. Syst. Dec. Control* **260**, 167–183 (2020)
15. Bolshakov, A.A., Kulik, A., Sergushov, I., Scripal, E.: Decision support algorithm for parrying the threat of an accident. *Stud. Syst. Dec. Control* **260**, 237–247 (2020)
16. Anufriev, D., Petrova, I., Kravets, A., Vasiliev, S.: Big data-driven control technology for the heterarchic system (building cluster case-study). *Stud. Syst. Dec. Control* **181**, 205–222 (2019)
17. Liu, Y., Lu, J., Long, X., Zhou, R., Wu, Y.: Dynamic identification of model parameters for energy storage batteries. *J. Natl. Uni. Def. Technol.* **41**(5), 87–92 (2019)
18. Cao, Y., Kou, X., Wu, Y., Jermisittiparsert, K., Yildizbasi, A.: PEM fuel cells model parameter identification based on a new improved fluid search optimization algorithm. *Energy Rep.* **6**, 813–823 (2020)
19. Sembiring, N., Dewi, I.S.: Sustainable supply chain model by using digital partial least square method. *IOP Conf. Ser. Mater. Sci. Eng.* **830**(4), 042006 (2020)
20. Siperkovskis, V.: Using T. Saaty method in transport systems planning. *Trans. Telecommun.* **10**(4), 18–27 (2009)
21. Veshneva, I., Bolshakov, A., Kulik, A.: Increasing the safety of flights with the use of mathematical model based on status functions. *Stud. Syst. Dec. Control* **199**, 608–621 (2019)
22. Ismagilov, I.I., Khasanova, S.F.: Algorithms of parametric estimation of polynomial trend models of time series on discrete transforms. *Acad. Strat. Manage. J.* **15**(SpecialIssue), 21–28 (2016)
23. Last, M., Bunke, H., Kandel, A.: *Data mining in time series and streaming databases*, p. 171. World Scientific Publishing Co. Pte. Ltd (2018)
24. Perfilieva, I.G., Yarushkina, N.G., Afanasieva, T.V., Romanov, A.A.: Web-based system for enterprise performance analysis on the basis of time series data mining. *Adv. Intell. Syst. Comput.* **450**, 75–86 (2016)
25. Katasev, A.S., Emaletdinova, L.Y., Kataseva, D.V.: Neural network model for information security incident forecasting. In: *Proceedings of the International Conference on Industrial Engineering, Applications and Manufacturing*, 8728734 (2018)
26. Dykin, V.S., Musatov, V.Yu., Varezhnikov, A.S., Bolshakov, A.A., Sysoev, V.V.: Application of genetic algorithm to configure artificial neural network for processing a vector multi-sensor array signal. In: *Proceedings of the International Siberian Conference on Control and Communications*, 7147049 (2015)
27. Yarushkina, N., Filippov, A., Moshkin, V., Namestnikov, A., Guskov, G.: The social portrait building of a social network user based on semi-structured data analysis. *CEUR Workshop Proc.* **2475**, 119–129 (2019)
28. Anikin, I.V., Makhmutova, A.Z., Gadelshin, O.E.: Symmetric encryption with key distribution based on neural networks. In: *Proceedings of the 2nd International Conference on Industrial Engineering, Applications and Manufacturing*, 7911640 (2016)
29. Alekseev, A., Katasev, A., Kirillov, A., Khassianov, A., Zuev, D.: Prototype of classifier for the decision support system of legal documents. *CEUR Workshop Proc.* **2543**, 328–335 (2020)
30. Lomakin, N., Shokhnekh, A., Sazonov, S., Lukyanov, G., Gorbunova, A.: Hadoop and Deductor based digital ai system for predicting cost of innovative products in conditions of digitalization of economy. *ACM Int. Conf. Proc. Ser.* 3373810 (2019)

Society 5.0: Smart Cities and Regions

Smart City: Cyber-Physical Systems Modeling Features



**Nikolai A. Fomin, Roman V. Meshcheryakov, Andrey Y. Iskhakov,
and Yuri Y. Gromov**

Abstract This chapter presents an approach to modeling smart city subsystems based on the principle of their cyber-physical integration into a single infrastructure. The uniqueness of the approach is the provision of the smart city topology as a set of cyber-physical clusters that ensure the functioning of the city. Unlike existing solutions, the proposed approach allows to consider the relationship of the information flows between different levels of smart city systems. Considered features of modeling of the cyber-physical water supply system (including active control system models, resource allocation models, scenario modeling) were put into practice by the smart city infrastructure in Saint Petersburg, Russia. The analysis of the functioning of the megalopolis water supply enterprise revealed significant vulnerabilities in ensuring the strategic security of the Leningrad Region, Russia. A basic list of recommendations for reducing the probability of water shortage has been developed based on scenario modeling of disabling of the main water supply source. These results are the basis for creating simulation models of interrupting the main source of smart city water supply and considering issues of improving the efficiency of water supply cyber-physical systems control in the smart city.

Keywords Cyber-physical systems · Management systems modeling · Smart city · Water supply systems · Improvement of control models · Digital water utility

N. A. Fomin (✉) · R. V. Meshcheryakov · A. Y. Iskhakov
V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow
117997, Russia
e-mail: science-fomin@yandex.ru

Y. Y. Gromov
Tambov State Technical University, Tambov 392000, Russia

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2021
A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced
Human-Centered Society*, Studies in Systems, Decision and Control 333,
https://doi.org/10.1007/978-3-030-63563-3_7

1 Introduction

The problem of resource management in modern cities is particularly relevant. Urbanization imposes certain requirements on the development of models of resource management. In the modern world, the issue of improving the efficiency of water resources management in modern cities is an important task. Cities are evolving, becoming more technological and smart while facing serious challenges—in addition to water scarcity due to climate change, the scarcity also occurs due to the degradation of water sources, anthropogenic impact on sources, an increase in the number of urban residents, which leads to significant problems of reducing the quality of life of urban residents, and deterioration of health. In December 2016, the United Nations General Assembly unanimously adopted the resolution “International Decade (2018–2028) for Action—Water for Sustainable Development” to help put a greater focus on the water during ten years [1], confirming the international importance of the issue of security water.

Modeling and management of environmental systems is an important aspect of ensuring the strategic development of cities, of which the quality of management of urban water supply systems plays a special role. Urbanization of the modern world, the dynamic development of society, due to the need to provide residents of modern cities with clean fresh water. At the same time, public water supply systems face a number of problems, including outdated infrastructure, growing regulatory requirements, problems of water quantity and quality, and lack of resources. According to [2], the United States will have to spend up to \$ 200 billion on water systems over the next 20 years to upgrade transmission and distribution systems. Of this amount, \$ 97 billion (29%) is estimated to be needed to control water losses. The average loss of water in systems is 16%—up to 75% of this can be recovered.

A similar situation is typical in most countries, which gives this topic of international significance. Anthropogenic impact on the environment, degradation of water resources, climate change increases the risk of water scarcity of cities, the occurrence of threats to water security of agglomerations. The potential of using modern technologies, including complex adaptive systems for controlling water supply in a modern city, allows minimizing risks and improving the quality of water provided to residents of cities.

The technological development of modern cities has certainly affected cities that are evolving into smart cities. The modern city is an abundance of systems that carry potential threats. Monitoring and security of critical water supply infrastructure systems in modern cities is an integral part of the security of vital support systems. At the same time, the possibility of integrated monitoring of systems increases the efficiency of using both the systems themselves in terms of maintenance and increasing the operating life, including involving robotic systems and systems and improves the quality of consumed water resources [3].

An important direction in ensuring the strategic security of cities is modeling potential threats and adjusting management based on the identified shortcomings. An integrated approach will increase the level of strategic security of modern cities.

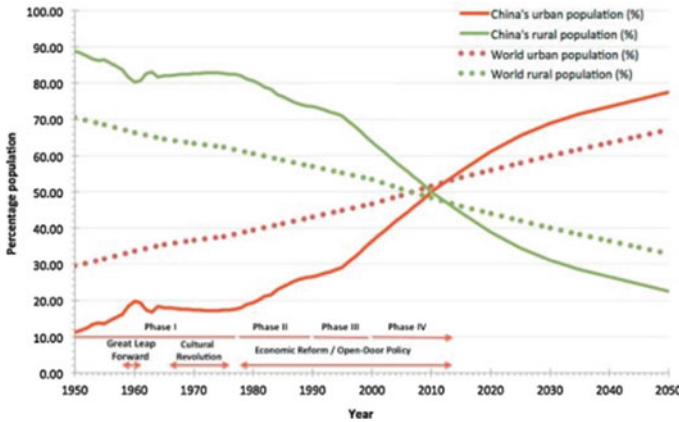


Fig. 1 Temporal dynamics of the urbanization levels (% urban population) for China and the world

Urbanization trends place higher demands on water quality. Often, previously designed water supply systems cannot provide a quantitative indicator of the increase in the volume of water consumed. The dependence on water is increasing at times, provided that additional external anthropogenic impacts on the environment are affected by the degradation of water resources and, as a result, the risk of lowering the quality of water provided to the urban population is realized. Population growth combined with economic growth has fueled recent urban land expansion in North America. Between 1970 and 2000, the urban land area expanded at a rate of 3.31% creating unique challenges for conserving biodiversity and maintaining regional and local ecosystem services [4]. This trend also confirms [5], demonstrating the dynamics of the urbanization levels for China and the world (Fig. 1).

An audit of water supply systems (WSSs) in the United States, Netherlands and United States [6] shows that in the Netherlands, at least half of the water distribution pipes have been replaced since the 1970s; as a result, pipe networks are, on average, 33–37 years old. Although there are regional differences, an estimated 22% of the pipes in the United States are more than 50 years old; the average age of pipe at failure is 47 years, and only 43% of pipes are considered to be in good or excellent condition. In the United Kingdom, as much as 60% of pipe inventory does not have a record of pipe age, and estimates of average pipe age are on the order of 75–80 years overall.

Also, the aquatic environment acts as a major pool for antibiotics and antibiotic-resistant genes (ARGs). Antibiotics in the aquatic environment generally originate from effluents of wastewater treatment plants, industrial sites, hospitals, and livestock farms. It was estimated that in 2013 the total usage of 36 antibiotics reached 92,700 tons in China, making it the world’s largest producer and user of antibiotics [7]. This allows concluding that, in addition to the level of deterioration of water supply systems, dividing by climatic zones, in the future, the model should include parameters for a technical assessment of the state of WSS in cities, as well as the

quality of water supply sources as an integral part of assessing the level of water security in cities.

Based on the problems, we formulate the goal of the study—to identify the features of modeling the CP WSS of the Smart City.

Research Objectives:

1. To analyze the systems of a smart city as a combination of CPS and classifying systems.
2. Identify the features of CPS of a smart city by the example of water supply systems.
3. Build practical basic models for managing the cyber-physical water supply system in Saint Petersburg, Russia.

2 CPSs Within the Smart City

2.1 *Concept Overview*

To solve the first task, it is necessary to form a comprehensive view of the specifics, determine the concepts used in the study. The basic concepts are CPSs and smart cities. According to [8], cyber-physical systems (CPS) are complex systems with organic integration and deep interaction of computing, communication, and control (3C) technologies. Currently [9], CPS are managed, reliable, and extensible networked physical systems that are further integrated with computing, communication, and control capabilities that can interact with people through many new modalities. CPS are the foundation and core of Industry 4.0 and the Industrial Internet. As we see, the obligatory criterion of CPSs is the presence of functions and human participation in the performance of a certain function. The classification of CPSs in a smart city will be considered in more detail below.

According to [10] assert that a smart city has six main dimensions: smart economy, smart people, smart governance, smart mobility, smart environment, and smart living. The authors propose replenishing the concept, namely: smart city—is a set of systems operating in the digital ecosystem of the city, with the aim of increasing the efficiency of the use of city resources, ensuring safety, improving the quality of life and health of the city residents.

In the Russian Federation, for example, the term Safe City is used along with the term Smart City. The functionality of Safe City systems in the Russian Federation is a set of systems aimed at coordinating the interaction of services responsible for maintaining law and order, preserving the life and health of city residents, increasing the efficiency of management of administrative and administrative authorities, and improving interagency cooperation.

2.2 *Classification of CPS of the Smart City*

The next stage of the study is to consider the Smart City as a set of CPS and, based on this set, divide CPS into directions, creating a basic classification. At the same time, the classification requires a more advanced division and not just the creation of the main and supporting clusters.

The city is a holistic management object that operates according to certain rules and depends, in terms of management theory, on management decisions, managerial impact, quality, and timeliness. At the same time, the city can be represented as a holistic distributed active management mechanism in which there is a plentiful amount of active elements, for example, providing services, residents, enterprises. In turn, they have a mechanism for influencing the functioning of the governing body, and managerial influences depend on the preferences of the active elements.

Existing approaches to decomposing smart cities into cyber-physical systems do not allow a comprehensive assessment of the interaction between different cyber-physical systems. The approach to classification is either very superficial [9]. Only the basic functions of the city and the division into main and supporting clusters are taken into account. Or there is no classification at all [19], which does not allow us to form a consolidated topology of cyber-physical systems of a smart city.

In this chapter, we propose to consider a smart city as a set of basic cyber-physical systems. These systems include a set of modern tools for protecting and processing the received information using basic Analytics technologies for making the best management decisions. The procedure for decomposing a smart city into cyber-physical systems is fundamental in the approach proposed by the authors. The extended classification requires a comprehensive infrastructure audit to identify cyber-physical systems. Each of these complexes includes a wide range of modern tools and interfaces that ensure the processing of various types of data and the provision of digital services and services, methods of automated processing, recognition, and ensuring all the properties of information security (confidentiality, integrity, availability of data).

In contrast to analogs, this approach allows us to take into account the peculiarity of the disparity and diversity of cyber-physical systems. Classification involves the decomposition of a smart city into cyber-physical systems and the division into categories (clusters). The classification is based on active system management functions. This makes it possible to comprehensively form the topology of cyber-physical smart city systems and evaluate their interaction. The basic classification of smart city cyber-physical systems developed by the authors is presented in Table 1.

The “Security” cluster can be interpreted both as “Providing” and “Supporting”. In turn, the authors decided to separate this function into a separate cluster, due to the potential increase in the number of threats to the functioning of the Smart City, especially cyber threats.

Table 1 CPS classification of smart city

Cluster	Description	Example CPS
Control	It is aimed at ensuring timely decision-making on the basis of reliable data obtained from CPS	Support and decision-making system, secure management, analytical systems
Control	Supervision of compliance with established requirements. Often presented as a subsystem of the cluster “management” or “safety”	Monitoring systems and condition, environment, traffic congestion, equipment condition
Safety	It is aimed at identifying potential threats to the security of both city residents and their property, as well as city infrastructure, enterprises	CCTV, video surveillance, warning systems, intrusion detection
Providing	Ensuring the functioning of the city	Water supply, gas supply, electricity supply, telecommunication, logistics
Supporting	Supporting systems responsible for maintaining the functioning of the city	Smart manufacturing, smart buildings

3 Modeling the Cyber-Physical Water Supply System of the Smart City

3.1 Analysis of the Specificity of the Subject

Based on the classification of Smart city CPS described in Chap. 2, let’s look in more detail at the water supply system as an example of a CPS. To evaluate the features of modeling, it is necessary to analyze the principle of operation of the system.

The system is based on the principle of resource allocation using the infrastructure of the water supply enterprise. The supply function is the basic one for the enterprise and the supplied water (resource) must meet the established parameters (characteristics) approved in this particular locality (city, country). The characteristics are the quality and guarantees of quantitative indicators of providing residents in case of compliance with the established requirements, often contractual relations and pricing policy [11].

The system consists of a set of water supply sources (resources), a management body (water supply enterprise), and water supply infrastructure (main and local communications).

Let’s imagine the functioning of a water supply company as a basic business process (Fig. 2).

A distinctive feature of digital Vodokanal as a CPS of a Smart city from an “analog” Vodokanal is the availability of specialized equipment, sensors, for example, based on IoT and IIoT, data centralization and Analytics that allow you to quickly monitor the state of systems, optimize maintenance costs, and indirectly increase the service life of the infrastructure.

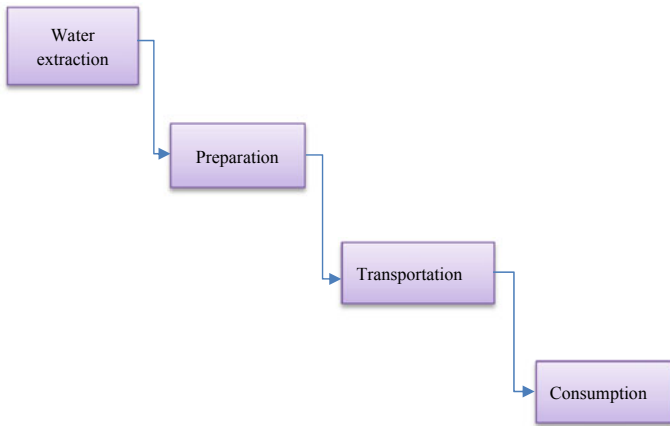


Fig. 2 The Basic business process of functioning of a water supply company

If we consider the water supply system as an active system, then with the appearance of additional intelligence of the system, increasing the active elements of the system, the complexity of managing the system increases. In this case, part of the control becomes automated or automatic, thereby reducing the human factor, but there are additional potential threats that can affect the system’s control mechanism.

3.2 Building Basic Models of a Cyber-Physical Water Supply System

Based on the study of the specifics of the cyber-physical water supply system of a Smart city, we will build basic models. A cyber-physical water supply system is a multi-link active resource distribution system, the resource is water. The system is managed by the management body. In control theory, active and passive systems are distinguished. In addition to being able to select a state, the AC elements have their interests and preferences, that is, they select the state purposefully (otherwise their behavior could be considered passive).

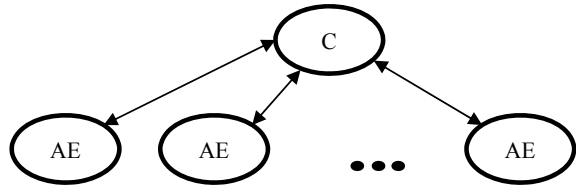
The General formulation of the task of managing an active system is as follows: The state of the system that belongs to some valid set A , is represented by the variable (1):

$$y \in A \tag{1}$$

The state of the system depends on control actions (2):

$$\eta \in U: y = G(\eta) \tag{2}$$

Fig. 3 Two-level automated fan-type system



Value $K(\eta) = \Phi(\eta, G(\eta))$ it is called control efficiency, where $\eta \in U$ —control the action.

The task of the governing body is to choose such an acceptable control that maximizes the value of efficiency, provided that the reaction of $G(\eta)$ of the system to the control actions is known (3):

$$K(\eta) \rightarrow \max_{\eta \in U} \tag{3}$$

Accordingly, the management task is reduced to the need to choose the optimal control (4):

$$\eta^* = \tilde{\eta}(y) \in U, \tilde{\eta}: A \rightarrow U \tag{4}$$

Let’s imagine a two-level active fan-type system (Fig. 3).

The AU structure is a set of information, management, and other relationships between AU participants, including relationships of subordination and distribution of decision-making rights. The governing body, active elements are members of the system, and resources are distributed according to the specified rules [12].

In turn, the section of contract theory is also applicable to modeling the functioning of water supply CPS. Contract theory-studies incentive mechanisms in active systems operating under conditions of external probabilistic uncertainty are also discussed in Sect. 3.4.

Consider the following approach based on the resource allocation mechanism. When distributing a resource between n active elements центр. The resource value of the i—element is determined by its utility function (5), where x_i —is the amount of resource it receives, a r_i —is the type of active element.

$$\varphi_i(x_i, r_i) \tag{5}$$

In turn, the main task of the center will be to distribute the resource itself with a given goal, for example, to maximize the utility of all elements (6)

$$\sum_{i \in I} \varphi_i(x_i, r_i) \rightarrow \max_{x \geq 0} \tag{6}$$

When modeling CPS, it is necessary to take into account the feature of the multi-links of the active system, which affects the system control model.

3.3 Consideration of the Features of Modeling Integrations of the Cyber-Physical Water Supply System of a Smart City

Modeling integrations of the cyber-physical water supply system of a Smart City has certain features. Integration refers to the process of combining disparate system elements into a single unit or dividing them into secure segments in order to improve the efficiency of system management. In our case, we will consider the process of integrating the integration of a cyber-physical water supply system with the management center of Smart city systems. Which is often a hardware and software platform that meets the security requirements for data storage and processing.

Integration in order to centralize data, improve the management and control mechanism, should be considered as additional potential threats to the functioning of the system.

To evaluate the performance of the CPS modeling integration of water Smart city model underlying the data exchange between the elements of the active system as a “Digital water supply system” (DWSS) and “Analog water supply system” (AWSS), it will also give an initial expert assessment of changes in the number of parts of the system, additional potential vulnerabilities.

Figure 4 simulates the functioning of an “analog” water utility, i.e. a water

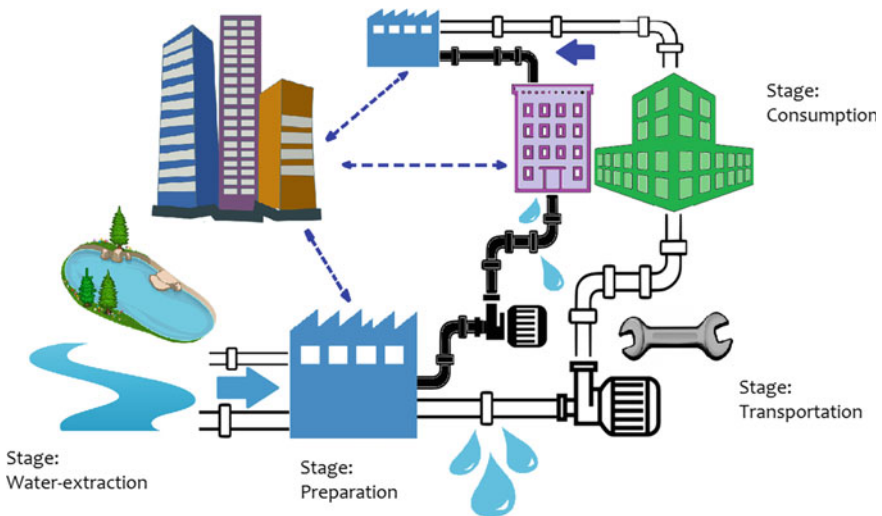


Fig. 4 Modeling of the operation of the “analog water supply system”

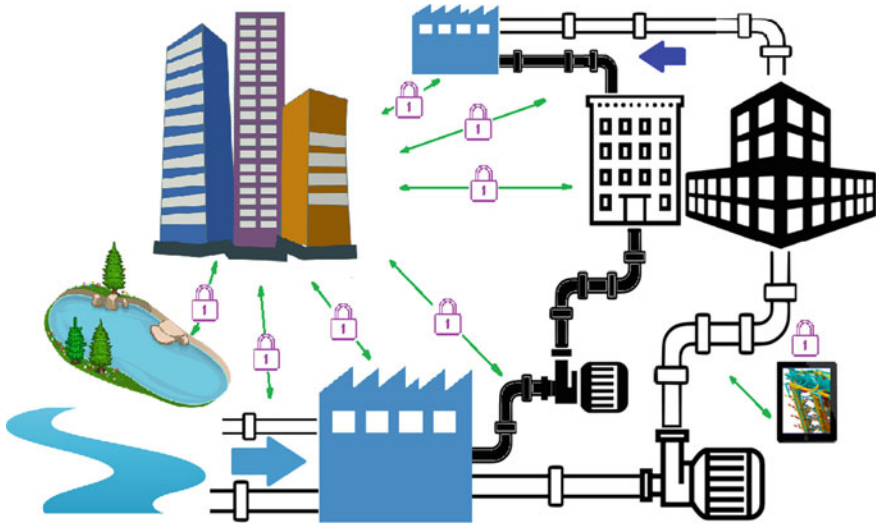


Fig. 5 Modeling of the operation of the “digital water supply system”

supply company that does not use centralized systems for accounting and monitoring resources, equipment condition, operational monitoring of resource losses, and predictive Analytics. Often, the information provided to the management body is unreliable, untimely, and incomplete, which leads to the inability to form an objective full-fledged situation for the water supply enterprise, the quality of the resource, equipment wear, planned and actual costs for repairs, and to ensure the level of safety.

Figure 5 simulates the functioning of the “Digital water supply system” as a complete CPS. Data is exchanged in a secure, centralized, and automated way at each production stage, according to the business process Fig. 2—from source to consumer. It analyzes the chemical and biological composition and resource losses, the condition of equipment and its level of wear, and quantitative indicators of consumption. Innovative technologies become a tool to improve the control efficiency of water supply systems of a modern city [13]. By automating the operation of the “digital water channel”, the potential number of security threats increases, which, as already mentioned, must be taken into account in the management models of the CPS of water supply in a smart city [14].

Based on the analysis, we compare the advantages and disadvantages of the two types of water utilities (Table 2).

The digital water channel control model is different from the analog water channel management model. Digitalization of the industry has a positive impact on the level of development of water utilities, centralizes the exchange of data over secure communication channels, provides the opportunity to build models of both deterministic and stochastic, the result of which is to optimize costs [13] and improve the quality of services provided to consumers. Often, the main problem of the transition to the

Table 2 Comparison of advantages and disadvantages of the operation of “analog water supply system” and “digital water supply system”

	The analog water supply system	The digital water supply system
Advantages	Less qualified personnel to maintain systems “Used” to manage this way and unwillingness to change the strategy	Improving the quality of delivered water Increasing the level of security Loss reduction Cost optimization
Disadvantages	Additional threat The lack of systems for monitoring the quality of the resource There is no reliable operational information about losses	Increasing the number of threats Additional infrastructure maintenance costs

management of a “digital water channel” is the lack of funding even to maintain the existence of an “analog water channel”. The authors believe that this problem is a serious test and challenge for the modern world in the period of transition to digital format, and it is important that the technological solutions of Smart cities also meet the security requirements, including in the field of CPS water supply for Smart cities.

3.4 Practical Implementation

Let’s take a practical example of the created basic models and analyze their applicability. One of the largest cities in the Russian Federation – Saint Petersburg with a population of more than 5 million people-was taken for the study. Vodokanal St. Petersburg, Russia’s oldest and one of the biggest suppliers and sewage operators in Russia. Vodokanal of Saint Petersburg is one of the best water supply and sanitation companies in Russia. The installed technological equipment allows us to provide high-quality water to consumers, while the negative factors of the company’s operation for environmental pollution are minimal and meet the requirements of international standards. Water in the Baltic Sea in the area of Saint Petersburg has a low level of anthropogenic pollution, which allows us to conclude about the effectiveness of measures and technologies used to ensure the management of the functioning of the cyber-physical water supply system.

Build two almost basic models, as applied to St. Petersburg:

1. Resource allocation model, according to the theory of contracts.
 2. Scenario modeling of failure of the main water supply source (degradation, infection).
- A. *Resource allocation Model, according to contract theory.*
 An active element, such as a consumer, selects an action $y \in A$, which under the influence of the external environment leads to the implementation of the result of the i -th active element (7), where $\Gamma \subseteq A$ —is the set of feasible actions of the

active element, A_0 —is the set of valid results of activity of the active element [12].

$$z \in A_0 \quad (7)$$

Suppose that on the available set of possible actions of the active element is finite and has the form

$$A = \{y_1 \dots, y_n\}, A_0 = \{z_1 \dots, z_n\} \quad (8)$$

Set formula

$$\sigma_i = \sigma(y_i), c_i = c(y_i), p_{ij} = p(z_j, y_i) \quad (9)$$

At the first stage, you need to define a set of actions to be implemented: for each possible action

$$y_k, k = \overline{1, n} \quad (10)$$

We are looking for an incentive system σ_j^k , that satisfies the constraints of (11) and implements it.

$$0 \leq \sigma_j^k \leq C, \quad j = \overline{1, n} \quad (11)$$

Looks like:

$$\begin{cases} \sum_{j=1}^n \sigma_j^k p_{kj} - c_k \geq \sum_{j=1}^n \sigma_j^k p_{ij} - c_i, A_i = \overline{1, n} \\ 0 \leq \sigma_j^k \leq C, j = \overline{1, n} \end{cases} \quad (12)$$

Conclusion: this basic approach allows us to conclude that contract theory can also be applied to modeling the functioning of water supply CPS.

After analyzing the basic theoretical mechanism of resource allocation, according to the theory of contracts, let's move on to a practical assessment of the resources themselves, namely the sources of water supply in St. Petersburg.

- B. *Scenario modeling of failure of the main water supply source (degradation, infection).*

During the study, it was revealed that according to the approved state scheme of water supply and sanitation of the city of Saint Petersburg [15], the main source of water supply in the city is the Neva river (98% of the total amount of water consumed; the remaining 2% are underground sources). The proportional ratio makes it possible to consider that it is actually the main and only one.

In case of contamination/degradation of the main source of water supply, it is necessary to promptly ensure the transition to a backup source of water supply, if there is one. It often happens that in many settlements there are no redundant sources for a variety of reasons—from natural to economic [1].

Reservation of water supply sources dates back to antiquity, for example, in the Byzantine period in Istanbul, one of the great creations in the field of water supply reservation was built, combining with a special architectural monument—the Big Basilica Cistern [16]. CIS tank was about 100,000 tons. The 4.80 m thick exterior walls were made waterproof by covering them with a 3.5 cm thick brick dust solution.

According to the Russian legislation [17], the reservation of sources of drinking and household water supply in the event of an emergency is carried out on the basis of underground water bodies protected from contamination and clogging. At the same time, such sources should be provided with special protection zones, the regime of which corresponds to the regime of zones of sanitary protection of underground sources of drinking and household water supply.

Source redundancy can be implemented with the help of additional wells, backup protected storage (open and closed), rapid transportation of the resource by transport, and storage of large-size reserves at enterprises (Fig. 6). Switching to a backup is also possible in an automated way, similar to the power supply when in case of disconnection of the main power source, the backup power is switched on either by

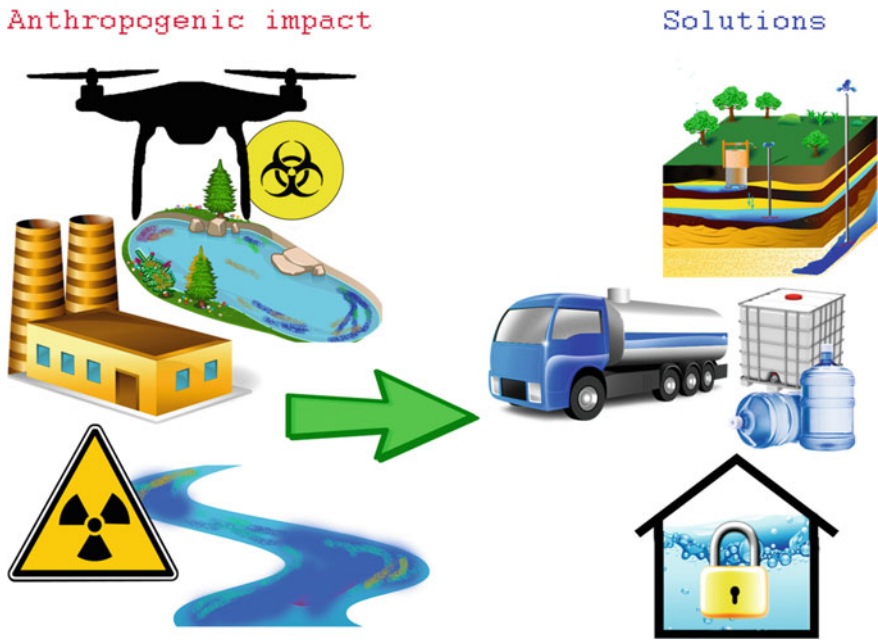


Fig. 6 Modeling potential water degradation scenarios

switching to an independent backup power line or by using a specialized installation like a fuel generator or solar/wind power and battery-stored backup power.

At the same time, it is often problematic for many cities to consider the possibility of quickly entering a backup water supply, especially if management systems do not provide for such a possibility. Unfortunately, many cities face this situation, which leads them to disrupt the stable functioning of systems and increase the risk of threats to water security. Active chemical and biological contamination, deliberate anthropogenic contamination of water sources for the purpose of terrorist and extremist activities, including sabotage [18], which are displayed in this scenario modeling, are potentially possible. Such studies allow us to justify the feasibility of modeling the functioning of CPS and the introduction of reservation systems in order to increase the level of strategic security of a modern city, improve the level of water resources control.

Recommendations based on scenario modeling:

1. Saint Petersburg is one of many cities that has a “single” source of water supply. Very often, cities are not provided with backup sources of water supply and in case of emergency situations—epidemics, biological infections, anthropogenic pollution, can cause significant economic and social damage. When building smart cities, it is necessary to take into account the fact of a “single” source and to forecast and analyze the scenario in the models in order to improve the strategic security of cities.
2. The presence of virtually the only source of water supply for the city of Saint Petersburg with more than 5 million inhabitants is unacceptable. The largest freshwater lake in Europe—Lake Ladoga, which is the source of the Neva River, is located near Saint Petersburg. This can potentially increase the security level to ensure water security of the region requires an increase in monitoring of compliance with measures of the purity of lake Ladoga, equipping CPS water for the purpose of operational monitoring of the chemical and biological status of the water body.
3. Consider the possibility of building closed protected freshwater storage facilities in Saint Petersburg.
4. Further research will include additional simulation modeling of various scenarios for the functioning of the cyber-physical water supply system in Saint Petersburg, taking into account the identified threats, trends, and the level of equipment wear, as well as integration with Smart city systems. Special attention should be paid to improving the level of security of both the functioning of systems and assessing the potential consequences due to the failure of supporting CPS.

4 Conclusion

Advanced technologies of smart city governance (including water resources management) are based on the use of cyber-physical systems. At the same time, the existing global challenge of improving water resources management is particularly acute in

the context of water source degradation, water shortages, and continuing population growth. This area of research is a pressing issue. The analysis of the theory and practice of smart city functioning as a set of cyber-physical systems allowed to conclude that it is advisable to refine the methods of classification of smart city cyber-physical systems. In particular, the application of the clustering approach of the cyber-physical systems of the smart city became the basis for the developed taxonomy of cyber-physical systems. This classification implies the decomposition of cyber-physical systems on the basis of control functions, providing consolidated construction of the cyber-physical system's topology and assessment of their interaction.

The authors created the basic models of functioning of cyber-physical water supply systems in a modern city and considered the specifics and features of the modeling process. When modeling cyber-physical systems of a smart city, it is necessary to take into account that these systems are multi-link active ones. By building models, it is possible to improve water management and reduce potential risks [20]. The development of integrated models of cyber-physical water supply systems functioning in a modern city requires additional research. The basic models of cyber-physical water supply systems simulated by the authors have been practically tested. Conducted scenario modeling implied disabling of the main water source in a city with more than 5 million inhabitants - on the example of the city of Saint Petersburg, Russia. In order to increase the strategic security of the region, a list of measures has been compiled to reduce the likelihood of water scarcity, including the creation of reserve protected closed-type freshwater storage facilities and equipping Lake Ladoga with cyber-physical water supply systems. These results are the basis for considering the improvement of smart city water supply management and the creation of simulation models for the scenario of the smart city's main water source disabling.

Acknowledgements The reported study was partially funded by RFBR according to the research project №19-01-00767 and by ICS RAS according to the state project.

References

1. United Nations website (Water Action Decade, 2018–2028). <https://www.un.org/sustainabledevelopment/water-action-decade/>. Accessed 24 Apr 2020
2. USEPA. Water audits and water loss control for public water systems, Report 816-F-13-002 (2013)
3. Galin, R.R., Meshcheryakov, R.V.: Review on human—robot interaction during collaboration in a shared workspace. In *Interactive Collaborative Robotics. ICR 2019*. LNCS, vol. 11659, pp. 63–74. Springer, Istanbul (2019)
4. McPhearson, T., Auch, R., Alberti, M.: Regional assessment of North America: urbanization trends, biodiversity patterns, and ecosystem services. In: *Urbanization biodiversity and ecosystem services: challenges and opportunities*, pp. 279–286. Springer, Dordrecht (2013)
5. Wu, J., Xiang, W.N., Zhao, J.: Urban ecology in China: historical developments and future directions. *Landscape and urban planning* **125**, 222–233 (2014)
6. Rosario-Ortiz, F., et al.: How do you like your tap water? *Science* 351(6276), 912–914 (2016)

7. Zhang, Q.Q., Ying, G.G., Pan, C.G., Liu, Y.S., Zhao, J.L.: Comprehensive evaluation of antibiotics emission and fate in the river basins of China: source analysis, multimedia modeling, and linkage to bacterial resistance. *Environ. Sci. Technol.* **49**, 6772–6782 (2015)
8. Liu, Y., Peng, Y., Wang, B., Yao, S., Liu, Z.: Review on cyber-physical systems. *IEEE/CAA J Automatica Sinica* **4**(1), 27–40 (2017)
9. Yang, S.H., Lyu, X., Ding, Y.: Safety and security risk assessment in cyber-physical systems. *Theory Appl., IET Cyber-Physical Systems* (2019)
10. Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovic, N., Meijers, E.: Smart cities-ranking of european medium-sized cities. *Rapport technique*, Vienna Centre of Regional Science (2007)
11. Burkov, V.N., Opoicev, V.I.: The allocation of resources in the active system, a collection of “active system” ICS RAS. Russia, Moscow (1973)
12. Novikov, D.A., Petrakov, S.N.: Course of the theory of active systems. SYNTEG, Moscow (1999)
13. Fomin, N.A.: Strategic innovative technologies as a tool to improve the control efficiency of water supply systems of a modern city. *IFAC-PapersOnLine*, vol. 52, Issue 25, 421–423. Elsevier, Sozopol, Bulgaria (2019)
14. Shcherbakov, M. V., Glotov, A. V., Cheremisinov, S. V.: Proactive and predictive maintenance of cyber-physical systems. In: Kravets A., Bolshakov A., Shcherbakov M. (eds) *Cyber-Physical Systems: Advances in Design & Modelling Studies in Systems, Decision and Control*. Springer, Cham, vol. 259, pp. 263–278. Springer, Cham (2020).
15. Resolution Of the government of Saint Petersburg of December 11, 2013 N 989 on approval of the scheme of water supply and sanitation of Saint Petersburg for the period up to 2025, taking into account the prospects up to 2030. <https://pravo.gov.ru/proxy/ips/?docbody=&prevDoc=131103190&backlink=1&&nd=131047822>. Accessed 25 Apr 2020
16. Yilmaz, E.N.: Byzantium period water architecture and a masterpiece in Istanbul: the big basilica cistern. *Turkish Neurosurg.* **24**(6), 823–827 (2014)
17. Water Code of the Russian Federation from 03.06.2006 N 74-FZ. https://www.consultant.ru/document/cons_doc_LAW_60683/814f76c933059091b59d1e16017ae944260a729e/. Accessed 25 Apr 2020
18. Bastrykin, A.I., Kvint, V.L., Kirilenko, V.P., Shamahov, V.A.: Extremism in the modern world. NWIM RANEP. Saint Petersburg, Russia (2018)
19. Leitao, P., Karnouskos, S., Ribeiro, L., Lee, J., Strasser, T., Colombo, A.W.: Smart agents in industrial cyber-physical systems. *IEEE* **104**(5), 1086–1101 (2016)
20. Iskhakov, A., Meshcheryakov, R.: Intelligent system of environment monitoring on the basis of a set of IOT-sensors. In: 2019 International Siberian Conference on Control and Communications (SIBCON), pp. 1–5, IEEE, Tomsk, Russia (2019)

Controlling Traffic Flows in Intelligent Transportation System



Alexander Galkin  and Anton Sysoev 

Abstract The approach to control traffic flows in intelligent transportation systems is proposed. The algorithm is based on optimization of the transportation system functioning criterion which is speed (or time) of movement. The system is represented as a graph. The control consists of changing the traffic flows rate on individual sections of the system (graph edges), for example, by regulating the operation of traffic lights, and changing the capacity of sections, for example, by using reverse lanes.

Keywords Intelligent transportation systems · Effectiveness criteria · Optimization · Control · Graphs

1 Introduction

Regional transportation systems have a significant impact on social and economic development of the region. Therefore, the issues of modernizing transportation systems are always relevant. The upgrade may involve powerful infrastructure changes [cf. 1]: construction of new roads, transport hubs, development of new modes of transport, etc. These activities require significant financial costs and a long time to implement them.

Effective control of existing transportation infrastructure is another area of development of transportation systems, and, as a rule, the cost of implementing such measures is much lower, since no significant infrastructure changes are required. The duration of their implementation is also shorter. But such events have a limited scale of positive effects.

A. Galkin (✉) · A. Sysoev
Lipetsk State Technical University, Moskovskaya str. 30, Lipetsk 398055, Russia
e-mail: galkin_av@stu.lipetsk.ru

A. Sysoev
e-mail: sysoev_as@stu.lipetsk.ru

The relevance and possibilities of effective control of transportation systems have increased with the development of information and telecommunications technologies [2]. The intellectualization of transportation systems is developing all over the world [3], and various technologies are being used. But all of them are based on collecting information about the current state of transportation systems and forecasts of the future state [4]. This information may include [5]: geo-information data (road and road network diagram, transport nodes, terrain, buildings, traffic control equipment, traffic flow diagram), traffic flow characteristics (speed, flow rate, density), information about meteorological characteristics of the environment (relative humidity, air temperature, pressure, precipitation), and others. The distinctive features of each intelligent transportation system are the ways in which such data is collected, processed, and used in control decisions [6]. Almost all transportation systems use graph structures to describe the road scheme, which is a universal tool in this case allowing to solve optimization problems [7]. The vertices of the graph can describe the transport nodes corresponding to, for example, the intersection of roads. And the edges describe the roads themselves.

2 Controlling Traffic Flows

2.1 *Graph-Structural Approach to Modelling of Transportation Systems*

Each vertex of the graph can be compared with a vector value that describes its various characteristics, such as the presence/absence of a traffic light or its signal, the incoming traffic flow rate, capacity, the value of carbon dioxide emissions, etc. For each edge of the graph, we can also match its own vector value $e^{(i,j)} = [x_i; x_j] = (e_1^{(i,j)}, e_2^{(i,j)}, \dots, e_m^{(i,j)})$, appropriate, for example, for traffic capacity, traffic flow rate, etc. In general, these characteristics are not constant and can change over time. It is also assumed that some of these characteristics can be controlled. For example, by adjusting traffic lights, it is possible to change the capacity of nodes, and by switching reverse roads, it is possible to change the capacity of roads and, as a result, the flow rate of incoming traffic at transport nodes [cf. 8]. The most common graph-structural object, modelling the traffic flows, is the transportation network [9, 10]. The main tasks include the equilibrium distribution in the transportation network [11, 12] and the search for the maximum flow [13, 14]. Control in an intelligent transportation system consists of setting conditions that optimize certain characteristics [15] of the transportation system state. In this case, these characteristics act as optimality criteria. There are many types of optimality criteria [16]: traffic safety, environmental impact, etc. But one of the main things is the speed (time) of movement. One of the possible values that characterizes this criterion is

$$k(e^{(k)}) = \frac{\lambda(x^{(i)})}{\mu(e^{(k)})}, \quad (1)$$

where $\lambda(x^{(i)})$ is the flow rate entering the vertex i , which is the beginning of the edge k , and $\mu(e^{(k)})$ is the capacity of the edge k . If the value of this criterion is less than 1, the movement is free and the speed is limited by permission signs. The criterion for the entire transportation system can be defined as follows

$$K_1 = \min \sum_{k=1}^m c(e^{(k)}), \quad (2)$$

where

$$c(e^{(k)}) = \begin{cases} 1, & k_1(e^{(k)}) > 1; \\ 0, & k_1(e^{(k)}) \leq 1. \end{cases} \quad (3)$$

Thus, the optimal situation is one in which the flow rate of incoming traffic does not exceed their capacity on as many edges corresponding to roads as possible. Let us consider the problem of choosing optimal controls for a transportation system, defined as a graph, and characterized by criterion (2). Let the graph have two parameters for each vertex i : $x_1^{(i)}$ is the flow rate of the incoming streams and $x_2^{(i)}$ is the maximum capacity, and each edge also contains two parameters $e_1^{(i,j)}$: $e_1^{(i,j)}$ is the flow rate of the incoming streams and $e_2^{(i,j)}$ is the maximum capacity. Let us assume that different traffic light control modes can be used as control actions, which change $x_2^{(i)}$ and reverse road switching, which change $e_2^{(i,j)}$.

Parameters set for graph vertices and edges are related. Figure 1 shows possible variants of the transport node scheme.

For the option (a) in Fig. 1 the link has the following form

$$\alpha_1 e_1^{(i,i_{out1})} + \alpha_2 e_1^{(i,i_{out2})} + \dots + \alpha_n e_1^{(i,i_{outn})} = \min(e_1^{(i_{in},i)}, x_2^{(i)}), \quad (4)$$

where $\alpha_1, \alpha_2, \dots, \alpha_n$ are weight coefficients that characterize the separation of traffic flows along different edges that exit from the same node.

For the option (b) in Fig. 1 the link has the following form

$$e_1^{(i,i_{out})} = \min(e_1^{(i_{in1},i)} + e_1^{(i_{in2},i)} + \dots + e_1^{(i_{inm},i)}, x_2^{(i)}). \quad (5)$$

For the option (c) in Fig. 1 the link has the following form

$$\begin{aligned} & \alpha_1 e_1^{(i,i_{out1})} + \alpha_2 e_1^{(i,i_{out2})} + \dots + \alpha_n e_1^{(i,i_{outn})} \\ & = \min(e_1^{(i_{in1},i)} + e_1^{(i_{in2},i)} + \dots + e_1^{(i_{inm},i)}, x_2^{(i)}). \end{aligned} \quad (6)$$

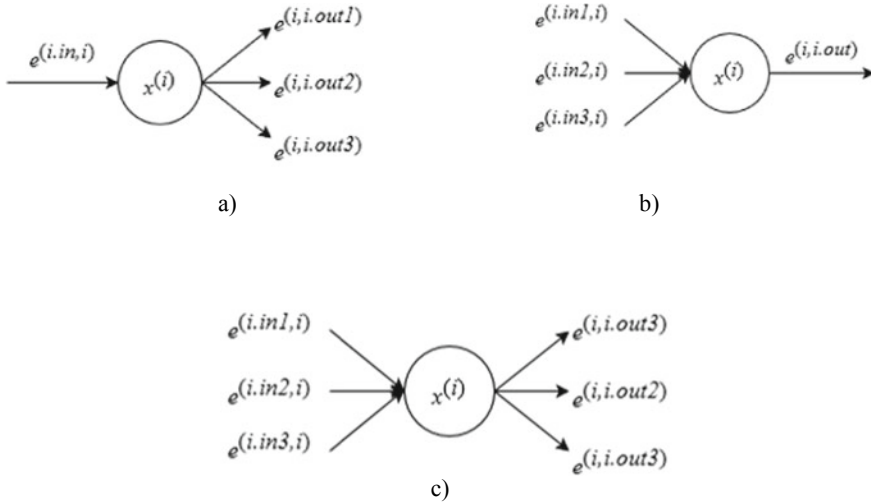


Fig. 1 Schemes of transport vertices: **a** one input—several outputs, **b** several inputs—one output, **c** several inputs—several outputs

However, it should be kept in mind that for the option (a), if $e_1^{(i_{in},i)} > x_2^{(i)}$, changes the maximum capacity of the edge $e^{(i_{in},i)}$ as follows $e_2^{(i_{in},i)} = x_2^{(i)}$. For options (b) and (c): if $e_1^{(i_{in1},i)} + e_1^{(i_{in2},i)} + \dots + e_1^{(i_{inm},i)} > x_2^{(i)}$, the maximum capacity of the edges changes $e^{(i_{in1},i)}, e^{(i_{in2},i)}, \dots, e^{(i_{inm},i)}$ as follows $e_2^{(i_{in1},i)} = \beta_1 x_2^{(i)}, e_2^{(i_{in2},i)} = \beta_2 x_2^{(i)}, \dots, e_2^{(i_{inm},i)} = \beta_m x_2^{(i)}$, where $\beta_1, \beta_2, \dots, \beta_m$ are weight coefficients that characterize the share of transport vertex capacity allocated to the corresponding incoming edge.

In order to account for the capacity of only edges, each vertex $x^{(i)}$ of the highway graph can be split into two vertexes— $x_{in}^{(i)}$ and $x_{out}^{(i)}$, at the same time setting a new edge $e^{(i_{in};i_{out})} = (x_{in}^{(i)}, x_{out}^{(i)})$. Then all incoming edges in the vertex $x^{(i)}$ will be included in $x_{in}^{(i)}$, while the outgoing ones will come from $x_{out}^{(i)}$. The vertex capacity of $x^{(i)}$, divided into two, will become the capacity of the edge $e^{(i_{in};i_{out})} = (x_{in}^{(i)}, x_{out}^{(i)})$. If there are vertices $x^{(j)}$, such that the graph has edges $e^{(i,j)}$ and $e^{(j,i)}$, then the edge $e^{(i_{out};i_{in})} = (x_{out}^{(i)}, x_{in}^{(i)})$ is also added.

2.2 Approach to the Redistribution of Traffic Flows in Transportation System

Control by adjusting the operating modes of traffic lights and switching reverse roads leads to changes in the capacity of individual graph edges. The total capacity of adjacent edges remains constant: $\sum_{j \in X} e_2^{(i,j)} + \sum_{j \in X} e_2^{(j,i)} = const$, where X is the vertex set of the graph.

Let us set the matrix of maximum edge capacities for a graph

$$E_2 = \begin{bmatrix} e_2^{(1,1)} & e_2^{(1,2)} & \dots & e_2^{(1,n)} \\ e_2^{(2,1)} & e_2^{(2,2)} & \dots & e_2^{(2,n)} \\ \dots & \dots & \dots & \dots \\ e_2^{(n,1)} & e_2^{(n,2)} & \dots & e_2^{(n,n)} \end{bmatrix} \quad (7)$$

and the matrix of traffic flow rate

$$E_1 = \begin{bmatrix} e_1^{(1,1)} & e_1^{(1,2)} & \dots & e_1^{(1,n)} \\ e_1^{(2,1)} & e_1^{(2,2)} & \dots & e_1^{(2,n)} \\ \dots & \dots & \dots & \dots \\ e_1^{(n,1)} & e_1^{(n,2)} & \dots & e_1^{(n,n)} \end{bmatrix}. \quad (8)$$

Let us compose the vector U of dimension n with elements presented as $\sum_{j \in X} e_1^{(j,i)} - \sum_{j \in X} e_2^{(i,j)}$ (difference between the sum of elements of the column i of the matrix (8) and the column i of the matrix (7)). The positive value of element means that the intensity of movement along the edges of the vertex i is higher than the capacity of the edges exiting it. The negative value indicates that there is a reserve of capacity of the edges emanating from the vertex i .

Next for each element of the resulting vector, we perform the following actions. If $U_i > 0$, we choose all the edges $e^{(i,j)}$, such as $U_j < 0$. Then the overall ability to increase the capacity of the node i will be $\sum_{j \in X_j} |U_j|$, where X_j is the set of vertices $x^{(j)}$, adjacent the vertex $x^{(i)}$.

For selected edges, according to the control capabilities, we increase the capacity as follows:

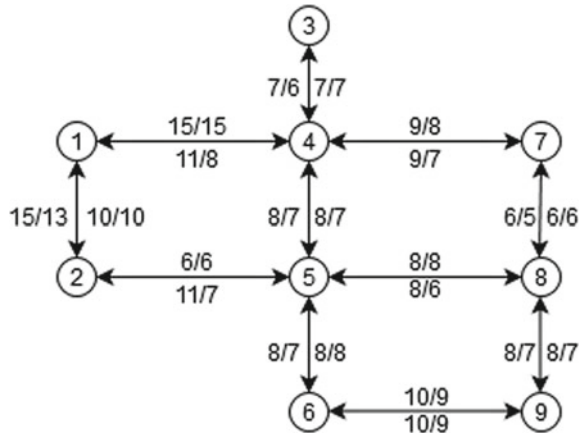
$$e_{2new}^{(i,j)} = e_2^{(i,j)} + \frac{|U_j|}{\sum_{j \in X_j} |U_j|} \min[|U_i|, \sum_{j \in X_j} |U_j|]. \quad (9)$$

This method is used to unload congested areas without loading them. The direct adjustment of the operating modes of the traffic light and reverse motion at the specified capacity capabilities is carried out according to the specialized algorithms [17–19]. The recalculation procedure (9) can also be repeated, if one considers the range of capacity not only of adjacent vertices, but also of vertices to which one can construct a route limited to a certain number of edges.

2.3 Numerical Example

Let us consider the example of traffic flow management in the transport network presented in Fig. 2.

Fig. 2 Transportation network graph



The structure of the graph corresponds to the road sections in Lipetsk The lateral orientation of the edges in Fig. 1 corresponds to two opposite edges. Arrow signs stand for the the edge traffic flow respectively. The label above refers to an edge from left to right, the label at the bottom—from right to left. The label on the left refers to an edge with a direction from the top to the bottom, and the label on the right refers to the direction from the bottom to the top. In addition, the capacity of the vertices $x^{(2)} = [25; 30; 8; 32; 30; 15; 14; 20; 20]$ is specified. It corresponds to the total possibilities of the intersection determined by the given vertex when organizing movement in all possible directions. In this case, the redistribution of the total capacity of the vertices in all directions can be considered as a test to check its effectiveness. One of the most common ways to reach this is to adjust traffic lights.

To connect this section of roads to the entire urban transportation system is ensured by the formation of additional incoming traffic at vertices 1, 2, 7, 9 and the capacity of those leaving the transportation network at the same points. This can be taken into account by putting an extra vertex to the graph, as shown in Fig. 3.

In fact, the new vertex 0 is the analogue of the source and run-off in the classical transportation network. To implement the proposed flow control algorithm, a number of changes must be made to the graph: (1) each vertex having multiple inputs and multiple outputs from the same vertices (see option c) in Fig. 1), must be split into as many vertices as there are exits. The resulting vertices have one exit and contain all inputs except the input from the vertex where the exit is going. This is necessary to ensure that the flow from the input vertex is not redistributed to the vertex itself; (2) to transfer the capacity from vertices to edges, each vertex must be split into two. The first will include all edges. One exit with the capacity of the shared vertex will be directed to the second. The other one will have all the edges coming out. The example of this partition for the first vertex is shown in Fig. 4.

Vertex 1 was split into three: 1.0, 1.2 and 1.4. The second character in the number defines the vertex with which the output edge is connected. In turn, each of the three vertices obtained was split into two to transfer the capacity from the vertex to the

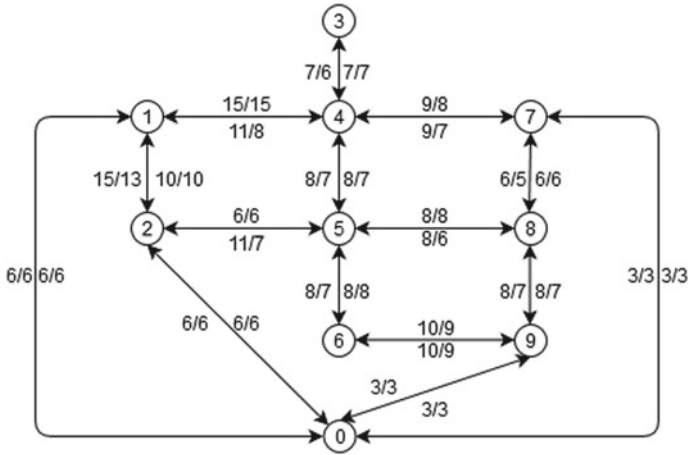
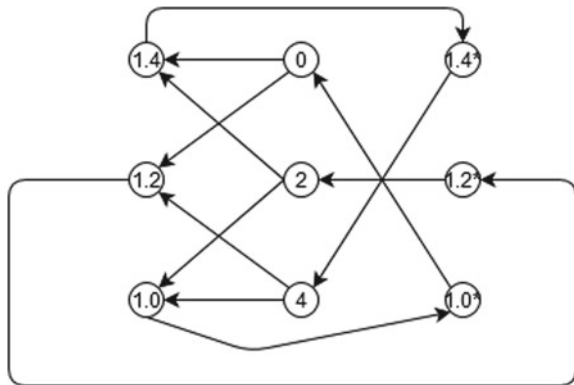


Fig. 3 Modified transportation network graph

Fig. 4 Partition of vertex 1



edges. The total capacity of the edges [1.0; 1.0*], [1.2; 1.2*] and [1.4; 1.4*] is equal to the capacity of vertex 1. The edge distribution is determined by the current traffic light setting. In the initial conditions of the task, let's assume that the capacity is proportional to the weight of the output edge among all exits, for example,

$$e_2^{(1.0,1.0^*)} = x_2^{(1)} \cdot \frac{e_2^{(1,0)}}{e_2^{(1,0)} + e_2^{(1,2)} + e_2^{(1,4)}} = 25 \cdot \frac{6}{6 + 15 + 15} = 4, 17.$$

Severed edge intensity values, e.g., [0; 1.2] and [0; 1.4], are determined by the weighting factors that characterize the division of the traffic across the different edges leaving the same vertex. Under the conditions of the problem, let the weights be proportional to the weight of the output edge among all exits, for example,

0	0	3	3	0	0	0	0	3.75	2.25	0	0	0	0	...
0	0	0	0	4.167	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	10.417	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	10.417	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	7.5	0	7.5	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	8.182	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	13.636	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	8.182	0	
6	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	2.857	0	7.143	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	3.143	7.857	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
⋮														⋮

Fig. 5 Capacity matrix fragment

$$e_1^{(0,1,2)} = e_1^{(0,1)} \cdot \frac{e_2^{(1,2)}}{e_2^{(1,2)} + e_2^{(1,4)}} = 6 \cdot \frac{15}{15 + 15} = 3.$$

In the end, capacity and flow rate matrices are determined for a graph (Fig. 3) of size $[52 \times 52]$. Figures 5 and 6 show the fragments of the obtained matrices.

The value of the vector U is calculated from the obtained matrices as follows:

$$U = [0.; 0.97619; -1.70238; -0.27381; -1.83333; -6.28571;$$

$$-4.85714; 0.943182; -5.51136; 0.568182; 2.18182; -1.875;$$

$$2.18182; 0.; -1.88661; 1.67443; 1.55385; 1.65833; -2.82947;$$

$$-0.6; 0.685714; -0.771429; -4.13228; -1.00529; -3.07937;$$

$$-2.78307; 0.407407; -1.85185; -0.0740741; 0.518519; 2.33333;]].$$
 The

$$-1.33333; -1.33333; -3.; -0.833333; -0.7; -3.46667; 3.83333;$$

$$-2.7; 5.86667; -0.272727; 0.974026; 0.298701; -1.5;$$

$$0.428571; -0.928571; -1.01099; -1.7033; -6.28571;$$

$$3.52448; -2.94872; 6.42424$$

value of the criterion obtained is $\sum_{k=1}^{52} c(e^{(k)}) = 18$.

Among the many positive elements selected are those whose capacity can be modified. These elements are $[2.5; 4.3; 4.5; 4.7; 6.5; 8.7; 8.9]$. So for the elements 2.5 and 8.7 there are no edges $e^{(2.5,j)}$ and $e^{(4.8,j)}$, such as $U_j < 0$, and their capacity cannot be increased. Among the remaining edges the edge i is selected, for which the maximum value is $\sum_{j \in X_j} |U_j| - U_i$, where X_j is the set of vertices $x^{(j)}$, adjacent the vertex $x^{(i)}$, with negative values of U_j . Such a vertex is 8.9. So we recalculate

0	0	3	3	0	0	0	0	3.75	2.25	0	0	0	0
0	0	0	0	4.167	0	0	0	0	0	0	0	0	0
0	0	0	0	0	8.714	0	0	0	0	0	0	0	0
0	0	0	0	0	0	10.143	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	6.5	0	6.5	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	8.182	0	0	0
0	0	0	0	0	0	0	0	0	0	0	8.125	0	0
0	0	0	0	0	0	0	0	0	0	0	0	8.182	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2.857	0	7.143	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2.286	5.714	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	7
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	2.625	4.375	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
⋮													

Fig. 6 Traffic flow rate matrix fragment

the new edge capacity with formula (9):

$$\begin{aligned}
 e_{2new}^{(8.9,8.9*)} &= e_2^{(8.9,8.9*)} + \frac{|U_j|}{\sum_{j \in X_j} |U_j|} \min[|U_{8.9}|, \sum_{j \in X_j} |U_j|] \\
 &= 7.272727 + 0.298701 = 7.571428
 \end{aligned}$$

That is the edge capacity must be increased by 0.298701. Increase in capacity in the example presented can only be achieved by redistributing the total capacity of the vertex between its outgoing edges (traffic lights regulation). In this case, the limit on the value of the total capacity of the vertex should be respected. In our case $e_{2new}^{(8.5,8.5*)} + e_{2new}^{(8.7,8.7*)} + e_{2new}^{(8.9,8.9*)} = const$. The edge $e_2^{(8.5,8.5*)}$ has the capacity reserve of 0.772727.

As a result, the following capability modifications are made:

$$e_{2new}^{(8.9,8.9*)} = 7.272727 + 0.298701 = 7.571428,$$

$$e_{2new}^{(8.5,8.5*)} = 7.272727 - 0.298701 = 6.974026.$$

We get the value of the criterion $\sum_{k=1}^{52} c(e^{(k)}) = 17$. and repeat this procedure with new capacity and intensity matrices as long as the criterion is reduced. In the example given, the value of the criterion can be reduced to 14.

3 Conclusion

Thus, the traffic control algorithm in Intelligent Transportation Systems is presented in this chapter. The intelligent transportation system is expected to collect and analyze real-time traffic information. This information is used by the traffic control algorithm. The algorithm is based on optimization of the performance criterion of the transportation system—speed (time) of movement. In the above example, using the algorithm it is possible to reconfigure traffic lights modes in such a way that the total number of congestion sections decreases from 18 to 14. In case of optimizing not only current state of the system but also taking into account the future possible states the presented approach could be extended to time-dependent graphs described in [20].

Acknowledgements The study was carried out with a grant from the Russian Science Foundation (Project 18-71-10034).



References

1. Krylatov, A.Yu.: Optimal strategies for road network's capacity allocation. *Vestnik of Saint Petersburg University. Appl. Mathem. Comput. Sci. Contr. Process.* **13**(2), 182–192 (2017)
2. Zhu, L., Yu, F.R., Wang, Y., Ning, B., Tang, T.: Big data analytics in intelligent transportation systems: a survey. *IEEE Trans. Intell. Transp. Syst.* **20**, 383–398 (2019)
3. Dimitrakopoulos, G., Uden, L., Varlamis, I.: *The future of intelligent transport systems*. Elsevier (2020)
4. Huang, D., Bai, X.: A Wavelet neural network optimal control model for traffic-flow prediction in intelligent transport systems. In: *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence. ICIC 2007. Lecture Notes in Computer Science*, vol 4682. Springer, Berlin, Heidelberg (2007)
5. Chandra, Y. R. V. S., Shiva Harun, M., Reshma, T.: Intelligent transport system. *Int. J. Civil Eng. Technol.* **8**, 2230–2237 (2017)
6. Sysoev, A.S., Khabibullina, E.L.: Forming production rules in intelligent transportation system to control traffic flow. *Open Semant. Technol. Intell. Syst.* **4**, 317–322 (2020)
7. Sicuro G.: *Graphs and optimization*. In: *The Euclidean Matching Problem*. Springer Theses (Recognizing Outstanding Ph.D. Research, pp. 5–23. Cham, Springer (2017)
8. Portilla, C.R., Cortes, L.G., Valencia, F., Lopez, J.D., Espinosa, J.J.: Optimal control of urban and highway traffic. In: *2012 IEEE Colombian Intelligent Transportation Systems Symposium (CITSS)*, pp. 1–6. Bogota (2012)
9. Iri, M. (1996) *Network flow—Theory and applications with practical impact*. In: *Proceedings of the Seventeenth IFIP TC7 Conference on System Modelling and Optimization*, pp. 24–36. Boston, Springer (1996)
10. Steenbrink, P.A.: *Optimization of transport networks*. Wiley, New York and London (1974)

11. Omarova, G.A., Chernov, K.Y.: Design and research of transport algorithms. *Inform. Probl.* **3**, 27–36 (2014)
12. Chistyakov, P.: Integrated transport system: analytical review. Centre for Strategic Development, Moscow (2018)
13. Shvetsov, V.I.: Algorithms of transport flow distribution. *Autom. Telemekh.* **10**, 148–157 (2009)
14. Skorokhodov, V.A., Chebotareva, A.S.: The maximum flow problem in a network with special conditions of flow distribution. *J. Appl. Ind. Math.* **9**, 435–446 (2015)
15. Rassafi, A. A., Vaziri, M.: Sustainable transport indicators: definition and integration [In Russian: Ustoychivyye transportnyye pokazateli: opredeleniye i integratsiya]. *Int. J. Environ. Sci. Technol.* 83–96 (2005)
16. Galkin, A., Sysoev, A.: Formalizing criteria of intelligent transportation and logistic systems functioning. *Transp. Res. Proc.* **45**, 514–521 (2020)
17. Sysoev, A.S.: Analysis of approaches to calculating optimal traffic cycle time. *Inf. Technol. Modell. Manag.* **4**, 300–307 (2012)
18. Zhao, Z., Chen, W., Wu, X., Chen, P.C.V., Liu, J.: LSTM network: a deep learning approach for short-term traffic forecast. *IET Image Proc.* **11**, 68–75 (2017)
19. Transportation networks. *Mathem. Sci. Eng.* **90**, 1–15 (1972)
20. Wang, Y., Yuan, Y., Ma, Y., et al.: Time-dependent graphs: definitions, applications, and algorithms. *Data Sci. Eng.* **4**, 352–366 (2019)

Development of Scenarios for Modeling the Behavior of People in an Urban Environment



Alexander Anokhin, Sergey Burov, Danila Parygin , Vyacheslav Rent, Natalia Sadovnikova , and Alexey Finogeev

Abstract The aim of the study is to improve intelligent methods for supporting city management tasks by monitoring the state of processes in the urban environment and deliberately changing their parameters in accordance with decisions obtained using predictive modeling. The chapter provides an analysis of the current state of the cyber-physical problem of modeling processes in systems with people interaction, existing methods for modeling the people movement in an urban environment, and projects for modeling the people movement in a city based on a multi-agent approach. The process of developing scenarios for moving agents in an urban environment is shown. The main components of the software solution responsible for simulating human behavior are presented.

Keywords Multi-agent modeling · Urban environment · Behavior scenarios · Decision support · Urban mobility · Modeling system

A. Anokhin (✉) · S. Burov · D. Parygin · V. Rent · N. Sadovnikova
Volgograd State Technical University, Lenina Ave. 28, 400005 Volgograd, Russia
e-mail: alex.anokhin.st@gmail.com

S. Burov
e-mail: sergey.burovic@gmail.com

D. Parygin
e-mail: dparygin@gmail.com

V. Rent
e-mail: vyacheslav61g@gmail.com

N. Sadovnikova
e-mail: npsn1@ya.ru

A. Finogeev
Penza State University, Krasnaya Str. 40, 440026 Penza, Russia
e-mail: alexeyfinogeev@gmail.com

1 Introduction

Decisions effectiveness improving when choosing options for urban environment development is one of the main problems of modern large cities. New approaches to the management of urban processes are required because of the increase in population, the complexity of urban processes, and the tightening of requirements for the quality of urban services provided. Models suitable for studying various options for changing the urban environment, organizing activities, and choosing the most effective solutions can be built on the basis of an analysis of the functions of the city, the processes, and behavior of residents in it [1].

The multi-agent approach is considered among many methods for modeling such systems. Such an approach is to reduce the initial complex problem into a set of simple tasks, and each “simple” problem is solved by a special program called an agent. An agent is an independent software system that consists of program objects that have the ability to receive an impact from the outside world, determine its reaction to this impact and, in accordance with this, form a response action.

Agents as software entities are able to perceive the environment and respond depending on the situation in which they are [2]. They are goal oriented, i.e. they receive tasks and perform them, interacting with each other and with the environment.

The problem of increasing the adequacy of the developed models requires the development of new approaches to modeling the behavior of actors taking into account the interaction with the environment, other actors and the implementation of a wide range of behavioral scenarios.

2 Existing Approaches to Modeling Human Behavior

Multi-agent modeling is currently an actively developing and promising area of IT [3]. A multi-agent system is a system of several interacting intelligent agents (programs). Entities in such a system are active (agents that respond to each other and the environment) and have individual characteristics. Modeling takes place from the bottom up, thus, behavior at the individual level forms the global level of the system, which allows achieving from low to high levels of abstraction. The simplicity of the initial implementation of such a system is due to incomplete data on the environment and the agents themselves that may cause subsequent difficulties in the interactions formalization [4].

The agent-based approach has been shown to be effective in describing a wide variety of processes. Nevertheless, there are too many unsolved problems associated with the implementation and ensuring the adequacy of multi-agent systems [5]. So, there are many approaches to modeling the behavior of intelligent agents in multi-agent systems. It makes sense to consider the main ones.

2.1 *Modeling the Behavior of Intelligent Agents in Multi-Agent Systems*

Rule Based Behavior Model. The organization of a rule-based behavior model assumes that the agent will behave identically in each similar situation, guided by the rule that the programmer set for him. Rules can be entered based on the current state of the agent, or they can be absolute and not depend on it. Setting rules based on the state of the agent or its parameters allows to provide some variety of behavior, but can't achieve realistic behavior [6].

Behavior Model Based on the Finite State Machine. The behavior of an agent with various states can be modeled on the basis of finite state machines. The state may reflect the conditions in which the agent is: hunger, thirst, fatigue, and others. The most famous variant of this approach is called GOAP and is successfully used in game projects on a global scale (F.E.A.R., S.T.A.L.K.E.R and others) [7]. This approach implies that the activity of the agent is regulated by his state and the purpose of this state. For example, the agent in a state of "hunger" must achieve the goal of "eat" before it can switch to another state. The need for an agent to achieve a goal determines the need to build a chain of actions aimed at achieving this goal.

Finite state machines allow you to simulate a fairly varied and realistic behavior of the agent. However, the addition of new states to the model involves a substantial revision of the entire model in order to add new relationships between the new and each of the existing states [8].

Behavioral Model Based on Behavior Tree. The behavior tree is an oriented acyclic graph whose nodes are agent behavior options. The tree starts from the root node, which sends the child a signal for execution and receives the state of the nodes in response: "Running" if the node is still running, "Success" in the case of successful implementation and "Failure" in case of failure. Variable agent behavior can be modeled using such a system. Also, the behavior tree, being a kind of finite state machine, has an undeniable advantage in the form that with the increase in the number of available states, the complexity of the tree does not grow as fast as the complexity of the automaton [9, 10].

Model of Social Forces. The model is based on Newtonian dynamics to describe the movement of pedestrians and shows several natural behavioral phenomena of pedestrians in the process of movement: pedestrians choose the shortest path; move at an individual speed, taking into account the situation, gender, age, restrictions; keep a certain distance from each other. The distance depends on the pedestrian flow density and speed. The pedestrian movement in the model of social forces is described by the sum of the forces acting on it. The position in space, the speed and acceleration of a pedestrian at any given time can be learned by compiling and solving a system of differential equations describing the action of social forces [11, 12].

The necessary accuracy of calculations and sufficient realism of pedestrian behavior can be obtained with the right settings. But the operation of the algorithm is characterized the by low speed with the high complexity of implementation, as

well as the need for high computing power. Moreover, the model has a low level of abstraction (micro level) [13].

2.2 Application of Multi-Agent Modeling in Gaming and Research Systems

It is also worth considering specific applications of multi-agent modeling. There are a number of applications designed to solve modeling problems in transport systems. Among such applications can be noted TRANSIMS, MIRO, MobiSim, ARCHISIM, SimMobility Freight and others.

There are also alternative multi-agent platforms such as GAMA, MadKit, Repast, Jade and NetLogo, which include basic components for creating agents from any data set, as well as the ability to perform large-scale modeling with millions of agents [14]. High-level agent programming languages, such as NetLogo or GAML, are offered to refine such systems.

The platforms under consideration can be universal, like GAMA, or highly specialized, like MATSim [15, 16], which has tools for modeling traffic flow using queue theory. The main advantage of the platforms under consideration over competitors is the use of multi-agent modeling, due to which the possibility of micromodeling of traffic flows and the environment is achieved [17, 18].

A multi-agent approach is also used in most computer games to simulate the behavior of non-player characters (NPC) [19]. It makes sense to pay attention to some specific games in the context of considering ways of modeling the behavior of the urban population.

Thus, in the series of games The Elder Scrolls [20], the Radiant AI game artificial intelligence system [21] is implemented, which describes the behavior of NPC and their behavior throughout the day. The behavior of the characters is subject to a plan according to which at a certain time of the day the character sleeps, takes food or walks around the city, while visiting places of interest and meeting with other characters for short conversations. This approach to modeling enables to create a realistic model of pedestrian movement in a settlement.

The Grand Theft Auto [22] series of games implements a larger-scale model of the city, which includes, in addition to pedestrians, vehicles, rail and air services. The optimization in this series is designed so that it does not clearly track the movement of all agents: characters that are not in the player's field of vision do not physically exist in the game, and appear only when they enter the field of view. A traffic system in which cars drive only on roads and observe traffic signals can be noted among the advantages of the model implemented in these games.

More realistic traffic system implemented in the Mafia game series [23]. Cars in these games observe the speed limit, independently determine from which row they can turn, and wait for the police in the event of an accident. In addition, the game is

able to track traffic violations: police can write a fine or detain a player for speeding, driving in the oncoming lane or through a red traffic signal, and for traffic accidents.

Multi-agent modeling can be successfully used in games where it is designed to diversify the game environment and provide a comfortable immersion of the player in its world, and in serious geographic information systems (GIS). The ability to model each agent with special properties individually is actively used in GIS. Due to this, it is possible to carry out accurate modeling of urban processes and predict changes in the urban environment.

3 The Proposed Approach to the Development of an Agent's Behavior Model in an Urban Environment

An adequate model in some approximation, which corresponds to the simulated part of the space, can be obtained by describing the basic rules of interaction and the necessary properties of agents. Each agent in the developed system is a model of a person (an actor with free will) or other entities of a living and artificial origin, who are participants in a conditionally dependent dynamic interaction and capable of transporting material and information resources.

The proposed method for constructing a model of state change allows working with events of a measurable scale. Rules of behavior are based on the essence of the roles and properties of the personality that shape its needs. The realization of needs launches certain scenarios of behavior, which, ultimately, are expressed in the sequence of actions of agents. The consequence of each action is a certain state in which the actor goes into and which causes a reaction due to rules, physical, social, legal and other restrictions in accordance with the conditions of the real world. A set of behavioral patterns of actors forms an integral model of the macro level [24].

An approach using the BDI paradigm [25] (a model of beliefs, desires, and intentions) was used in this study.

Beliefs is information about the patterns and condition of the environment that an agent can receive. The assumption is made that the agent's information about the world can be erroneous and incomplete, therefore it is only the agent's view of the world, and not reliable fundamental laws and environmental information.

Desires is the aggregate of all the goals that the agent would like to achieve. Desires must be consistent with each other. It is assumed that the agent will not be able to realize all his desires (goals), so he must limit their list and choose the most important ones.

Intentions is a collection of plans and scenarios for achieving desires. Intentions determine the direction of activity. The agent is trying to find ways in which he could fulfill his intentions. Intentions limit future choices, because an agent cannot create and accept new intentions that are incompatible with those that are already in the process of implementation. Intentions have a long life. The agent will make new

plans in case of failure of the current, until it succeeds. An intention can be deleted from the list only if the agent realizes that the intention is not feasible [26].

The BDI approach is based on the desire to analyze a person's mental activity, processes and understanding how and on what basis people make decisions. The basic algorithm of the BDI agent actions is as follows:

1. Determination the goals to which the agent will strive.
2. The choice of those goals, the implementation of which the agent will try to achieve.
3. Determination of methods and scenarios for achieving selected goals.

The behavior of an agent can change under the influence of factors such as weather conditions in the area of its location, the presence in the field of activity of certain physical objects, as well as the activities of other agents. The result of choosing a scenario for achieving goals based on a combination of data from the world and other agents is the impact: the movement of a person in space, as well as the opposite effect on other agents.

4 Development of the Module for Human Behavior Simulation

It was decided to implement a client–server architecture in order to separate the functionality of the application. The C# programming language, which has proven itself to be the best in the speed of developing complex projects, was used to write the server side of the system.

The client part of the system is written using the JavaScript language. The OpenLayers open source library was chosen as the framework for working with the map [27]. The WebSocket communication protocol is used to implement closer interaction between the browser and the platform, with the ability to work in interactive mode with support for real-time applications.

“.NetTopologySuite” framework used to work with map objects [28]. This framework is a port of the JTS Topology Suite framework written in Java.

Geometry classes support modeling points, lines, polygons, and collections. The geometry is linear in the sense that the boundaries are determined by linear interpolation between the vertices. The geometry is embedded in a 2-dimensional Euclidean plane. The vertices of the geometry can also have a value of Z .

User-defined accuracy models are supported for geometric coordinates. Computations are performed using algorithms that provide reliable geometric calculations for all exact models.

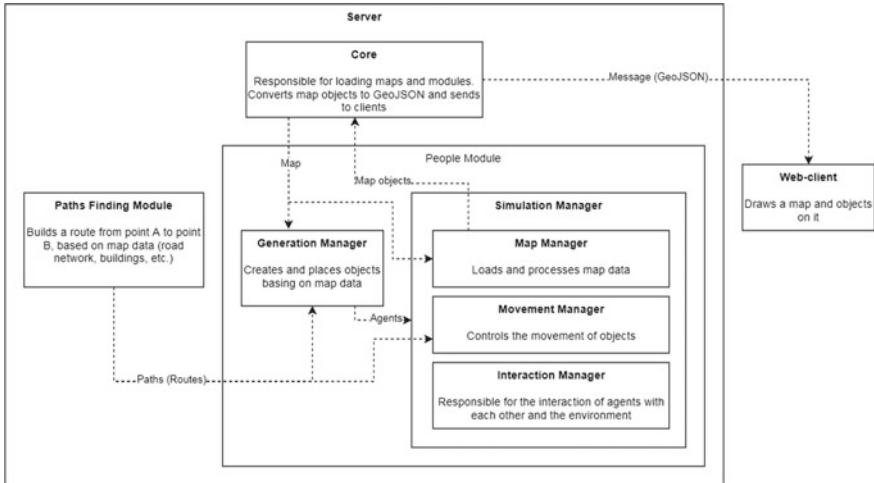


Fig. 1 Brief diagram of system components in the form of UML 2.0 components diagram

4.1 System Architecture

An extensible motion simulation system has been developed to meet the requirements described above. The architecture of this system is client–server, and also divides the program into a kernel and modules that implement individual tasks. This architecture is shown in Fig. 1.

The main part of the system is the server, which is responsible for the simulation. MapManager loads the map data on which the simulation will take place before starting the simulation. GenerationManager creates all objects and agents and places them on the map using its submodules. After that, each agent has a destination on the map to which it must get.

MovementManager is responsible for moving agents around the map and updating the values associated with it. The InteractionManager module is used to implement the interaction of agents with objects and with each other. StatisticManager collects statistics for all agents and objects in the system from the moment the simulation starts, and also provides access to statistics data from past simulations.

4.2 Development of Pedestrian Moving Algorithms

The construction of the pedestrian behavior algorithm will be considered as an example. In general, the process of modeling the movement of people is reduced to updating the state of objects with a certain frequency and changing this state, as well as the number of objects themselves and many other.

The system needs to create new agents to get started (see Fig. 2a). The scene is filled with the necessary number of agents (pedestrians) at the start of the program according to the developed algorithm. Each agent is assigned a route consisting of the starting and ending points of the path, as well as a set of points through which the agent must pass along the road. The agent disappears from the scene and must be replaced by a new agent as soon as it reaches the endpoint of its path.

The process of modeling agent behavior is as follows. The agent is initially assigned a standard scenario: following from the start point to the end with periodic updating of the state of the agent under the influence of external and internal factors. An agent may be assigned a new, non-standard scenario, as a result of updating the state: for example, following another agent, stopping movement or changing a route (see Fig. 2b).

The process of executing standard and non-standard scripts will be discussed later. The standard scenario involves following from the starting point of the route where the agent was created (left the house, from work or from transport) to the destination

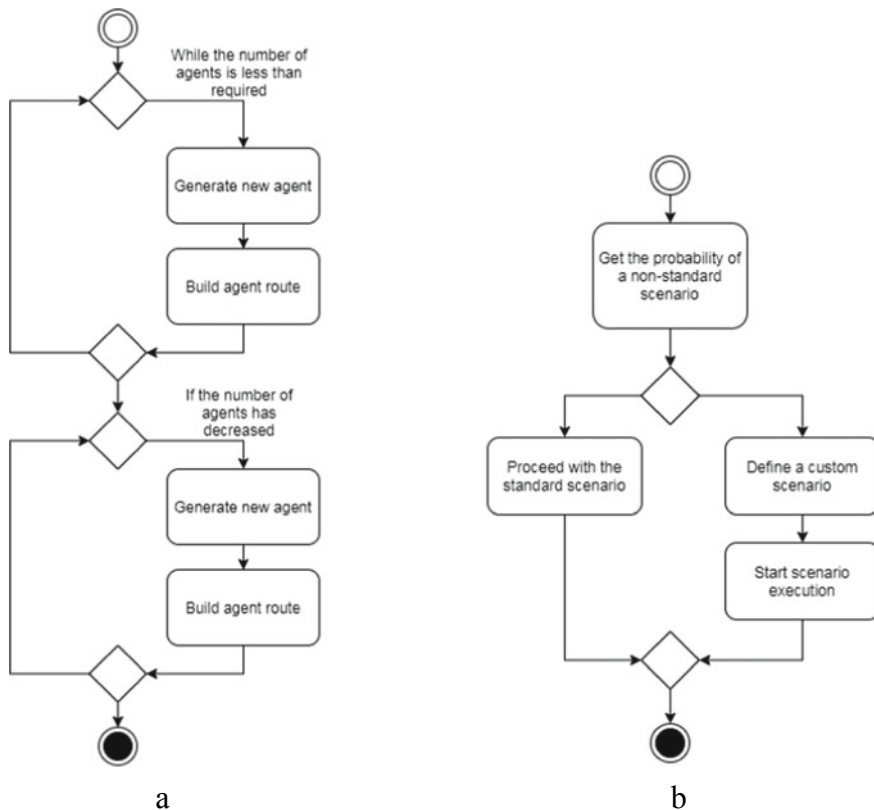


Fig. 2 Algorithms to run simulations: **a** algorithm of filling the scene with agents, **b** scenario assignment algorithm

point (work, home, store), where it will be deleted (see Fig. 3a). The agent path is indicated by dots. The option of waiting one agent for another with subsequent joint movement to a given point is a non-standard scenario (see Fig. 3b).

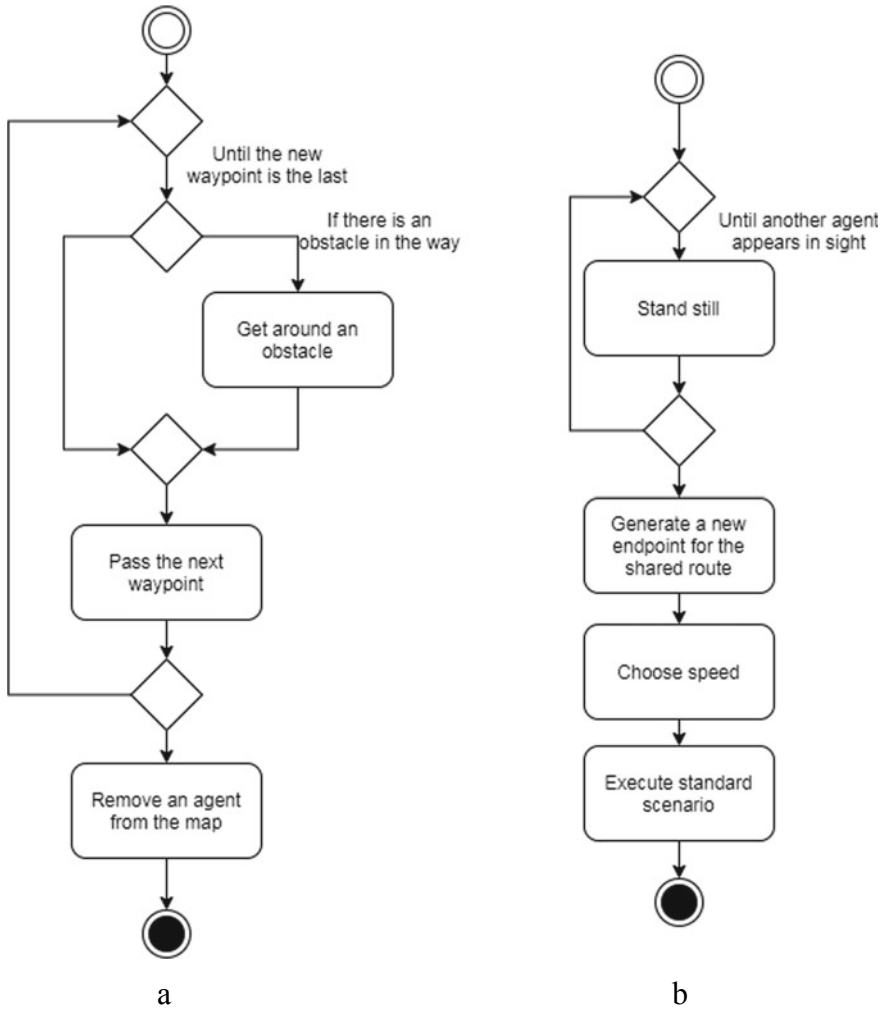


Fig. 3 Pedestrian moving algorithms: a standard scenario, b non-standard scenario

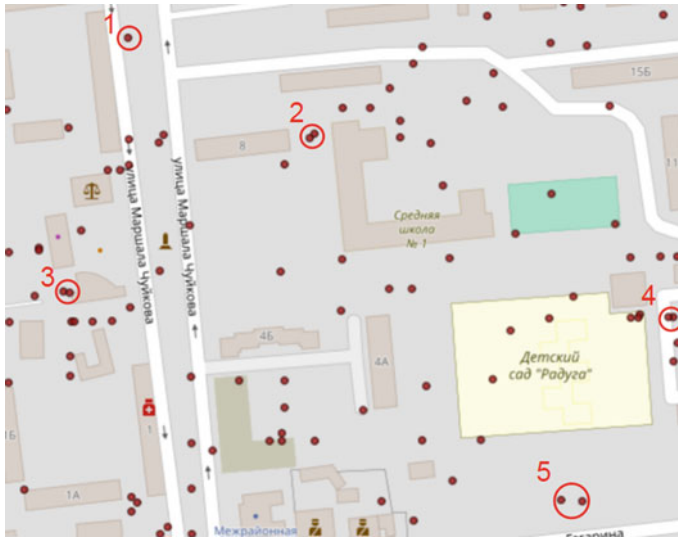


Fig. 4 Visualization of pedestrian behavior

4.3 Pedestrian Behavior Implementation Example

At the moment, several models of behavior of actors in the urban environment have been implemented. Modeling pedestrian behavior will be discussed in more detail below.

Figure 4 shows the result of modeling pedestrian behavior. The main program window is occupied by a map, on top of which individual agents (pedestrians) are shown in circles. As a result, points (simulated objects) appear on the map at the same coordinates where they are located on the server.

The main scenarios of pedestrian behavior are indicated by the numbers in Fig. 4:

1. The standard scenario of a pedestrian following from a start point to an end point.
2. The movement of pedestrians relative to each other if the speed of one pedestrian is higher than the speed of another.
3. The situation of stopping both pedestrians in one place, for example, for conversation.
4. Joint walking of pedestrians in one direction (the speed of this pair is equal to the speed of the slowest pedestrian).
5. The scenario in which one pedestrian waits for another in order to continue moving together.

The running model visualization is always available on the project website <https://live.urbanbasis.com/> [29]. The software package core implements simulation at the multiprocessor computing cluster of VSTU [30].

5 Conclusion

The main methods for modeling agent behavior and tools for its implementation are considered in the study. The main approaches to the construction of scenarios of the behavior of actors are studied: the behavior of individual agents in a multi-agent system is modeled on the example of pedestrians. The developed model demonstrated such advantages as the ability to individually customize each actor, due to which high variability of the behavior of the same type of entities is achieved. The bulkiness and the need for high computing power for operation can be noted among the shortcomings of the resulting model. The tasks for further correction of the model taking into account the current results are identified.

A generalized behavior model of the urban system, as a system describing the interaction of independent agents, system actors, is planned to be developed by complicating the behavior models of actors, as well as increasing the complexity of interactions between elements of the system. This will make it possible to predict changes in processes depending on changes in parameters (properties) of the urban environment. The presence of this forecast will allow to timely influence the current situation by choosing options for changing the situation depending on the set performance indicators. It is especially important that the simulation results can help in decision-making in the event of emergencies that have not previously occurred.

Acknowledgements The reported study was funded by Russian Foundation for Basic Research (RFBR) according to the research project No. 18-37-20066_mol_a_ved. The results of part 3 were obtained within the Russian Science Foundation (RSF) grant (project No. 20-71-10087). The authors express gratitude to colleagues from UCLab involved in the development of Live.UrbanBasis.com project.

References

1. Ustugova, S., Parygin, D., Sadovnikova, N., Finogeev, A., Kizim, A.: Monitoring of social reactions to support decision making on issues of urban territory management. *Procedia Computer Science* **101**, 243–252 (2016)
2. Timm, I.J., Woelk, P.-O., Knirsch, P., Tönshoff, H.-K., Otthein, H.: Flexible mass customisation: managing its information logistics using adaptive co-operative multiagent systems. In: 6th International Symposium on Logistics, Salzburg, Austria, pp. 227–232 (2001)
3. Parygin, D., Nikitsky, N., Kamaev, V., Matokhina, A., Finogeev, A., Finogeev, A.: Multi-agent approach to distributed processing of big sensor data based on fog computing model for the monitoring of the urban infrastructure systems. In: 5th International Conference on System Modeling & Advancement in Research Trends, pp. 305–310. IEEE (2017)
4. Yanishevskaya, A.G., Pesterev, P.V.: The architecture of the multi-agent search. *Dyn. Syst. Mech. Mach.* **6**(2), 94–101 (2018)
5. Anokhin, A., Sadovnikova, N., Kataev, A., Parygin, D.: Modeling of agents behavior to implement gaming artificial intelligence. *Caspian J. Contr. High Technol.* **2**(50), 85–99 (2020)
6. Tsalgatidou, A., Loucopoulos, P.: Rule-based behaviour modelling: specification and validation of information systems dynamics. *Inf. Softw. Technol.* **33**(6), 425–432 (1991)

7. Goal-Oriented Action Planning. <https://alumni.media.mit.edu/~jorkin/goap.html>. Last accessed 2020/04/20.
8. Syahputra, M.F., Arippa, A., Rahmat, R.F., Andayani, U.: Historical theme game using finite state machine for actor behaviour. *J. Phys: Conf. Ser.* **1235**, 012122 (2019)
9. Sekhavat, Y.A.: Behavior trees for computer games. *Int. J. Artif. Intell. Tools* **26**(02), 1730001 (2017)
10. Anokhin, A., Kataev, A.: Finite-automaton model for controlling the behavior of intelligent agents in educational games. *Inf. Technol. Sci. Educ. Manag.* **4**(14), 75–80 (2019)
11. Helbing, D., Molnar, P.: Social Force Model for Pedestrian Dynamics. *Phys. Rev. E* **51**(5), 4282–4286 (1998)
12. Wang, P.: Understanding social-force model in psychological principles of collective behavior. <https://arxiv.org/abs/1605.05146>. Last accessed 2020/05/12
13. Benjamin, P., Erraguntla, M., Delen, D., Mayer, R.: Simulation modeling at multiple levels of abstraction. In: 1998 Winter Simulation Conference, IEEE, Washington, DC, pp. 391–398 (1998)
14. Parygin, D., Usov, A., Burov, S., Sadovnikova, N., Ostroukhov, P., Pyannikova, A.: 2020) Multi-agent approach to modeling the dynamics of urban processes (on the Example of Urban Movements. *Commun. Comput. Inf. Sci.* **1135**, 243–257 (2020)
15. Umnitsyn, M., Nikishova, A., Omelchenko, T., Sadovnikova, N., Parygin, D., Goncharenko, Y.: Simulation of malicious scenarios using multi-agent systems. In: 7th International Conference on System Modeling and Advancement in Research Trends, IEEE, Moradabad, pp. 3–9 (2018)
16. Global MATSim scenario for the 2018 FIFA World Cup in Russia. <https://www.otslab.ru/en/the-global-matsim-scenario-for-the-fifa-world-cup-2018-in-russia>. Last accessed 2020/03/20
17. MATSim scenario for Krasnoyarsk. <https://www.otslab.ru/en/the-matsim-scenario-for-krasnoyarsk>. Last accessed 2020/02/02
18. Borshchev, A., Filippov, A.: From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools. In: 22nd International Conference of the System Dynamics Society, Oxford, England (2004)
19. Warpefelt, H.: The non-player character: exploring the believability of npc presentation and behavior. <https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A912617&dsid=9143>. Last accessed 2020/05/17
20. The Elder Scrolls. <https://elderscrolls.bethesda.net/ru>. Last accessed 2020/05/14
21. Bethesda's Radiant AI as the future of role-playing games, <https://dtf.ru/flood/16882-radiant-ai-ot-bethesda-kak-budushchee-rolevyh-igr>, last accessed 2020/04/22.
22. Grand Theft Auto V. <https://www.rockstargames.com/>. Last accessed 2020/03/25
23. Mafia. <https://mafia-game.com/>. last accessed 2020/03/10
24. Johnson, J., Sarkisian, N., Williamson, J.: Using a micro-level model to generate a macro-level model of productive successful aging. *Gerontologist* **55**(1), 107–119 (2015)
25. Caillou, P., Gaudou, B., Grignard, A., Truong, C.Q., Taillandier, P.: A simple-to-use BDI architecture for agent-based modeling and simulation. *Adv. Intell. Syst. Comput.* **528**, 15–28 (2017)
26. Samigulina, G.A., Samigulina, Z.I.: Cognitive agent development for SMART management systems. *Inf. Technol. Sci. Educ. Manag.* **4**(14), 39–43 (2019)
27. A high-performance, feature-packed library for all your mapping needs. <https://openlayers.org/>. Last accessed 2019/09/16
28. .NetTopologySuite. <https://github.com/nettopologysuite/nettopologysuite>. Last accessed 2019/10/09
29. OsmLifeSimulation. <https://live.urbanbasis.com/>. Last accessed 2020/06/01
30. Multiprocessor computing complex (cluster). <https://evm.vstu.ru/index.php/labs/hpc-lab/abou-ut-hpc>. Last accessed 2020/05/31

Forecast of the Impact of Human Resources on the Effectiveness of the Petrochemical Cyber-Physical Cluster of the Samara Region



Natalya Baykina, Pavel Golovanov, Mikhail Livshits ,
and Elena Tuonosova 

Abstract The chapter analyzes the impact of the training system in the field of cyber-physical production processes for the oil industry of the Samara region. Mathematical models in the form of the production function are Cobb–Douglas linking the efficiency of the petrochemical cluster of indicators of activities supporting the University in the Samara region. On the obtained mathematical models the forecast for oil production and refining volumes is shown, depending on the training of qualified specialists in the university taking into account the regular nature of economic conditions. Based on the DEA (Data envelopment Analysis) methodology, in the period under review, the comparative performance indicators of the oil industry in the Samara region as a cyber-physical production system are evaluated.

Keywords Petrochemical cluster · Oil production · Cyber-physical production system · The cobb–douglas production function · Mathematical model · Data envelopment analysis (DEA) · Comparative efficiency

1 Introduction

A significant share of the budget of the Samara region is provided by the work of the petrochemical cluster, which includes enterprises for the oil extraction and refining, as well as scientific organizations and educational institutions that supply the industry with personnel. The depth increase of oil refining characterizing the effectiveness of the oil industry is demanded strong and powerful cyber-physical industrial complex of scientific and technical potential and various specialists who support the digitalization of the industry. The personnel needs of industry are provided by specialists in many profiles: geologists, operational staff, workers in pipeline transport and oil and chemical processing enterprises, mechanics and electricians, as well as the staff

N. Baykina · P. Golovanov · M. Livshits (✉) · E. Tuonosova
Samara State Technical University, 244 Molodogvardeyskaya, St. Samara 443100, Russia
e-mail: usat@samgtu.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_10

of the necessary infrastructure support: mathematicians, IT specialists, economists, etc.

The preparation of qualified specialists for such a complex, interconnected cyber-physical infrastructure requires a long time—student training at the university with an established educational process continues for at least 4–5 years. This fact does not allow to experiment for effective management for personnel policy of the industry. Assessing the relationship between the indicators of training specialists at a university and the efficiency of oil production is an urgent task, both for the development of a university and the effectiveness of digitalization of the industry as a whole. An effective tool for analyzing this relationship in the context of cyber-physical production is mathematical modeling.

It is necessary to mention that the problem of such mathematical models for crisis periods (e.g. war periods, epidemic, etc.) has not been solved yet. But for regular periods, rather well-known approaches for mathematical modeling is existed [1–5]. The relevance of forecasts of these models does not decrease moreover it is even increased during the crisis period due to several reasons.

The forecast of these models allow evaluating the loss rate not relatively retrospective industry performance but relatively predicted achievable in the absence of crisis phenomena. Besides, chaotic economic processes become regular after the crisis. And this process management requires prediction by mathematical modeling showing long-term fundamental patterns.

2 Mathematical Modeling

The method of constructing mathematical models in the form of an inhomogeneous production function (PF) of Cobb–Douglas is widely used in modern literature; it is used to describe processes and objects of various fields: social, economic, and production. PFs allow one to take into account several influencing factors and evaluate their contribution to the overall effect. For example, in the article [6], a mathematical model was constructed of the relationship and mutual influence of economic growth and innovative technologies of firms operating in monopolistic and oligopolistic markets.

The group of authors in the article [7] built a food industry growth model based on the modified Cobb–Douglas production function, taking into account the innovative factor, and received a typology of Russian regions according to the level of development of light and processing industries.

In the article [8], the dynamics of changes in the volume of production in the mining industry for the period from 2005 to 2015 are considered and a model is constructed in the form of the production function of Cobb Douglas using the example of mining in Russia. The author considered the output volume (billion rubles) as the output characteristic of the model, and the fixed capital (billion rubles) and the number of staff (people) as inputs.

In the article [9], models are presented for assessing the operational efficiency of distribution electric networks using the DEA method and based on the least-squares method with the Cobb–Douglas production function. A comparison is made between the estimation methods and a conclusion is made that the model in the form of a Cobb–Douglas PF is less preferable from the point of view of evaluating efficiency points.

The author in the article [10] gives a mathematical model of GDP based on the Cobb–Douglas PF models. Agriculture, industry, and services. The conclusion is drawn about the influence of the final price of the goods on the prices of factors of production.

In [11], the author considers the issue of the Cobb–Douglas PF and offers a methodology for analysis in accordance with the elasticity of output associated with factors of production.

To describe regular development production systems, the method of constructing mathematical models in the form of an inhomogeneous production function (CF) of Cobb–Douglas (1) [12] is widely used:

$$Y(t) = A \cdot K(t)^\alpha \cdot L(t)^\beta \cdot e^{\lambda t}, \quad (1)$$

where Y is the release of the final product; $K(t)$ —capital resources, $L(t)$ —labor resources, α , β —characteristic of resource use efficiency—elasticity indicator, A —scaling transformation coefficient, λ —factor of influence of scientific and technological progress (STP), α —factor elasticity for factor K , β —for factor L . Cobb–Douglas model supposes to the exponential growth of the product at constant labor and capital costs as a result of scientific and technological progress (STP), the degenerate variant excluding the “no STP” option is easily obtained by and $\lambda = 0$ [12, 13].

On this basis, we will build mathematical models that link the effectiveness of the cyber-physical petrochemical complex with indicators characterizing the effectiveness of training specialists at the Samara Technical University (Samara State Technical University), a reference university in the Samara Region, which trains specialists for the oil industry in all of the above areas. For the output parameters of the model, we take quantitative indicators characterizing the productivity of the oil industry in the Samara Region: oil production— Y_1 , the amount of oil received for refining— Y_2 and the number of processed products: gasoline, diesel fuel, heating oil— Y_3 .

As input characteristics $K(t)$ $L(t)$ formula (1) we will take indicative indicators of the basic higher education institution of the Samara region: graduation of students of Samara State Technical University— S_i , people (characterizes the performance of the university); the total number of scientific publications— P_i , pcs.; performance of scientific and technical works (R&D) on grants— G_i , units; and the generation of intellectual property— I_i , units (characterize the level of practical significance of scientific work). Given these input factors, the model (1) of the inhomogeneous Cobb–Douglas PF with allowance for the NTP is written in the form (2).

$$Y(t) = A \cdot S(t)^\chi \cdot P(t)^\kappa \cdot G(t)^\phi \cdot I(t)^\rho \cdot e^{\mu t}, \quad (2)$$

Table 1 Factor elasticity

Factor elasticity	Input resource
χ	For the factor S_i , graduation of students of Samara State Technical University
κ	For the factor P_i , total number of scientific publications
φ	For the factor G_i , grant research
ρ	For the factor I_i , intellectual property generation
μ	For the influence factor of scientific and technological progress (NTP)

The identification of model parameters is carried out using the least-squares method (LSM) [13, 14]. We will evaluate the quality of modeling by the determination coefficient (R^2) and the criterion of F-statistics, and the predicted properties of the model by the Darbin-Watson criterion (DW). We will carry out the smoothing of the initial data based on the moving average method, providing averaging of the effect of random outliers of statistical information [15].

The sensitivity of model solutions (2) to the corresponding input resources is characterized by sensitivity coefficients, factorial elasticity, which is presented in Table 1.

We develop mathematical models in the form of an inhomogeneous Cobb–Douglas PF taking into account scientific and technological progress (2).

In these models, as an output parameter Y , we will consider the performance indicators or the oil industry that are most dependent on human resources.

oil production, $Y = Y_1$

$$Y_1(t) = A_1 \cdot S(t)^{\chi_1} \cdot P(t)^{\kappa_1} \cdot G(t)^{\phi_1} \cdot I(t)^{\rho_1} \cdot e^{\mu_1 t} \quad (3)$$

by the amount of oil received for processing, $Y = Y_2$

$$Y_2(t) = A_2 \cdot S(t)^{\chi_2} \cdot P(t)^{\kappa_2} \cdot G(t)^{\phi_2} \cdot I(t)^{\rho_2} \cdot e^{\mu_2 t} \quad (4)$$

by the number of processed products: gasoline, diesel fuel, and heating oil, $Y = Y_3$

$$Y_3(t) = A_3 \cdot S(t)^{\chi_3} \cdot P(t)^{\kappa_3} \cdot G(t)^{\phi_3} \cdot I(t)^{\rho_3} \cdot e^{\mu_3 t} \quad (5)$$

3 The Result of Modeling

For identification, we will use statistical data on the oil complex of the Samara region [16–18]. Table 2 shows the coefficients of mathematical models (3), (4), (5) obtained with the help of method LSM separately for smoothed and unstated statistical information.

Table 2 Characteristics of mathematical models (3–5)

	$Y_1(t)$		$Y_2(t)$		$Y_3(t)$	
	Oil production		Amount of oil received for processing		Amount of processed products	
Elasticity	Unstated	Smoothed	Unstated	Smoothed	Unstated	Smoothed
χ	0.0245	0.8098	0.1537	0.4567	0.1325	1.0605
κ	0.0494	0.0650	0.0155	0.0323	0.0196	0.1046
ϕ	0.0109	0.0808	-0.0236	-0.0239	-0.0222	-0.0489
ρ	0.0836	-0.4503	0.0380	-0.1531	0.0545	0.5506
μ	0.01341	0.06705	0.00106	0.02139	-0.01480	0.04674
<i>Simulation quality</i>						
DW	1.1461	2.0982	1.8553	3.0115	1.5339	2.9943
Ra	0.4270	-0.0491	0.0723	-0.5058	0.2331	-0.4971
R^2	0.8494	0.9698	0.9358	0.9893	0.7822	0.9769
F	5.6395	32.0996	14.5714	92.2261	3.5904	42.3796

Figures 1 and 2 show the results of modeling for oil production from the unstated and smoothed initial data of the inhomogeneous Cobb–Douglas PF taking into account the scientific and technical progress (3) from 2008 to 2018.

Factor elasticities are the components of the logarithm gradient of the corresponding industry indicator (3)–(5) in the region of the input coordinates $S(t)$, $P(t)$, $G(t)$ and $I(t)$ that is why it allows estimating the sensitivity of this parameter. The greatest factor elasticity model (3), constructed based on non-smoothed initial statistical data has the generation of intellectual property objects

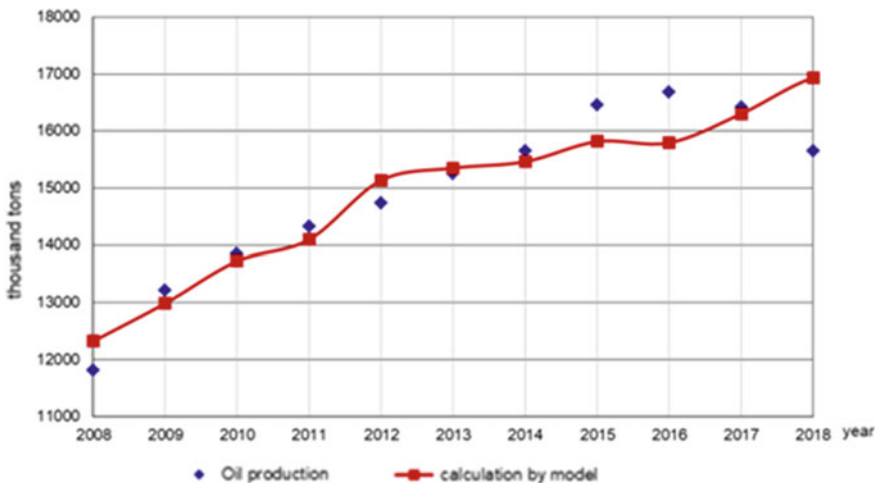


Fig. 1 Calculation of oil production (Y_1) by the model (3)

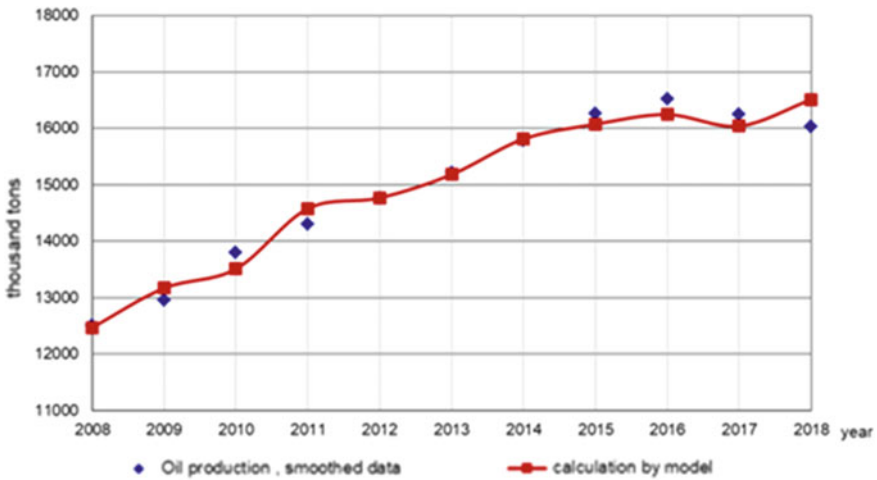


Fig. 2 Calculation of oil production (Y_1) by the model (3)—smoothed data

I_i . When smoothing data, the greatest influence for oil production Y_1 is made by the implementation of research on grants (factor G). When smoothing data, the quality of this model improves—the coefficient of determination R^2 increases by 14% to almost one. This indicates a trend of a significant influence of the scientific work of university graduates on the effectiveness of their work in the oil industry.

Similarly, the modeling results and model parameters for the amount of oil received for refining (Y_2) are presented in Figs. 3 and 4.

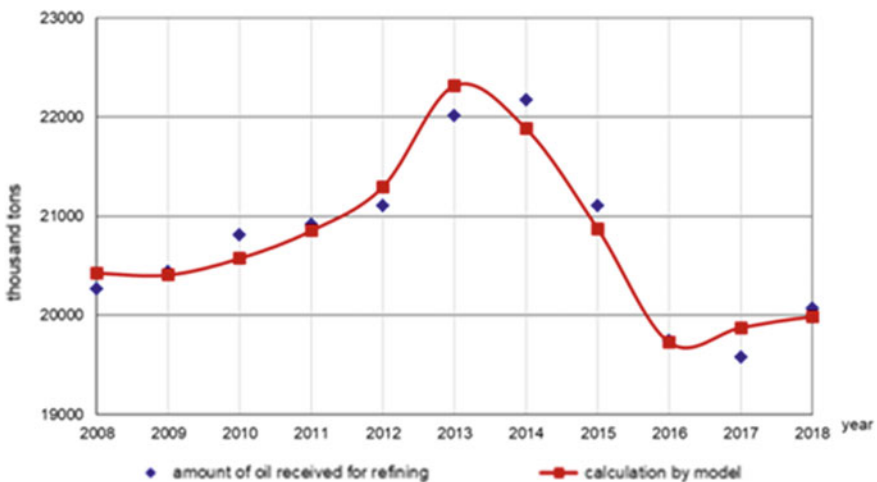


Fig. 3 Calculation amount of oil received for refining (Y_2) by the model (4)

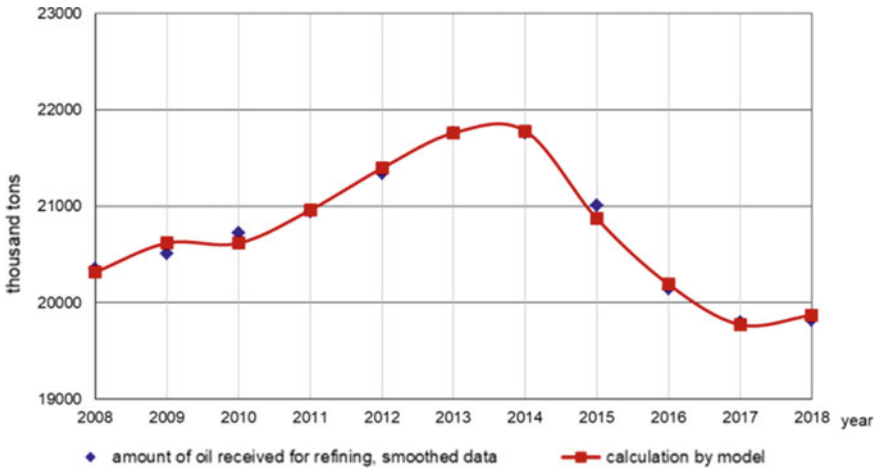


Fig. 4 Calculation amount of oil received for refining (Y_2) by the model (4)—smoothed data

The greatest factor elasticity in the model (4), taking into account scientific and technological progress, has the generation of intellectual property objects I_i , which reflects its greatest influence on the amount of oil received for refining, and when smoothing the initial data, the total number of scientific publications has the greatest impact P_i . The models are distinguished by their good approximative and prognostic properties. The value R^2 is more than 0.9, and the Darbin–Watson criterion DW in the model constructed from the smoothed data increases by 62% compared to the model constructed from the unsmoothed data and reaches a value of 3.

The result for the production of petroleum products in the Samara region (5) based on uncoated and smoothed source data is presented in Figs. 5 and 6.

When modeling the number of oil products produced, the factor generation, as in the case of the model (4), has the highest factors of the generation of intellectual property objects I_i and performing research on grants G_i . The processing of gasoline, diesel fuel, and heating oil is a complex technological process, where the scientific base of young industry experts is very important, which is largely characterized by joint research developments of teachers and university students. When smoothing the initial data, the modeling error decreases and the determination coefficient (R^2) increases by 25%.

All the obtained models are characterized by good convergence with statistical data, but at the same time, the factor elasticity χ of the number of graduates of SSTU students S_i is either the smallest or negative. This can be explained by the inefficiency of the extensive approach to training for such a high-tech and knowledge-intensive industry as the oil industry.

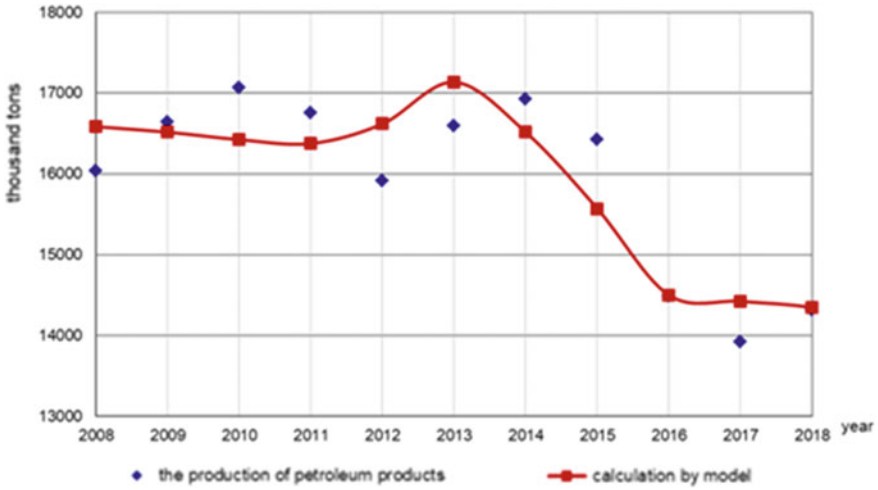


Fig. 5 Calculation processed products (Y_3) by the model (5)

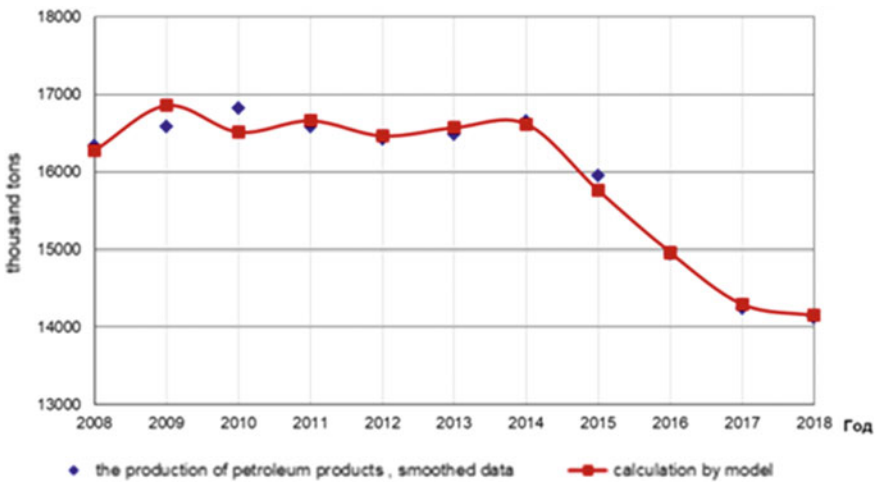


Fig. 6 Calculation processed products (Y_3) by the model (5)—smoothed data

4 Forecasting

We check the predicted properties of the Cobb–Douglas PF models taking into account the scientific and technical progress for 2008–2018 and construct forecast until 2022. Provided regular crisis-free development the student graduation S_i for the next 4 years is easily predicted. We extrapolate the data on the generation of intellectual property objects I_i by a quadratic polynomial, research on grants G_i by

Table 3 Forecast of university resources for 2019–2022

Year	Input data			
	S , people	P , pieces	G , units	I , units
2019	5272	4899	50	52
2020	4731	5081	51	60
2021	4235	4764	52	71
2022	3856	4987	52	87

a cubic polynomial, and the total number of scientific publications P_i in the forecast will be assumed to correspond to the roadmap of the reference university [19, 20] (Table 3).

The forecast will be built on models based on unstated source data. Figure 7 shows the forecast values obtained by the model (3) of oil production (Y_1), Fig. 8 - the amount of oil received (Y_2) for refining (4), and Fig. 9—the production of petroleum products (Y_3) (5).

From the forecast (Fig. 7) according to the initial data (graduation of specialists and scientific work at Samara State Technical University), it can be seen that in the case of regular crisis-free development oil production would be increased during the study period.

According to the model (4), the forecast for the amount of oil received for refining, presented in Fig. 8, shows that in 2019 there will be no changes in oil refining, and in the future, an increase would be expected.

Regarding the forecast of the amount of produced gasoline, diesel fuel, and heating oil (Fig. 9), in 2019 there is a slight decrease in the volume of processed products,

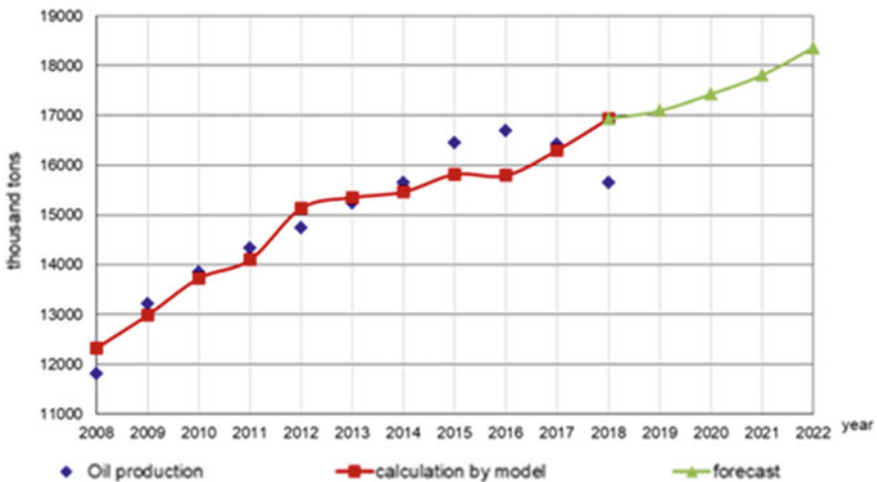


Fig. 7 Forecast of oil production (Y_1) by the model (3) until 2022

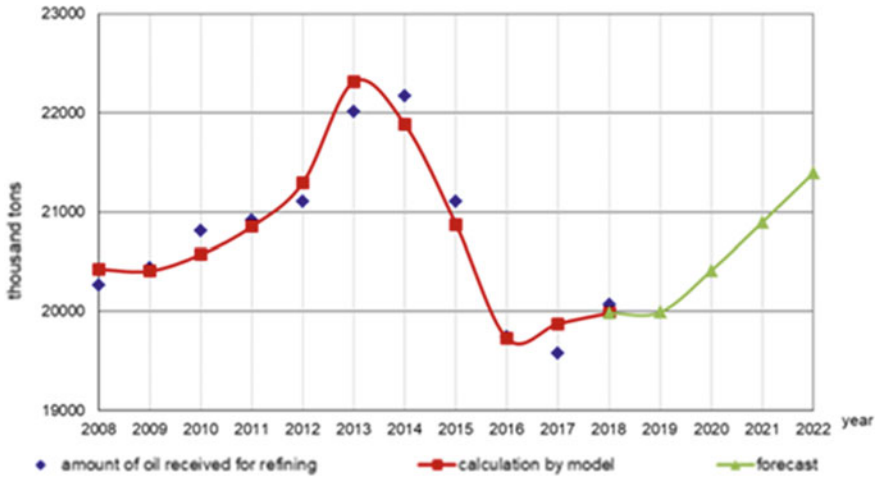


Fig. 8 Forecast of the amount of oil received for refining (Y_2) by the model (4) until 2022

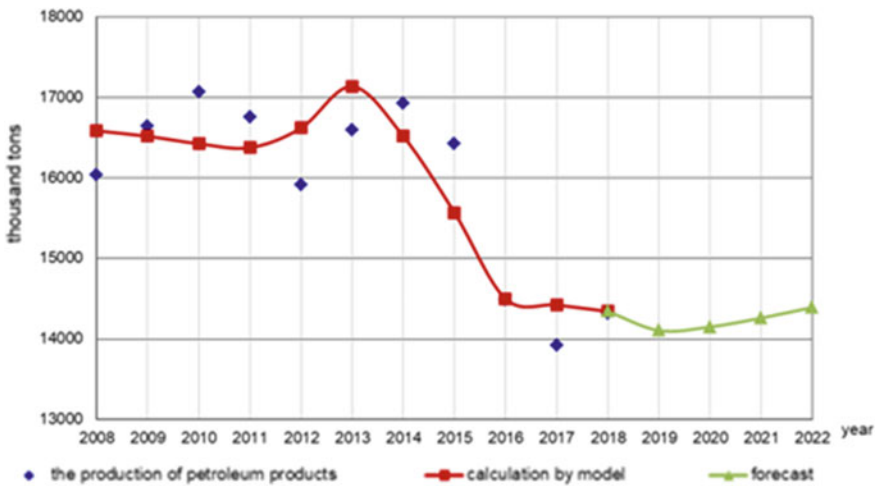


Fig. 9 Forecast of processed products (Y_3) by the model (5) until 2022

and then in case of regular crisis-free development, there would be a slight increase, which in 2022 would ensure a return to the level of 2018. It should be noted that the personnel influence on the operation of the oil complex is very important but not absolute for the high performance of output production indicators. For example, as practice shows, if the state of financial support, equipment, and arrangement of fisheries are predictable, then external factors make significant adjustments, and they can not always be taken into account.

5 Comparative Assessment of the Effectiveness

For a comparative assessment of the effectiveness in the considered period time of staffing the oil cluster of the Samara region for all three indicators Y_1, Y_2, Y_3 , we use the methodology of multi-criteria evaluation of comparative efficiency—Data Envelopment Analysis (DEA) [21].

The structure of the comprehensive performance indicator DEA will form as follows:

$$f_i = \max \frac{u_1 \cdot Y_1 + u_2 \cdot Y_2 + u_3 \cdot Y_3}{v_1 \cdot S_i + v_2 \cdot P_i + v_3 \cdot G_i + v_4 \cdot I_i}, \tag{6}$$

where u_1, u_2, u_3 are the positive weighting coefficients characterizing the relative contribution of each of the output factors Y_i to the total efficiency coefficient f , and accordingly, v_1, v_2, v_3, v_4 the weighting coefficients of the input quantities to be determined during the DEA procedure.

Figure 10 shows the comparative performance indicators of the oil industry in the Samara region.

In 2008 and 2013, the comparative effectiveness of staffing in the industry was maximum and equal to 1. From 2013 to 2018, there has been a decline in production efficiency and a decrease in relative efficiency to a minimum level— $f_{\min} = 0.54$. Then the growth is possible in the absence of crisis to 0,71 by 2022. The results presented in Fig. 10 confirm the conclusions of the simulation by formulas (3)–(5), shown in Figs. 3–9.

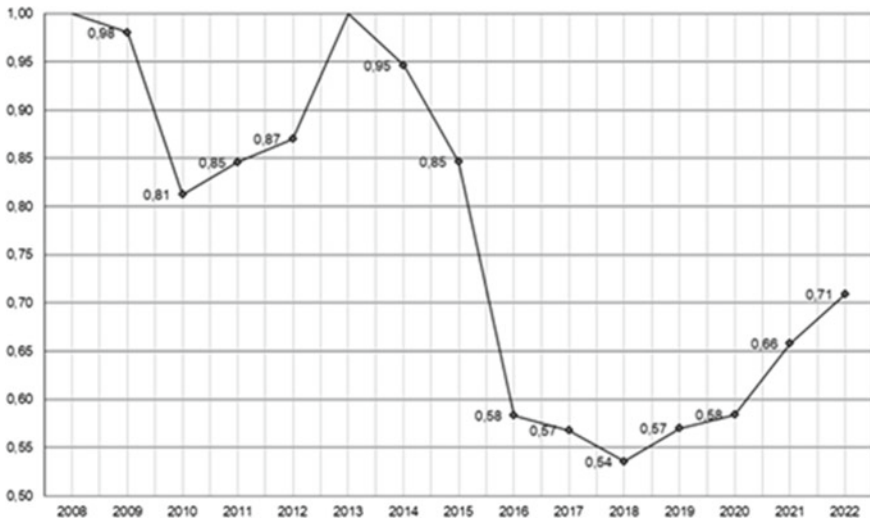


Fig. 10 Indicators of the comparative efficiency of the oil industry from 2008 to 2022

6 Conclusion

Mathematical modeling, forecasting, and comparative assessment of the effectiveness allow making the following conclusion: the practical scientific work skills together with the University scientists acquired by young specialists during their study provide high efficiency of staff replenishment contribute most to the development of the oil industry.

The obtained models allow us to forecast the absence of crisis production development and outline measures for the development of the industry, university, and forms of their interaction. Confirmation of the forecast for 2019 should be expected with the publication of statistical reports in 2020.

It should be noted that the chapter analyzes the influence of exclusively the training and graduation of specialists of SSTU on the effectiveness of the regional oil cluster and establishes a significant effect on the production of this factor. In this case, one should take into account the influence of numerous other factors, in particular, the equipment of deposits, the economic, social, and political conditions, etc.

It is necessary to highlight that maintaining the high relevance of forecasting on mathematical models in the conditions of the absence of crisis development in crisis periods. These forecasts determine long-term prospects; reflect the fundamental tendencies which are very important while industry staffing. In the long run, staffing will determine the productivity, competitiveness, and financial performance of the oil cluster.

Acknowledgements The reported study was funded by the Russian Foundation for Basic Research (RFBR), according to research project No. 20-08-00240.

References

1. Klopchenko, V.: Elements of theory of prediction in education, No.4 (93), pp. 235–241. Published by: Technological university of the Federal state educational institution of higher professional education “Southern Federal University” in Taganrog (2009)
2. Golovanov, P., Livshits, M., Tuonosova, E.: Mathematical model of issuing specialists by university. *Mathem Methods Tech Tech MMTT J* **2**, 114–119. Yu.A. Gagarin Saratov State Technical University, Saratov (2018).
3. Bondyreva, I.: Interaction of universities and enterprises in training personnel for innovative development. *Theor. Econ. Mag.* **3**, 19–25 (2011)
4. Evelev, A., Livshits, M., Tuonosova, E., Frank, E., Tsapenko, M.: Mathematical modeling process of training specialists in the engineering industry at a technical university. In: IOP CONFERENCE SERIES: materials science and engineering. «13th international conference on mechanical engineering, automation and control systems, meacs 2018». Novosibirsk, 12–14 December’s 2018 r. Institute of Physics Publishing, p. 012026. (2019)
5. Golovanov, P., Livshits, M., Tuonosova, E.: Analysis of impact made by the flagship university on the efficiency of petrochemical complex. In: *Studies in systems, decision and control*, vol. 260, pp. 289–300. Springer International Publishing (2020)

6. Kislickij M.M., Chumachev A.A., Gan E.P.: Evaluation of the company's performance based on the analysis of the relationship and mutual influence of economic growth and innovative technologies. In: *Agri-food policy of Russia*, № 8 (20), pp. 43–49. Ural research Institute of economic and food security (2013) (In Russian)
7. Glinskij, V.V., Serga, L.K., Samotoj, Simonov, N.V., Yu, E.: Development of the food and processing industry as a condition for improving Russia's national security. *Vestnik NSUEM*, № 3, 221–234. Publishing house of the Siberian branch of the Russian Academy of Sciences, Novosibirsk (2017) (In Russian)
8. Titkova, I.K.: Analysis of the current state of mining in Russia using the Cobb-Douglas production function. In: *Collection of student scientific papers based on the materials of the XV all-Russian student scientific conference with international participation*. Edited by L. A. Dremova. pp. 372–374 (2017)
9. Marcelo Azevedo Costa: Ana Lúcia Miranda Lopes, Giordano Bruno Braz de Pinho Matos: Statistical evaluation of Data Envelopment Analysis versus COLS Cobb-Douglas benchmarking models for the 2011 Brazilian tariff revision. *Socio-Econ. Plann. Sci.* **49**, 47–60 (2015). <https://doi.org/10.1016/j.seps.2014.11.001>
10. Muro, K.: A note on the three-sector Cobb-Douglas GDP function. *Econ. Model.* **31**, 18–21 (2013). <https://doi.org/10.1016/j.econmod.2012.11.008>
11. Pavelescu, F.M.: Methodological considerations regarding the estimated returns to scale in case of Cobb-Douglas production function. *Procedia Economics and Finance.* **8**, 535–542 (2014). [https://doi.org/10.1016/S2212-5671\(14\)00125-7](https://doi.org/10.1016/S2212-5671(14)00125-7)
12. Klejner, G.: *Production functions: theory, methods, application*, 239 p. Finance and statistics, Moscow (1986)
13. Diligensky, N., Tsapenko, M., Gavrilova, A.: *Mathematical models of management of production and economic systems: textbook allowance*, 112 p. Samar. state tech. un-t—Samara (2005)
14. Gavrilova, A., Golovanov, P., Diligensky, N., Tuonosova, E.: Designing models and building a short-term perspective forecast of graduation. *Bull. Samara State Tech. Univ. Ser. Eng. Sci.* **2**, 21–29 (2011)
15. Ayvazyan, S., Mkhitarian, V.: *Applied statistics and fundamentals of econometrics*, 1022 p. M: Unity (1998)
16. Samara Annual Statistical Report, 155 p. Stat.sb\Samarastat (2017) (In Russian)
17. Samara Annual Statistical Report. Enterprises and organizations. https://samarastat.gks.ru/wps/wcm/connect/rosstat_ts/samarastat/ru/statistics/organizations/ (2018) (In Russian)
18. Federal State Statistics Service Regions of Russia. The main characteristics of the constituent entities of the Russian Federation—2018 Volga Federal District. Samara Region. https://www.gks.ru/bgd/regl/b18_14s/Main.htm
19. Samara State Technical University. Normative documents. Roadmap of the development program. <https://su.samgtu.ru/files>
20. Samara State Technical University. Mission of the reference university. <https://su.samgtu.ru>
21. Diligensky, N., Gavrilova, A., Tsapenko, M.: Construction and identification of mathematical models of production systems: Textbook, 126 p. Samara: LLC “Etching” (2005)

Mathematical Model of Integration of Cyber-Physical Systems for Solving Problems of Increasing the Competitiveness of the Regions of the Russian Federation



Alexander Bolshakov , Irina Veshneva , and Dmitry Lushin

Abstract A mathematical approach to the creation of an integrated mathematical model of the socio-economic system for the study of various aspects of the development of the competitiveness of the region on the basis of the principles of building cyber-physical systems is proposed. For this, a modification of the Kolmogorov-Chapman equations has been developed, which allows one to describe the integration of enterprises into clusters and use fragmentary data on the state of indicators of the structures under study. A graph of cause-and-effect relationships of dangerous combinations of events is created. The results of model approbation are carried out. An example of numerical modeling is given, showing a high probability of timely response to the occurrence of elementary events included in the minimum sections of the graph. The results of the work are intended for use in the development of mathematical models of advising systems for monitoring and countering the violation of the sustainable functioning of an enterprise using modern mathematical models and information technologies.

Keywords Industry 4.0 · Cyber-physical systems · Competitiveness · Regions of the russian federation · Causal graph · Minimal sections · Kolmogorov-chapman equations

A. Bolshakov (✉)

Peter the Great St.Petersburg Polytechnic University, 29, Polytechnicheskaya, St. Petersburg 195251, Russia

e-mail: aabolshakov57@gmail.com

I. Veshneva · D. Lushin

Saratov National Research University Named After N.G. Chernyshevsky, 83, Astrakhanskaya Street, Saratov 410012, Russia

e-mail: veshnevaiv@gmail.ru

D. Lushin

e-mail: dman95csait@gmail.com

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_11

1 Introduction

The modern period of socio-economic development corresponds to a series of various successive crises. At the same time, one crisis entails others, such as, for example, the problems caused by the coronavirus pandemic, which initiate many economic, social, psychological, and other interrelated problems interacting in explicit causal products and in hidden, latently affecting the environment. Often it is latent connections that have a decisive influence on the dynamics of the system as a whole. Therefore, the task of identifying latent causal relationships is urgent [1–3].

The problem is complicated by the formation of new industry 4.0 [4–6] and a change in the technological order. The processes occurring in the world are of a systemic nature and testify to the uneven, cyclical nature of economic development and the need to detect emerging basic technologies, the use of some innovations in order to relatively quickly overcome alternating crises corresponding to the minimum of the great Kondratyev wave. One of the most noticeable trends in the change in the technological structure is the digitalization of the economy and all spheres of society [7–9]. The modern information technology concept of cyber-physical systems [10–13] implies the integration of computing resources into real objects of any kind, including industrial complexes, socio-economic systems, and biological objects. In cyber-physical systems, information, and communication technologies (ICT) are distributed throughout the system and synergistically linked with its constituent elements [14].

According to the researchers, the prospects for building cyber-physical systems, as well as the formation of Industry 4.0 on their basis, affect the interests of society. Therefore, the creation of such systems should be investigated in technical, as well as in broader social, cultural, and economic aspects, including taking into account the increase in the competitiveness of various regions [15]. The use of cyber-physical systems extends to many types of human activity, including different complexes: defense, industrial, transport, commercial, and energy, as well as various types of life support systems from medicine to smart homes and cities, as well as many economic systems. At the same time, when creating these systems, it is necessary to take into account and identify cause-and-effect relationships, which is especially important when building competitive regions of the Russian Federation on the principles of cyber-physical systems. At the same time, it is believed that the creation of full-fledged cyber-physical systems in the future will lead to changes in interaction, including the economic sphere, similar to the construction and implementation of the World Wide Web in everyday life. In the current conditions of revolutionary changes, the study of the processes of increasing the competitiveness of regions is one of the key tasks of forecasting and management. One of the obvious ways of research is the development of mathematical models based on nonlinear differential equations.

2 Mathematical Modeling Based on the Kolmogorov-Chapman Equations

One of the approved methods for studying failures in technical systems, the tree of failures, and the system of Kolmogorov-Chapman equations has been successfully used to model processes in socio-economic systems [16, 17]. In the latter case, the special development of the cause-and-effect graph is carried out like a tree of events. Logical Boolean operations are introduced into the vertices of the graph, which are consequences of causes. Then the possible minimum cross-sections are revealed as sets of sequentially realized events. An event graph is built for each of the minimum sections. For each graph of events, a system of Kolmogorov-Chapman equations is constructed and solved. For each of these sets of events that form the minimum cross-sections, the dynamics of the probabilities of the combined events are obtained. The rest of the events are then excluded by the model. For them, other minimal sections, state graphs, systems of equations, and solutions are constructed. As a result, sets of separate possible solutions are considered for joint consideration of the system.

The difficulty in applying this model is, firstly, the need to declare clear dependencies of cause-and-effect relationships, which is very difficult and often incorrect for socio-economic systems. Secondly, a significant increase in the number of equations. For 4 reasons, $2^4 + 1 = 32$ equations are obtained, for 100 reasons— 2^{101} equations, which is quite difficult for calculations based on a mathematical model. Thus, when building a model for the integration of cyber-physical systems, it is necessary to solve the following tasks:

- (1) take into account the intersection of the studied processes;
- (2) create a system of nested models to reduce the number of equations in the system;
- (3) to allow the structure to develop naturally, taking into account the effects of causes that are absent in the field of study.

This problem of developing a mathematical model of a complex cyber-physical system can be solved only on the basis of nonlinear dynamics methods. In accordance with the tasks set, we will present the model of the economy of the region of the Russian Federation, as a certain socio-economic environment that allows providing the infrastructure of interaction between clusters of various organizations in the region. The initiative to increase the competitiveness of the region is based on the leadership policy that unites the leading organizations and allows the development of the infrastructure of the socio-economic environment of the region based on information and communication technologies (ICT), linking all the constituent elements (Fig. 1).

To create a mathematical model of such a structure, we will be based on a relatively simple model of the Kolmogorov-Chapman equations. Let the models of individual organizational structures be based on a system of balanced indicators [18] for which a cause-and-effect graph is built that combines the 4 branches of risk indicators of the threat of loss of the organization's ability to work: finance, market, internal

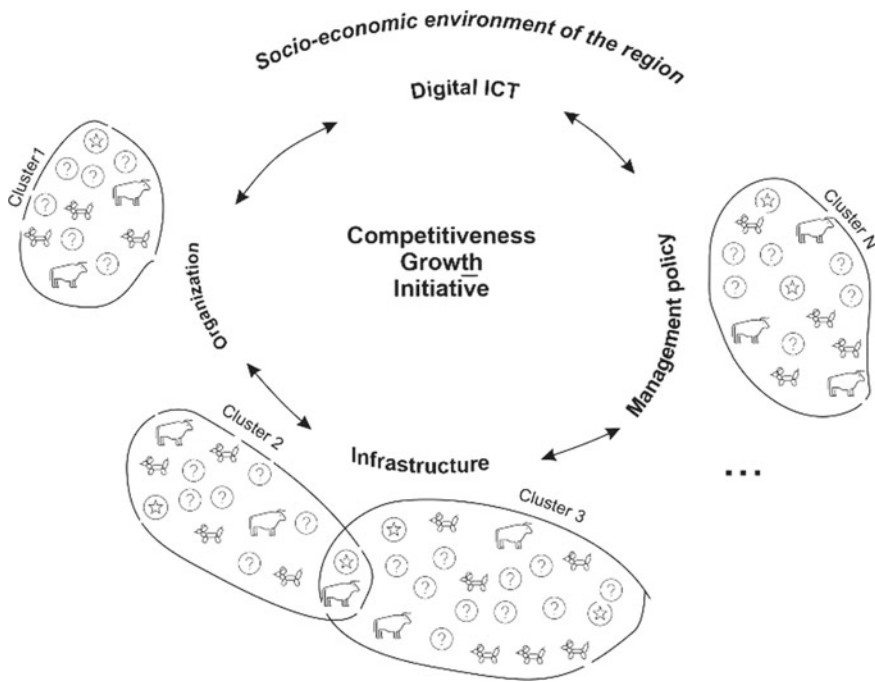


Fig. 1 An integrative model for describing the growth opportunities of the region's competitiveness

business processes, training, and growth. In the processes of the organization's functioning, risks arise from the impossibility of implementing the main processes, for example, industrial business processes, as a result, the risk of systemic disruption of the enterprise's activities and the implementation of the root top of the risk tree increases.

To implement the goal of monitoring the identification of dangerous combinations of risks for the successful functioning of an enterprise, it is advisable to use the system of linear differential equations of Kolmogorov-Chapman:

$$\begin{aligned} \frac{dP_0(t)}{dt} &= - \sum_{j=1}^k \lambda_j P_0(t) + \sum_{j=1}^k \mu_j P_j(t); \\ \frac{dP_i(t)}{dt} &= -P_i(t)\pi_i^- + \sum_{j=0}^{2^k-1} \pi_{ij}^+ P_j(t); \\ \frac{dP_{2^k-1}(t)}{dt} &= \sum_{j=1}^k \lambda_j P_{2^k-k+j-2}(t). \end{aligned} \tag{1}$$

This mathematical model is applicable to determine the probabilities of realizing the minimum cross-sections under certain conditions; it consists of 2^k equations for the functions $P_0(t), \dots, P_{2^k-1}(t)$ —the probabilities of events corresponding to the vertices of the cause-and-effect graph of the risk tree. The disadvantages of using such a model to characterize various socio-economic processes are the above-noted lack of taking into account the cross-section and mutual influence of risk events, a large number of 2^k equations taking into account the possibility of implementing k reasons, and the absence of a visible prospect of increasing the complexity of the modeled part of the system.

3 Modeling Statistics

Let us ask ourselves a question: what functions describe statistics? Consider the statistical data for the Saratov region of the Russian Federation. Let’s take the data on the factor “Volume of innovative goods, works, services”, which contain, in percentage terms, information on the total volume of goods that are shipped, as well as services and works [17] (Table 1).

Let’s approximate these data and get the following function:

$$\begin{aligned} \text{Volume InnProd} = & 0.102866 + 1.23901e^{-x^2} - 0.0582672\text{Cos}(x) \\ & + 0.0331014\text{Cos}(x^2) - 0.0377882\text{Sin}(x) \\ & + 0.0362681\text{Cos}(10x)\text{Sin}(x). \end{aligned} \tag{2}$$

The results of approximation (solid line) and statistical data (points) are shown in Fig. 2. If you use this pattern, you can build the following extrapolation curve (Fig. 3). The obtained dependence allows us to assume: on the one hand, the value of the Volume InnProd factor, which has fallen to close to zero, will remain fluctuating near zero for a very long time. This may be so, if we recall, for example, the last few decades of the development of the South African Republic. However, on the other

Table 1 Characteristics of innovative products (Volume InnProd) by volume, services, and works as a percentage of the total volume of goods shipped, as well as services and works in the Saratov region

Year	Volume InnProd	Year	Volume InnProd	Year	Volume InnProd
2000	0.49	2006	0.080	2012	0.103
2001	0.06	2007	0.036	2013	0.111
2002	0.20	2008	0.175	2014	0.054
2003	0.17	2009	0.328	2015	0.140
2004	0.189	2010	0.211	2016	0.000
2005	0.15	2011	0.001	2017	0.000

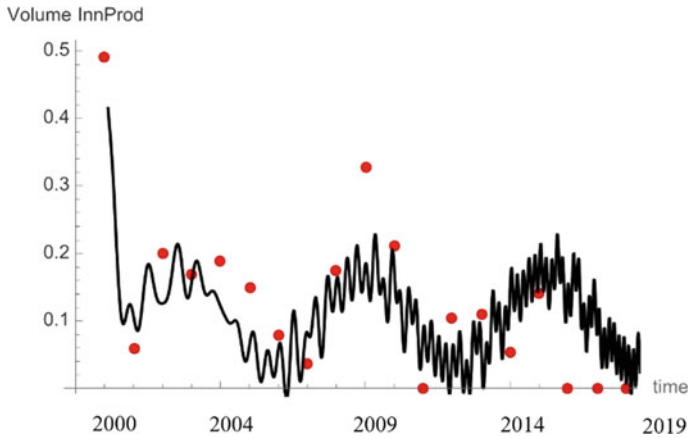


Fig. 2 Characteristics of the value of innovative goods, as well as services and works as a percentage of the total volume in the Saratov region for the period 2000–2018 according to [17], marked with red dots. The extrapolation results are shown by the solid line

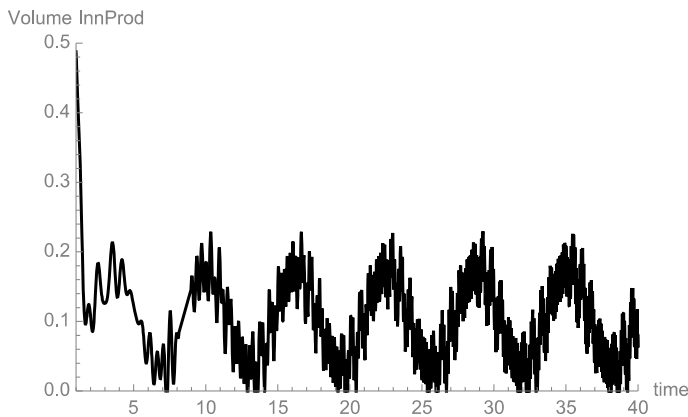


Fig. 3 Results of extrapolation of statistical data on the value of innovative services, goods, works as a percentage of the total volume in the Saratov region for 40 years, starting from 2000

hand, we can safely say that the period over which the extrapolation was carried out is too short. In addition, the Volume InnProd factor is dependent on many other components. It is not isolated and should be considered as a result of interaction with other factors of the region's development. Also, it should be taken into account that the region is an open system and interacts with the environment, consisting of other regions and the governing influences of the government.

It is important to note that the observed process is oscillatory. Let's carry out the transformation and look at the expression obtained as a result of statistical data interpolation. The resulting expression is complex-valued:

$$\begin{aligned}
\text{Volume InnProd} = & 0.102866 + 1.23901e^{-x^2} - 0.0582672\text{Cos}(x) \\
& - (0.0291336 - 0.0188941i)e^{ix} \\
& - 0.00906701ie^{-i9x} + 0.00906701ie^{-i9x} \\
& + 0.00906701ie^{-i11x} - 0.00906701ie^{i11x} \\
& + 1.23901e^{-x^2} + 0.0165507e^{-ix^2} \\
& + 0.0165507e^{ix^2}
\end{aligned} \tag{3}$$

Let us assume that it is possible to build a certain model that describes the dynamics of statistically observed factors of the region's competitiveness. Let this model be based on the well-proven model of the Kolmogorov-Chapman equations.

4 Modification of the Kolmogorov-Chapman Equations

Let's modify the mathematical model for predicting combinations of risks for an organization. First, consider the replacement of probabilities in the nodes of the event tree with status functions (SF) [18]. Here SF characterizes the status, which is a set of characteristics that are inherent to the subject or object and determine its position in the system. In this case, an SF will be called a mapping that establishes a rule for determining the correspondence of some ordered pair of arguments [18]. An ordered pair of arguments is a characteristic of the probability of an event occurring at a graph node and its direction along an adjacent edge, which can be interpreted as an action or reaction. Let's introduce the binary values of the characteristics at the nodes. We will estimate the probability of an event occurring at a node as low and high, and the direction to the root top (up) or away from it (down). In this case, there will be 4 possible estimates of the state in the node of the event tree: {low, down}; {low, up}; {high, down}; {high, up}. The use of SF as characteristics of the top of the cause-and-effect graph allows us to combine λ_j —the values of the intensities of the occurrence of an event π_j and μ_j —the intensity of counteraction to this event, which is used in the system of Kolmogorov-Chapman equations. Intensities describing the "direction of action" λ_j , π_j , μ_j , and the probabilities P_i of these events in complex-valued status functions. In this case, the real characteristic of the function describes the classical probabilities of occurrence of events at the vertices of the causal graph, the imaginary characteristic corresponds to the angle of rotation of the vector and allows you to represent the "direction of action".

In addition, such a representation allows one to take into account interference effects. If two causes A and B are in states described by the attributed functions f_A and f_B and if these states interact, then the result of the interaction is described as.

$$P = |f_A - f_B|^2. \tag{4}$$

Such a mathematical description reflects the features of various interdependent processes, as well as the internal structure of various real objects that are subject to research. These include, for example, the fundamentals and elements of the firm's balanced scorecard. These include financial performance, customer relationships, internal business processes, and personnel training and development indicators [19]. Moreover, they are internally interconnected. A common description of these states allows one to obtain a probability distribution of the following form:

$$P = |f_A|^2 + |f_B|^2 \quad (5)$$

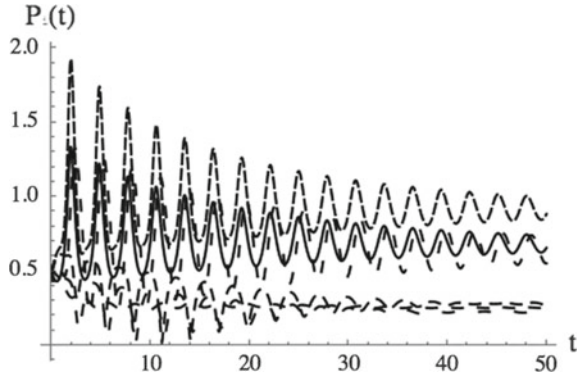
The methods used for estimation to obtain estimates introduce errors in these measurable states. This, in turn, influences the principle of fact-based governance. The use of the so-called complex functions for assessment will eliminate this discrepancy.

Note that SFs are attributed to the state of the object, similar to membership functions (FP) in the fuzzy set theory (TNM). They also depend on the base variable entered, such as r . However, unlike FP, the choice of which is determined by the conditions of the problem and the simplicity of presentation, the set of used SFs is formed as a set of orthonormal functions that form a set of orthogonal basis functions. The introduction of SF changes the interpretation of the calculation results. The probability is not limited by the supremum $FP = 1$, has a similar restriction on the integral $SF = 1$. To implement this condition, we introduce not a normalization equation as in the system of Kolmogorov-Chapman equations, but a restriction on the growth of the function into the equation. This limitation is introduced by analogy with the limitations of innovation diffusion models [20]. Then in the equations for P_0 the 2nd term of the equation appears under the integral— P_0 (2). The third term of the equation under the integral allows us to introduce the mixing of the SF, the implementation of which is described in the equation with the rest of the reasons for the set of reasons under study. Such a mathematical representation of the cross interaction of the processes under study is characteristic of synergistic models in the equations of chemical and biochemical kinetics. It is also useful in assessing the health of components in a production system that uses digital twin technology [21, 22].

Consider a model with nonlinear growth and mixing constraints introduced into the Kolmogorov-Chapman equations for the probabilities of realizing events introduced into the vertices of the event graph. The first term remains in the equations, the growth restriction is introduced in the second, and mixing is carried out in the third, which occurs nonlinearly. We obtain equations of the form (5):

$$\frac{dP_0(t)}{dt} = \int_{-1}^1 \left(\sum_{j=1}^k P_0(t) (\psi_j(r) - P_0(t) - \psi_j^*(r) P_j(t)) \right) dr$$

Fig. 4 Description of the results of digital modeling based on the modification of the Kolmogorov–Chapman equations. Nonlinearity is introduced for various initial conditions used: 1— $P(0) = 0.5$, 2— $P(0) = 1$. In this case, the values of the probabilities of other different events are taken equal to 0.5



$$\frac{dP_i(t)}{dt} = \int_{-1}^1 \left(\sum_{j=0}^{2^k-1} P_i(t) (\psi_i^* - \psi_j P_j(t) - \psi_j P_i(t)) \right) dr$$

$$\frac{dP_{2^k-1}(t)}{dt} = \int_{-1}^1 \left(\sum_{j=1}^k P_{2^k-k+j-2}(t) (\psi_j - \psi_{2^k-k+j-2} P_{2^k-k+j-2}(t)) \right) dr \quad (6)$$

where $P_i(t)$ are complex-valued SFs.

Numerical simulation of the system of Eq. (4) makes it possible to obtain a certain oscillatory process tending to an equilibrium stable state (see Fig. 4). The results obtained suggest that studies based on the proposed equations obtained on the basis of the modification of the Kolmogorov–Chapman equations for SF can be promising as a basis for describing cyber-physical systems that are a sub-system of the general socio-economic system of the region. This raises the complexity of the obtained results of digital modeling of interpretation for social applications. Therefore, when investigating a dangerous confluence of events, it is required to compare models for the same behavior of the system in a certain sense. It is desirable to form proactive management decisions based on forecasting [23].

5 Conclusions

Thus, a mathematical approach to the creation of an integrative model of the socio-economic system by studying the opportunities and risks of developing the competitiveness of the region is proposed. The model is based on a modification of the Kolmogorov–Chapman equations and implies the integration of individual enterprises into clusters with the possibility of using disparate fragmented data on the state of indicators of the structures under study and sections of the graph of cause-and-effect relationships of dangerous combinations of events. The preliminary approximation of the model has been carried out. An example of numerical calculations

is presented, showing a high probability of timely response to the occurrence of elementary events included in the minimum sections of the graph. Their occurrence and development lead to the realization of the event of the graph vertex. The results of the work are intended to be used in the development of mathematical models of advising systems for monitoring and countering the violation of the sustainable functioning of an enterprise using modern mathematical models and information technologies.

Acknowledgements This research was partially supported by the Russian Fund of Basic Research (grant No. 20-010-00465).

References

1. Anderson, P., Tushman, M.L.: Technological discontinuities and dominant designs: a cyclical model of technological change. *Adm. Sci. Q.* **35**(4), 604–633 (1990)
2. Cherry, E.: Veganism as a cultural movement: a relational approach. *Soc. Movement Stud.* **5**(2), 155–170 (2006)
3. Das, T.K., Teng, B.S.: Risk types and inter-firm alliance structures. *J. Manage. Stud.* **33**(6), 827–843 (1996)
4. Akaev, A.A., Sadovnichij, V.A.: Mathematical modeling of global, regional and national dynamics, taking into account the impact of cyclical fluctuations [Akaev, A.A., Sa-dovnichij, V.A.: Mathematical modeling of global, regional and national dynamics taking into account the impact of cyclical fluctuations] Modeling and forecasting global, regional and national development [Modeling and forecasting of global, regional and national development]. Moskva, LIBROKOM [Moscow, LIBROKOM], 5–67 (2011)
5. Demsetz, H.: Industry structure, market rivalry, and public policy. *J. Law Econ.* **16**(1), 1–9 (1973)
6. Uzzi, B.: Social structure and competition in interfirm networks: the paradox of embeddedness. *Adm. Sci. Q.* **42**(1), 35–67 (1997)
7. Kergrouch, S.: Industriya 4.0: new challenges and opportunities for the labor market. *Forsajt [Foresight]* **11**(4), 6–8 (2017)
8. Goodman, D., DuPuis, E.M.: Knowing food and growing food: beyond the production–consumption debate in the sociology of agriculture. *Sociologia ruralis* **42**(1), 5–22 (2002)
9. Kaplan, R.S., Mikes, A.: Managing risks: a new framework. *Harvard Business Review*. June. (2012)
10. Lee, E.A.: The past, present and future of cyber-physical systems: a focus on models, <https://citeweb.info/20150013436>. Last accessed 2018/04/12
11. Cyber-Physical Systems (CPS) <https://www.nsf.gov/pubs/2018/nsf18538/nsf18538.htm>. Last accessed 2019/04/14
12. Lee, E.: Cyber physical systems: design challenges. University of California, Berkeley Technical Report No. UCB/EECS-2008–8. <https://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-8>, last accessed 2019/05/16
13. Rawat, D.B., Rodrigues, P.C.: Stojmenovic I. Cyber physical systems: from theory to practice. CRC Press (2015)
14. Ronzhin, A.L., Basov, O.O., Sokolov, B.V., YUsupov, R.M.: Conceptual and formal model for the synthesis of cyber-physical systems and intelligent spaces. *News of higher educational institutions. Instr. Mak.* **59**(11), 897–905 (2016)
15. Meshalkin, V.P., Katerishchuk, M.Yu., Vasilenko, E.A.: Methodology for the formation of a comprehensive assessment of the effectiveness of reengineering of business processes at an

- industrial enterprise. *News of higher educational institutions. series: economics, finance and production management* **2**(20), 87–92 (2014)
16. Bolshakov A.A., Veshneva I.V., Kushnikov V.A. Mathematical modeling of the dynamics of the behavior of social groups to prevent dangerous combinations of events in critical periods of state development. *Actual problems of applied mathematics, informatics and mechanics: collection of articles. Proceedings of Int. scientific and technical. conf. – Voronezh: Publishing House “Scientific Research Publications”* (2017). ISBN 978-5-9500319-1-5
 17. Regions of Russia. Socio-economic indicators. https://gks.ru/bgd/regl/b19_14p/Main.htm. Last accessed 2020/04/24
 18. Veshneva, I.V., Chistyakova, T.B., Bolshakov, A.A.: The status functions method for processing and interpretation of the measurement data of interactions in the educational environment. *SPIIRAS Proc* **6**(49), 144–166 (2016). <https://doi.org/10.15622/sp.49.8.ISSN:20789181>
 19. Kaplan, R.S.: *Conceptual foundations of the balanced scorecard*. Harvard University, Harvard Business School (2010)
 20. Rogers, E.: *Diffusion of innovations*, 4th edn. Free Press, New York (1995)
 21. Buldakova, T.I., Suyatinov, S.I.: Assessment of the state of production system components for digital twins technology. Springer Nature Switzerland AG 2020. In: Kravets et al. A. G. (eds.) *Cyber-Physical Systems: Advances, in Design & Modelling, Studies in Systems, Decision and Control*, vol. 259, pp. 253–262 (2020). https://doi.org/10.1007/978-3-030-32579-4_20
 22. Suyatinov, S.I.: Conceptual approach to building a digital twin of the production system Springer Nature Switzerland AG 2020. In: Kravets et al. A. G. (eds.) *Cyber-Physical Systems: Advances, in Design & Modelling, Studies in Systems, Decision and Control*, vol. 259, pp. 279–290 (2020). https://doi.org/10.1007/978-3-030-32579-4_22
 23. Shcherbakov, M.V., Glotov, A.V. Cheremisinov, S.V.: Proactive and predictive maintenance of cyber-physical systems. Springer Nature Switzerland AG 2020. In: Kravets et al. A. G. (eds.) *Cyber-Physical Systems: Advances, in Design & Modelling, Studies in Systems, Decision and Control*, vol. 259, pp. 263–278 (2020). https://doi.org/10.1007/978-3-030-32579-4_21

Regional Competitiveness Research Based on Digital Models Using Kolmogorov-Chapman Equations



Irina Veshneva , Galina Chernyshova , and Alexander Bolshakov 

Abstract The concept of regional competitiveness is investigated taking into account the trends in the transition to Industry 4.0. The risk indicators analysis for the competitiveness increase of Russian regions was carried out. A systematization of indicators based on a cause-effect graph was proposed. The developed system of indicators allows us to identify competitive risks at the regional level. A simplified digital model of the cause-effect graph structure for competitiveness risks was presented. It is suggested to find a solution to the problem of high dimension by revealing the structures of minimal cut set in a causal graph. In this case, it is assumed that the reasons that trigger the chain of event-consequences realization which leads to the risk of a root event of the cause-effect graph should be identified. The minimal cut sets for the obtained graph of competitiveness factors that correspond to critical combinations of events and lead to a decrease in the region's competitiveness were determined. The digital model was validated using the indicators values for the regions of the Volga Federal District as an example.

Keywords Digital model · Industry 4.0 · Cyber-physical systems · Competitiveness · Regions of the Russian federation · Cause-effect graph · Minimal cut set · Kolmogorov-Chapman equations

I. Veshneva (✉) · G. Chernyshova
Saratov National Research University named after N.G. Chernyshevsky, 83, Astrakhanskaya
Street, Saratov 410012, Russia
e-mail: veshnevaiv@gmail.ru

G. Chernyshova
e-mail: cherny111@mail.ru

A. Bolshakov
Peter the Great St. Petersburg Polytechnic University, 29, Polytechnicheskaya, St. Petersburg
195251, Russia
e-mail: aabolshakov57@gmail.com

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced
Human-Centered Society*, Studies in Systems, Decision and Control 333,
https://doi.org/10.1007/978-3-030-63563-3_12

1 Introduction

In the context of the transition to Industry 4.0, the competitiveness of regions or other territorial items is an urgent issue concerning creating and studying the risks of adverse events, and their consequences. The concept of regional competitiveness has different ambiguous definitions. In many cases, the proposed interpretations correspond to the requirements of the specific research. The competitiveness of the region is a reflection of the complex interaction of various factors (economic relations, productive forces, institutional environment), and from such interaction, a synergistic effect appears.

Traditionally, for quantitative evaluation of competitiveness at the country level is applied to the gross domestic product (GDP). For a comprehensive assessment of regional development, Russian statistical sources use an indicator such as the gross regional product. Competitiveness can be interpreted in terms of efficiency, which is expressed through GDP, and industrial indicators of the region. However, the GDP as an indicator of wealth and economic development has been criticized [1].

A large number of competitiveness assessments apply a rating approach. One of the main theories, which is devoted to the study of regional competitiveness and its relationship with the characteristics of global competitiveness, was proposed by M. Porter. His concept is based on the method of determining the so-called global competitiveness indicator. It can be used to control the competitiveness of a particular region [2].

Application of the principles of cyber-physical systems creating is essential for obtaining modern models for assessing socio-economic objects. The estimations obtained using the Global Competitiveness Index (GCI) 4.0 reflect the ability of countries to compete in the era of the Fourth Industrial Revolution [3]. Indicators considered as prerequisites for long-term growth are united in a hierarchical structure containing four categories: the creation of favorable institutional, infrastructural conditions; human capital; innovation ecosystem; development of commodity markets, financial system.

The European Regional Competitiveness Index (RCI) provides a generalized submission of competitiveness for each of the territorial nomenclature units. Although RCI includes a wide set of indicators, it has some limitations, for example, it does not take into account any of the environmental aspects [4].

The Global Sustainable Competitiveness Index (GSCI) involves an extensive range of indicators, taking into account factors of sustainable development and future potential (environment, society, economy) for integrated competitiveness assessing [5].

Other proposed approach includes a large number of factors which affect regional system performance. These indicators are divided into three groups: social infrastructure and political institutions, monetary and fiscal policies, microeconomic competitiveness. This structure covers more than 120 indicators [6, 7].

A modern approach to competitiveness estimation is associated with innovations and new production technologies, including cyber-physical systems. They lead to the

growth of industrial production without increasing damage to the environment. In the framework of the World Economic Forum methodology, competitiveness estimation is formulated in terms of environmental sustainability. The efficient use of natural resources; health improvement; biodiversity for innovation are taken into account. As a rule, Green economies work better in terms of the labor force health (reduction of environmental pollution), innovation, and restoration of natural capital [8].

Significant attention is devoted to studies of the economy cluster composition. It is concluded that increasing competitiveness should have a cluster character. This point of view differs from the previous provisions on the targeted nature of industrial production. It means the economy gradually moves to related activities that rely on the corresponding fundamental comparative advantages [9, 10].

An essential aspect of the conceptual approach to regional competitiveness is the increase in labor productivity and innovative activity [11]. A distinctive feature of regional competitiveness models should ensure the inclusion of endogenous growth patterns. Key factors requiring targeted investments are human capital and knowledge. They will determine the differences in growth rates according to researchers [12].

It is especially important to give due consideration to the effectiveness of using the principles of cyber-physical systems constructing for managing socio-economic processes. To analyze the scientific, technical and innovative potential of the region in the process of assessing competitiveness, especially the interaction of innovations and socio-economic development of territories, it is necessary to take into account the specifics of assessing competitiveness in various conditions, especially the interaction of innovation and socio-economic development of territories [13].

The debatability of the methodology for assessing regional competitiveness has led to the emergence of a variety of ratings comparing the Russian regions by different sets of indicators. However, competitiveness is not a characteristic that can be measured directly; assessments are based on an indirect set of indicators. Econometric models have significant limitations in the process of application (formulation of a hypothesis about the nature of data distribution, sampling of the required volume, quantitative indicators, unstructured time series). Attempts to use artificial intelligence methods and models to expand the possibilities of numerical assessment of complex socio-economic systems in the face of uncertainty, and inaccuracy of information are limited to simpler objects of study than the competitiveness of regions. The proposed approach will allow the development of constructive digital mathematical models and methods for the dynamic assessment of the risks of competitiveness of the Russian regions on the principles of building cyber-physical systems.

2 Formation of a Risk Factors Hierarchical Structure for Regional Competitiveness Assessment

Digital modeling of regional competitiveness involves specifying the concept under consideration. In the context of the regional economy, there are different approaches to competitiveness. As an important criterion for this category, the ability of regional authorities to create conditions for achieving and maintaining a competitive advantage in certain areas is highlighted [14].

The experience of foreign studies cannot be fully applied to the Russian regions due to the specifics of the Russian economy. The development of the domestic regional economy is largely determined by differences in the conditions of reproduction: the regions are provided with natural resources to varying degrees; differ in climatic, social, economic, investment, and innovative conditions [15]. For a systematic assessment concerning identifying the risks of regional development, competitiveness is determined by the ability to use regional potential (resource, raw materials, labor, innovation, production) to achieve high living standards.

Risk is an integral part of the functioning of socio-economic systems at any level. Therefore, the risks have a significant impact on the governance of competitiveness, require consideration in the formation of managerial decisions. Existing approaches to the definition of risk concerning socio-economic systems focus on the uncertainty and probabilistic nature of the assessment. In particular, this refers to the probability of obtaining an unpredictable result as a consequence of a management impact. It should be noted that the risk is a real or perceived event or activity that could potentially cause uncertainty, harm, or destruction of the economic, natural, or social system [16].

The risks of regional competitiveness are understood as the risk of negative trends that arise in the process of socio-economic development as a result of an unsatisfactory degree of taking into account various types of dangers. These dangers threaten regional socio-economic development if the corresponding managerial influences were not appropriated during creating the concept of perspective directions for regional development and strategic management.

An important research stage is the formation of an indicator set to assess the regional competitiveness risks. The methods used for the quantitative estimation of regional competitiveness are diverse in the indicator set. They generally contain such categories as macroeconomic, infrastructural, demographic, technological, and innovative indicators. For a systematic assessment of competitiveness risks, it is necessary to include in the model a sufficiently large number of external (economic, political, organizational) and internal (financial and non-financial) factors that create threats at the regional level. Studies that do not take into account various aspects of competitiveness and their interaction do not provide the necessary tools to support decision-making at the territorial level [17]. Therefore, the application of ideas and methods for constructing cyber-physical systems for a given subject area is relevant.

To understand the phenomenon of competitiveness, various factors of competitiveness risks are highlighted by the proposed classification. They are presented as

a hierarchical structure. The root level of the tree corresponds to the integral assessment of regional competitiveness. At the next level of the hierarchy, two categories of indicators are distinguished: transactional and transformational factors. The applied method of event tree formation allows distinguishing semantically connected groups of indicators [18].

Technical, social, and natural resource risk indicators are denoted as transformational indicators at the next level of the event tree. The category of transactional indicators includes institutional, informational, and innovative indicators of competitiveness risks. An expanded set of indicators is presented for a comprehensive description of the regional competitiveness risk system at the next level of the hierarchical structure. As the leaves in the tree-like structure, it is proposed to use an extended set of characteristics: the value of all fixed assets; the value of fixed assets depreciation; the density of railway tracks; density of paved roads; the percentage of the population with money income below the subsistence minimum; average per capita money income of population; life expectancy at birth; migration growth coefficient; dilapidated housing stock; the number of medical personnel (doctors) per 10,000 population; the number of unemployed registered at the bodies of employment services; the number of recorded crimes; the volume of shipped own-produced goods, works performed and services (mining and quarrying); sowing areas of main crops; emission of pollutants into atmosphere; discharge of polluted sewage; use of the Internet electronic documentary exchange in organizations; R&D personnel; sales of innovative goods and services. This set of indicators sufficiently describes the level of regional development and allows assess the competitiveness risks. Besides, statistical data is available, and corresponding data has been collected for a fairly long time.

3 Building an Event Tree to Prevent Competitiveness Risks at the Regional Level

The proposed method to the analysis of risks associated with a decrease in the competitiveness of regions is to study the factors of competitiveness risks based on the formed causal graph. A special feature of the method is the development of a cross-functional tree graph to describe a complex multi-factor object, which is the competitiveness of the region. The use of cause-effect graphs has been well tested in the analysis of technical systems. The extension of these methods has been demonstrated for the analysis of socio-economic systems and cyber-physical systems. It can be argued that the systematization of competitiveness indicators for Russian regions based on the structures of cause-effect graphs is promising.

To modeling the structure of the cause-effect graph for systematization of competitiveness risk indicators, the following hierarchical risk indicators system of for assessment is proposed: regional competitiveness (E_0); transformational (E_1); transactional (E_2); technical (E_3); social (E_4); natural resource (E_5); institutional (E_6);

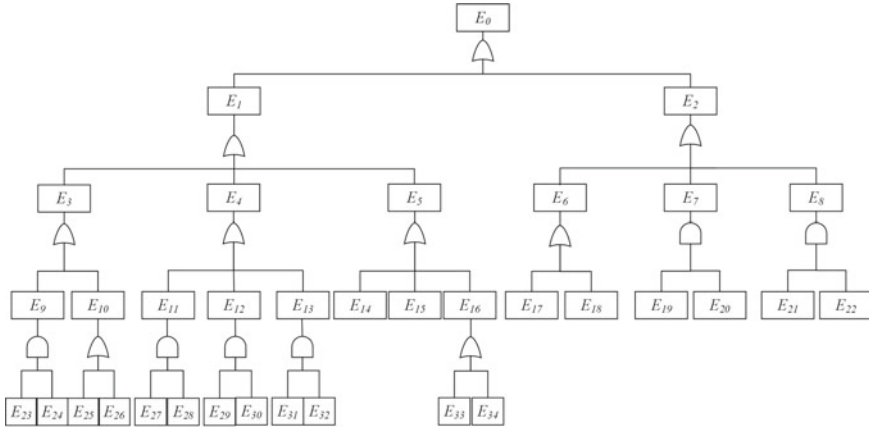


Fig. 1 The hierarchical structure of the region competitiveness risk indicators

information (E_7); innovative (E_8); using of fixed assets (E_9); transport infrastructure (E_{10}); population income (E_{11}); demographic (E_{12}); quality of life (E_{13}); mineral extraction (E_{14}); sowing areas of main crops (E_{15}); environmental (E_{16}); number of unemployed (E_{17}); number of crimes (E_{18}); using of the Internet (E_{19}); using the Internet electronic documentary exchange in organizations (E_{20}); R & D personnel (E_{21}); sales of innovative goods and services (E_{22}); all fixed assets (E_{23}); fixed assets depreciation (E_{24}); density of railway tracks (E_{25}); density of paved roads (E_{26}); average per capita money income of population (E_{27}); the percentage of the population with money income below the subsistence minimum (E_{28}); life expectancy at birth (E_{29}); migration growth coefficient (E_{30}); dilapidated housing stock (E_{31}); medical personnel (doctors) per 10,000 population (E_{32}); emission of pollutants into atmosphere (E_{33}); discharge of polluted sewage (E_{34}). Figure 1 shows a causal graph for assessing regional competitiveness risks.

Digital modeling based on the obtained structure allows identifying a key direction for management decision making that impact on the socio-economic system of the region to increase its competitiveness.

4 Kolmogorov-Chapman Model for Assessing Competitiveness Risks

The following algorithm, taking into account the values of the vector of managerial influences on the set of adverse events $u(t) \in U$ is used to assess the competitiveness of a particular region. For permissible values of the vector of estimation of the states of the system $x(t) \in X$ the criterion is minimized at a given time interval $[t_s, t_f]$:

$$K_A = \int_{t_s}^{t_f} Pr(t, x, u)dt, \quad (1)$$

где K_A is the criterion which characterizes the risks of reducing the region's competitiveness in the time interval $[t_s, t_f]$; Pr is the probability of the corresponding risk being realized and the occurrence of a certain process stopping due to a critical combination of events under consideration, while they do not separately cause the specified risk; X, U are a set of values for system state vectors $x(t)$ and management impacts on the combination of adverse events $u(t)$; the variable t indicates time.

The solution to this task poses certain difficulties because the construction of a sufficiently complex digital mathematical model for the active system under research is required for a full procedure for determining the values of the criterion $K_A(t, x, u)$ as well as the integrand $Pr(t, x, u)$. In the process of minimization, it is necessary to solve a complex system of differential equations of relatively high dimension. Thus, when solving the formulated problem, it is required to determine the control action $u(t) \in U$ which transfers the criterion value K_A into the area of the minimum risk probability for the vertices of the constructed cause-effect graph. In this regard, the basis of the algorithm contains a statement repeatedly confirmed by practice that to reduce the risks of the root event at the top of the graph, it is enough to develop and implement a detailed comprehensive plan of action for minimizing the selected combinations of risks.

As a result, task (1) is reduced to the search for such actions to achieve effective functioning of the region's competitiveness management system, in which the probability of occurrence and development of risks to reduce the region's competitiveness is minimized.

To assess the probability of the occurrence of a possible set of events that are atomic in the proposed representation, it is necessary to single out a finite number of states. The state graph for the complete cause-effect tree of events (Fig. 1) has $2^{35} = 34\,359\,738\,368$ vertices corresponding to the combinations of events. Solving a system of more than 34 million equations is not feasible. It is possible to consider mathematical models of separated fragments of the tree structure. These parts are formed by the minimal cut sets structure for the cause-effect graph.

To assess the probability of reducing competitiveness, the minimal cut sets [19] of the graph of the main factors affecting competitiveness are determined. Elements of these minimal cut sets correspond to groups of processes, the implementation of risks of non-fulfillment of which leads to the loss of competitiveness by the region. Classification is performed by the number of elements of minimal cut sets. There are two-, three-element, etc. minimal cut sets. They correspond to combinations of events that are critical and lead to a decrease in the competitiveness of the Russian region.

A state graph is formed for each such section, which is used to construct a system of differential equations. Such a system is called Kolmogorov-Chapman [20, 21]. As a result of solving the system of differential equations the probability of reducing

the region's competitiveness due to critical combinations of events associated with individual risk factors is determined.

Let us consider one variant of the minimal cut sets $E_{17}—E_{19}—E_{20}—E_{21}—E_{22}$. This minimal cut set is 5-element; it includes innovative, institutional, and informational risk factors for regional competitiveness. Possible causes of adverse events or the implementation of risks are presented in Fig. 2. The implementation of each risk combination is characterized by the probability Pr_i . Analysis of the state graph of the type shown in Fig. 2 can be performed by digital mathematical modeling methods.

For each of the states, it is necessary to compose a differential equation:

$$\frac{dPr_i(t)}{dt} = \sum_{j=0}^{31} l_j Pr_j(t) - \sum_{k=0}^{31} d_k Pr_k(t) \quad (2)$$

The constraints are formed based on the Kolmogorov-Chapman differential equations system and in matrix form have the following form:

$$\frac{dPr(t)}{dt} = A \cdot \begin{pmatrix} Pr_0(t) \\ Pr_1(t) \\ \dots \\ Pr_{32}(t) \end{pmatrix} \quad (3)$$

where $Pr_i(t)$ is the probabilities of the transition of the managed object into certain state i , A is the transition matrix containing parameters of the intensity of actions to restore d_i and actions to implement the risk of the corresponding vertex l_i in the state graph. Matrix A is constructed based on the adjacency matrix of the states graph (Fig. 2).

Thus, it is possible to compose the differential equations system describing the effect on the root vertex for the cause-effect graph. The system of differential equations in the case of the 5-element minimal cut sets takes the form.

$$\begin{aligned} \frac{dPr_0(t)}{dt} &= (-l_1 - l_2 - l_3 - l_4 - l_5)Pr_0(t) + d_1Pr_1(t) + d_2Pr_2(t) \\ &+ d_3Pr_3(t) + d_4Pr_4(t) + d_5Pr_5(t) \end{aligned} \quad (4)$$

$$\begin{aligned} \frac{dPr_1(t)}{dt} &= l_1Pr_0(t) + (-d_1 - d_2 - d_3 - d_4 - d_5)Pr_1(t) + d_2Pr_6(t) \\ &+ d_3Pr_7(t) + d_4Pr_8(t) + d_5Pr_9(t) \end{aligned} \quad (5)$$

$$\begin{aligned} \frac{dPr_2(t)}{dt} &= l_2Pr_0(t) + d_3Pr_{10}(t) + d_4Pr_{11}(t) + d_5Pr_{12}(t) \\ &+ (-d_2 - l_1 - l_3 - l_4 - l_5)Pr_2(t) + d_1Pr_6(t) \end{aligned} \quad (6)$$

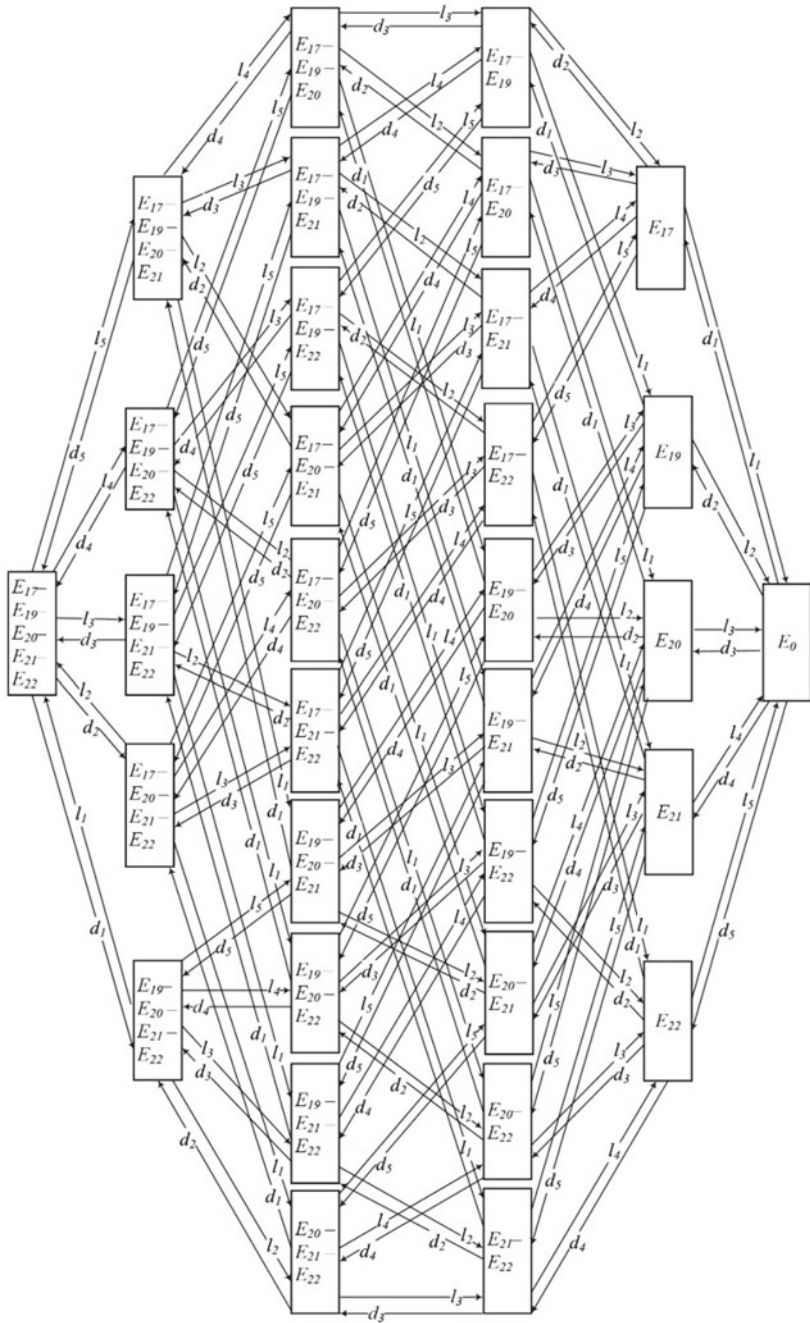


Fig. 2 State graph for a 5-element minimal cut set $E_{17}-E_{19}-E_{20}-E_{21}-E_{22}$

...

$$\begin{aligned} \frac{dPr_{31}(t)}{dt} = & l_5 Pr_{26}(t) + l_4 Pr_{27}(t) + l_3 Pr_{28}(t) + l_2 Pr_{29}(t) + l_1 Pr_{30}(t) \\ & + (-d_1 - d_2 - d_3 - d_4 - d_5) Pr_{31}(t) \end{aligned} \quad (35)$$

A solution of system (4)–(35) is carried out by the Runge-Kutta methods.

5 Conducting a Computational Experiment for Specific Regions Based on a Digital Model for Assessing Regional Competitiveness

The proposed approach was being tested on the example of several regions of the Volga Federal district. The research information and analytical database were developed. It used data from open sources [22, 23]. The values of leaf vertices of the cause-effect graph are determined for 2018. The analysis of competitiveness is carried out for the following regions of the Volga Federal district: Saratov region, Samara region, Ulyanovsk region, Republic of Tatarstan. These regions represent objects that differ significantly in terms of socio-economic development. They are indicative of the fact that, given the relative similarity of natural resource conditions, they belong to different categories in the ratings of the socio-economic status of the Russian Federation's subjects [24, 25].

A computational experiment was carried out for the same probabilities of risk development and managerial impact $l_1 = 0.04$; $l_2 = 0.4$; $l_3 = 0.3$; $l_4 = 0.4$; $l_5 = 0.5$; $d_1 = 0.3$; $d_2 = 0.4$; $d_3 = 0.5$; $d_4 = 0.5$; $d_5 = 0.5$. Preliminary calculations were performed to visualize the possible dynamic modes of the model (4)–(35). The obtained calculation results are presented in Figs. 3 and 4.

Initially, the low risk of losing the region's competitiveness at given parameter increases, passes the peak of the maximum value, decreases, and begins to grow again.

The results of a computational experiment (Fig. 3) with fixed values of the model parameters corresponding to the situation with a high level of crisis impacts and an intensive control action response to prevent crisis phenomena at the regional level, allows us to draw the following conclusions. The considered critical combinations of events in the case of deterioration in indicators of innovative and informational risk factors of development and increasing unemployment lead to the fact that the probability of a region's competitiveness decrease does not exceed 0.3 for all these regions. It should be noted that the risk for the leading region, to which the Republic of Tatarstan can be attributed based on existing ratings, is significantly lower and does not exceed 0.15. This is consistent with the widespread opinion of the expert community that leading regions retain their positions in the medium term.

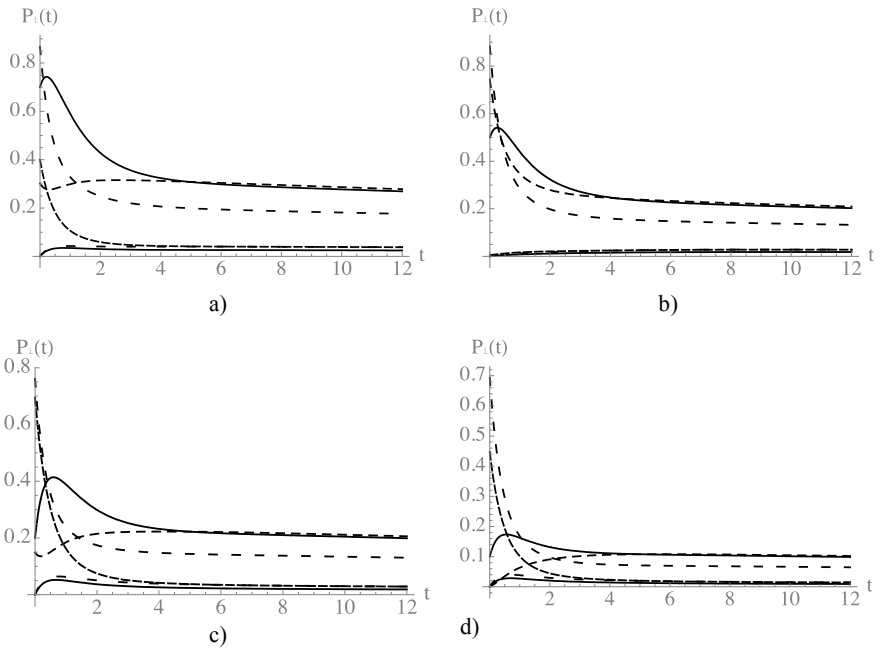


Fig. 3 The modeling results of analyzing the region’s competitiveness for Volga Federal district: **a** Saratov region; **b** Samara region; **c** Ulyanovsk region; **e** The Republic of Tatarstan

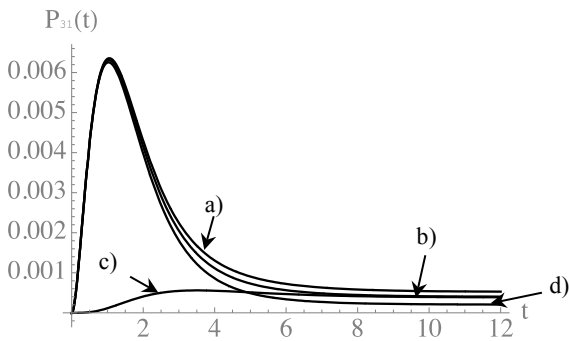


Fig. 4 The probabilities of joint critical realization of events: E_{17} —the number of unemployed, E_{19} —digitalization and using of the Internet, E_{20} —using the Internet electronic documentary exchange in organizations, E_{21} —R & D personnel, E_{22} —sales of innovative goods and services for the regions **a** Saratov region, **b** Ulyanovsk region, **c** Samara region, **d** the Republic of Tatarstan. The model uses the same values of the implementation of the risks and the control actions in the selected minimal cut set

In fig. Figure 4 shows the probabilities of a critical combination of events E_{17} — E_{19} — E_{20} — E_{21} — E_{22} for four different regions: Saratov Region, Ulyanovsk Region, Samara Region, Republic of Tatarstan. The probability values of the risk realization and the occurrence of process stopping due to a critical combination of the events in question, while they do not individually cause a certain risk to be initialized. In general, the simulation results correspond to the expected values, which confirms the adequacy of the model used.

The disadvantages of this approach include the need to neglect a large number of potentially significant indicators of vertices to reduce the size of the model. In the proposed scheme for constructing state graphs for minimal cut sets, this problem is insurmountable. However, due to the significant complexity of the structure of the considering indicators, digital modeling of the region's competitiveness based on the Kolmogorov-Chapman equations is an acceptable way for analyzing a multi-factor socio-economic system.

6 Conclusions

The competitiveness of the region has a dynamic nature and is formed in an evolutionary way depending on the influence of various direct and indirect factors. The dynamics of technical and technological changes, the development of management methods and systems, the change in the needs structure lead to a continuous increase in the number of factors that should be taken into account in the analysis of regional competitiveness. The competitiveness of the region is associated with the multifactorial concept of competitive potential, which characterized the possibility of successful participation in inter-regional and interethnic competitive relations. It assumes the presence and realization of the competitive potential of the regional system. In turn, the assessment of potential in the framework of competitive analysis involves the assessment of risks inherent in a given region. The formation of a hierarchical structure of indicators for risk assessment, as the initial stage of the study, allows the use of mathematical models based on the Kolmogorov-Chapman equations for analyzing the dynamics of the risks of regional competitiveness.

An important aspect of regional competitiveness researching is the analysis and forecasting of potential losses, the assessment of various risk types, and the identification of priority regional policy directions. The applicability of the methods based on the principles of building cyber-physical systems, using the Kolmogorov-Chapman equations to analyze the dynamics of regional development for assessing competitiveness risks is shown. For many strategies developed at the national level, the goal of competitiveness is to develop and maintain measures that encourage wealth (for example, low inflation, an effective institutional environment, open markets, etc.). At the level of subnational territories, the task is to determine priorities within the framework of the strategic plan in conditions of limited resources that have the most significant positive effect taking into account regional conditions.

Using modern digital mathematical models to assess such complex socio-economic objects as the regions by competitiveness analysis will allow us to obtain a more objective characterization. A feature of this study, in contrast to traditional approaches, is the study of digital models that provide assessing regional competitiveness in dynamics when implemented as a cyber-physical system.

Acknowledgements This research was partially supported by the Russian Fund of Basic Research (grant No. 20-010-00465).

References

1. Henderson, H.: Worldwide support found for measuring true wealth of nations. *Foresight* **10**, 67–69 (2008)
2. Porter, M.: The economic performance of regions. *Reg. Stud.* **37**, 549–578 (2008)
3. The Global Competitiveness Report 2019, http://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf. Last accessed 04.26.2020
4. Annoni, P., Dijkstra, L.: Measuring and monitoring regional competitiveness in the European Union. In: Huggins, R., Thompson, P. (eds.) *Handbook of Regions and Competitiveness, Contemporary Theories and Perspectives on Economic Development*, pp. 49–79. Edward Elgar Publishing Limited, London (2017)
5. SolAbility. *The Sustainable Competitiveness Report (2019)*. SolAbility Sustainable Intelligence, Zurich, Seoul (2019)
6. Delgado, M., Ketels, C.H., Porter, M., Stern, S.: *The Determinants of National Competitiveness*, No. 18249, 47 p. NBER Working Paper, Cambridge (2012)
7. Ketels, C.H.: Recent research on competitiveness and clusters: what are the implications for regional policy? *Camb. J. Reg. Econ. Soc.* **6**(2), 269–284 (2013). <https://doi.org/10.1093/cjres/rst008>
8. World Economic Forum. *Assessing the Sustainable Competitiveness of Nations*. In: *The Global Competitiveness Report 2013–2014*, pp. 53–82. World Economic Forum, Geneva (2013)
9. Enright, M.J.: Why clusters are the way to win the game? *World Link* **5**, 24–25 (1992)
10. Hausmann, R., Hidalgo, C.A., Bustos, S., Coscia, M., Chung, S., Jimenez, J., Simoes, A., Yildirim, M.A.: *The Atlas of Economic Complexity*. MIT Press, Boston (2013)
11. Martin, R.L.: Regional competitiveness: from endowments to externalities to evolution. In: Cooke, P., Asheim, B., Boschma, R., Martin, R., Schwartz, D., Tödtling, F. (eds.) *Handbook of Regional Innovation and Growth*, pp. 234–245. Routledge, London (2011)
12. Huggins, R., Izushi, H., Thompson, P.: Regional competitiveness: theories and methodologies for empirical analysis. *J. CENTRUM Cathedra Bus. Econ. Res. J.* **6**(2), 155–172 (2013)
13. Nezhivenko, E., Golovikhin, S., Dubynina, T., Dolinskaia, A.: Regional index of competitiveness of Russian regions. *Innov. Manag. Educ. Excellence Vis.* **2020**, 4921–4932 (2018)
14. Lomov, E.E.: The competitiveness of the region: elements, indicators, evaluation mechanisms. *News of Higher Educational Institutions. Volga Region. Economic Sciences.* **1**, 105 (2015)
15. Mazilkina, E.I., Panichkina, G.G.: *Fundamentals of competitiveness management*. Moscow: Omega-L, 328 p. (2008)
16. Stimson, R., Stough, R., Roberts, B.: *Regional Economic Development: Analysis and Planning Strategy*, 398 p. Springer Science & Business Media (2013)
17. Ketels, C.: Recent research on competitiveness and clusters: what are the implications for regional policy? *Camb. J. Reg. Econ. Soc.* **6**(2), 269–284 (2013)
18. Inshakov, O.V.: “The core of development” in the context of the new theory of factors of production. *Econ. Mod. Russian J.* **1**, 11–25 (2003)

19. Klyuyev, V.V., Rezchikov, A.F., Bogomolov, A.S., Filimonyuk, LYu.: Interaction of resources of complex man-machine systems in critical situation. *Control Diagn.* **4**, 41–45 (2013)
20. Bolshakov, A.A., Veshneva, I.V., Kushnikov, V.A.: Mathematical modeling of the dynamics of the behavior of social groups to prevent dangerous combinations of events in critical periods of the development of the state. In: *Actual Problems of Applied Mathematics, Computer Science and Mechanics: Sat. Proceedings of the International Scientific and Technical Conference*, pp. 592–599. Publishing House “Scientific Research Publications”, Voronezh (2017)
21. Rezchikov, A.F., Bogomolov, A.S., Ivashchenko, V.A., Filimonyuk, LYu.: Approach to ensuring and maintaining the security of complex systems based on automatic models. *Control Large Syst.* **54**, 179–194 (2015)
22. *Russian Statistical Yearbook 2019: Stat. Book*, 708 p. Rosstat, Moscow (2019)
23. *Regions of Russia. Socio-economic indicators*. https://gks.ru/bgd/regl/b19_14p/Main.htm. Last accessed 04.24.2020
24. Rating of the socio-economic situation of the constituent entities of the Russian Federation: results of 2018. http://vid1.rian.ru/ig/ratings/rating_regions_2019.pdf. Last accessed 05/04/2020
25. Skobelev, P.: Towards autonomous AI systems for resource management: applications in industry and lessons learned. In: *Proceedings of the XVI International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS 2018)*. vol. 10978, pp. 12–25, Springer (2018)

Society 5.0: Smart Technology for Ecology

Transformation of Reduced Compounds of Carbon and Nitrogen Oxides in the Cramped Aerodynamic Conditions



Vadim A. Zaytsev 

Abstract The joint transformation of typical urban atmosphere pollutants under cramped aerodynamic conditions is considered. These are the spaces between buildings with their highest concentrations, as a result of which reactions proceed quite intensively, with speeds greater than in unlimited volume. The kinetics of the process is described by a stiff system of differential equations, which is solved by the Rosenbrock method. The effect of the influence of reduced carbon compounds and nitrogen oxides in a wide range of their concentrations on the rate of formation of a highly toxic secondary pollutant is studied. In the course of numerical experiments, the conditions corresponding to its maximum value are determined. In this case, the transfer of substances from the reaction zone does not have time to occur.

Keywords Mathematical modeling · Systems of stiff differential equations · Environmental chemistry · Atmospheric chemistry · Chemical kinetics · Ecology

1 Introduction

A large number of publications are devoted to the study of atmospheric chemical reactions. For a quantitative assessment of the prevailing directions of these processes, depending on the conditions, as well as for use in problems of mathematical modeling, corresponding databases on chemical kinetics and photochemistry were created. In some of these works, the reactions taking place in the stratosphere are considered. At the same time, special attention is paid to the processes leading to a decrease in the concentration of ozone (the so-called “thinning” of the ozone layer). A completely different situation is typical for the troposphere, where an increase in ozone content is a negative factor. The starting point here was the study of photochemical smog [1–6].

V. A. Zaytsev (✉)

Dmitry Mendeleev University of Chemical Technology of Russia, Miuskaya sq., 9, Moscow 125047, Russia

e-mail: 8zaytsev@mail.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_13

157

Articles that contain material, already on the actual model study of the lower atmosphere, mainly cover the range of pollutant concentrations typical for an unlimited volume, i.e. when dilution has occurred or occurs over time with distance from emission sources. They often compare the results of a numerical experiment and direct measurements [7–10].

In contrast to the well-studied rates of processes occurring in an unlimited volume in the lower atmosphere and their dominant directions, depending on various factors, the intensity of chemical reactions in constrained aerodynamic conditions has not been studied in such detail [11–14]. This situation occurs quite often in other problems of hydrodynamics and mass transfer [15].

These works are based on the so-called base case, which corresponds to the averaged concentrations of pollutants for a number of stationary stations of GPBU “Mosecomonitoring” located near busy transport highways. For the formulation of the problem in them, it is not important that these values differ slightly at different stations and on different days (of course, since photochemical transformations are being studied, we are talking about the warm season with bright sunny weather). The articles [12, 13] analyze the situation that occurs when these concentrations increase by the same number of times or investigate the quantitative contribution of such primary pollutants as CH_4 and CO to the increase in the concentration of the secondary pollutant— O_3 , at concentrations of NO and NO_2 corresponding to the basic variant. Work [14] is devoted to the study of the effect of the relative content of the above-mentioned nitrogen oxides.

All these cases are quite possible when the flow of cars moves in the space between two groups of long and tall buildings located not far from each other, located on opposite sides near the motorway. There is a high probability of creating abnormally high concentrations of exhaust gases. The appearance of such new buildings in the zone of influence of high sources of emissions, for example, thermal power plants, initially designed for smaller houses, leads to similar consequences. In order not to create local concentrations in the area of the aerodynamic shadow of the building, the pipe must exceed 3.5 of its height.

An analysis of the possibility of the appearance and determination of concentrations of secondary pollutants—more toxic compounds formed during the process of transformation of primary pollutants—is of exceptional importance for atmospheric chemistry. In this case, the compounds of the first group include ozone and peroxyacetyl nitrates; to the second—hydrocarbons, carbon monoxide, nitrogen monoxide, nitrogen dioxide.

For this reason, in some works, it is precisely the negative effect of ozone on health and an increase in mortality of the population that is investigated. Thus, the article [16] studies the relationship between the content of ground-level ozone and the incidence rate of people in a low-urbanized area (Vyatskiye Polyany, Kirov region) and Moscow in the summer of 2010. In Vyatskiye Polyany, the absence of a relationship between these parameters at high air temperatures and low ozone concentrations. Thus, an attempt was made to isolate the effect of pure ozone. In contrast, when exposed to mean daily ground-level ozone concentrations above 60 mg/m^3 for 13 consecutive days, the number of ambulance calls increased. In the capital, the increase in ozone

content led to an increase in the frequency of respiratory diseases and mortality. Similar research is being carried out abroad. The works [17–22] show its contribution to the total additional mortality, mainly due to cardiovascular and respiratory diseases.

In view of the importance of the problem of secondary atmospheric pollution, whole reviews are regularly devoted to ozone, containing descriptions of the work of a fairly large number of researchers [23, 24].

2 Research Problem Statement

In works [12–14] the analysis of the influence of the concentrations of primary pollutants on the oxidation of methane in the atmosphere is carried out. Its content in the exhaust gases of cars with gasoline engines without catalytic converters is the highest in comparison with other hydrocarbons. The study of the process took place in the area between the buildings, which was presented as a very large room with a certain rate of air exchange. In these constrained aerodynamic conditions of the urban atmosphere, relatively high concentrations of primary pollutants can be achieved. Intense sunlight creates a favorable situation for photochemical reactions.

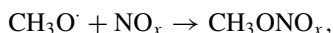
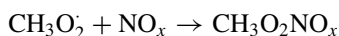
It is shown that the greatest contribution to the increase in the ozone content is made by a simultaneous increase in the concentration of nitrogen oxides or a relative increase in the concentration of NO_2 in comparison with NO . An increase in the content of nitrogen monoxide in comparison with its dioxide leads to the opposite result. The influence of the concentrations of reduced carbon compounds (CH_4 and CO) at various variations in the concentrations of NO and NO_2 remains unexplored. The effect of reduced carbon compounds in pure form has not been identified, which is very important from a theoretical point of view, since their presence in the air was a priori assumed. That is, the case was not considered when their concentrations, as well as the concentration of the product of the incomplete transformation of methane—formaldehyde, are equal to zero. Consequently, the estimation of the time for which there is a significant increase in the ozone concentration in these cases, as well as the search in the course of numerical experiments for conditions corresponding to its maximum value is the goal of the work.

The above estimates are important for studying parallel transformations of other hydrocarbons, their predominant directions under given conditions.

3 Developed Mathematical Model and Numerical Experiments

For the convenience of perception, the chemical equations underlying the mathematical model, described earlier in [12–14], are briefly presented again. Chemical reactions that are essential for the transformation of methane under the above conditions are presented in Table 1, as well as their rate constants [12–14].

The mathematical model does not take into account the process of ozone photolysis, since, at such high NO concentrations, the rate of O₃ sink according to reaction (14) is significantly higher than the rate of its photochemical decomposition. The loss of a part of intermediate radicals and molecules is not taken into account. This is due, for example, to the formation of so-called “reservoir gases”, which can then be decomposed by light into the original substances:



and with the formation of formic acid:

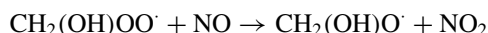
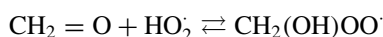
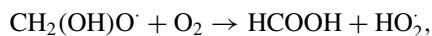


Table 1 Chemical reactions taken into account in the mathematical model and their rate constants

$\text{CH}_4 + \text{HO} \cdot \rightarrow \cdot\text{CH}_3 + \text{H}_2\text{O}$	$k_1 = 7.9 \times 10^{-15}$
$\cdot\text{CH}_3 + \text{O}_2 + \text{M} \rightarrow \text{CH}_3\text{O}_2 + \text{M}^*$	$k_2 [\text{M}] = 1.8 \times 10^{-12}$
$\text{CH}_3\text{O}_2 + \text{NO} \rightarrow \text{CH}_3\text{O} \cdot + \text{NO}_2$	$k_3 = 7.6 \times 10^{-12}$
$\text{CH}_3\text{O} \cdot + \text{O}_2 \rightarrow \text{CH}_2 = \text{O} + \text{HO}_2 \cdot$	$k_4 = 1.3 \times 10^{-15}$
$\text{CH}_2 = \text{O} + h\nu \rightarrow \cdot\text{CH} = \text{O} + \text{H} \cdot$	$k_5 = 3.7 \times 10^{-5}$
$\text{CH}_2 = \text{O} + h\nu \rightarrow \text{CO} + \text{H}_2$	$k_6 = 4.9 \times 10^{-5}$
$\text{CH}_2 = \text{O} + \text{HO} \cdot \rightarrow \cdot\text{CH} = \text{O} + \text{H}_2\text{O}$	$k_7 = 1.0 \times 10^{-11}$
$\cdot\text{CH} = \text{O} + \text{O}_2 \rightarrow \text{CO} + \text{HO}_2 \cdot$	$k_8 = 5.5 \times 10^{-12}$
$\text{CO} + \text{HO} \cdot \rightarrow \text{CO}_2 + \text{H} \cdot$	$k_9 = 3.0 \times 10^{-13}$
$\text{H} \cdot + \text{O}_2 + \text{M} \rightarrow \text{HO}_2 + \text{M}^*$	$k_{10} [\text{M}] = 1.4 \times 10^{-12}$
$\text{HO}_2 + \text{NO} \rightarrow \text{NO}_2 + \text{HO} \cdot$	$k_{11} = 8.3 \times 10^{-12}$
$\text{NO}_2 + h\nu \rightarrow \text{NO} + \text{O}(\text{}^3\text{P})$	$k_{12} = 8.9 \times 10^{-3}$
$\text{O}(\text{}^3\text{P}) + \text{O}_2 + \text{M} \rightarrow \text{O}_3 + \text{M}^*$	$k_{13} [\text{M}] = 1.8 \times 10^{-14}$
$\text{O}_3 + \text{NO} \rightarrow \text{NO}_2 + \text{O}_2$	$k_{14} = 1.8 \times 10^{-14}$



and also with the loss of the hydroxyl radical for the oxidation of other hydrocarbons.

The latter is partly justified by the application of modern norms, for example, the Euro-4 standard. It sets the benzene content in gasoline to no more than 1%. This slightly changes the hydrocarbon composition of exhaust gases in comparison with the same characteristic of previous years in the direction of the absolute predominance of methane in them, which is directly confirmed by the data of State Budgetary Institution “Mosecomonitoring”. Thus there is a decrease in the toxicity of emissions.

Therefore, the actual transformation times of a number of substances will be slightly longer than those obtained. However, as indicated above, the goal is only to estimate these values, which is necessary for a detailed analysis of the quantitative contribution to this effect of various variants of changes in the concentrations of reagents, as well as for constructing a more complex mathematical model that takes into account the parallel transformations of other hydrocarbons and their prevailing directions. under the given conditions.

Such modeling problems are reduced to solving systems of ordinary differential equations of the form (1):

$$\frac{dc_i}{d\tau} = F_i - R_i, \quad (1)$$

where c_i are the concentrations of each of the initial substances and products of these reactions (the unit of measurement is the number of particles in cm^3 of air), τ is the time, F_i and R_i are terms that take into account the generation and consumption of substances, respectively.

The resulting rigid system of 17 differential equations is solved numerically by one of the specially developed methods for systems of this class—the Rosenbrock method. It should be noted that the use of the Runge-Kutta-Felberg method here does not allow finding a solution for a time greater than the value of the order of 5×10^{-4} s, which is important; moreover, at a smaller value, a large solution error can be obtained. The initial concentrations correspond to the abnormally high ones reached in constrained aerodynamic conditions. According to the data of the State Budgetary Institution Mosecomonitoring, averaged for a number of stationary stations located near busy transport routes, the concentrations of such air pollutants as methane, carbon monoxide, nitrogen monoxide, nitrogen dioxide can, for example, be, respectively, 1.30, 0.83, 0.040, 0.060 mg/m^3 (basic variant). It is not very important here that there is a possible slight difference in these values at different stations and on different days [12–14].

At this stage, a simultaneous increase in the concentrations of nitrogen monoxide and nitrogen dioxide compared with the base variant by 2 and 5 times, while the content of reduced carbon compounds in each of these cases corresponds to either the base variant or increases by 5 times. These situations can relate to very severe air pollution from vehicles in the cramped aerodynamic conditions. Figure 1 shows the results of numerical experiments in which the influence of these parameters on the

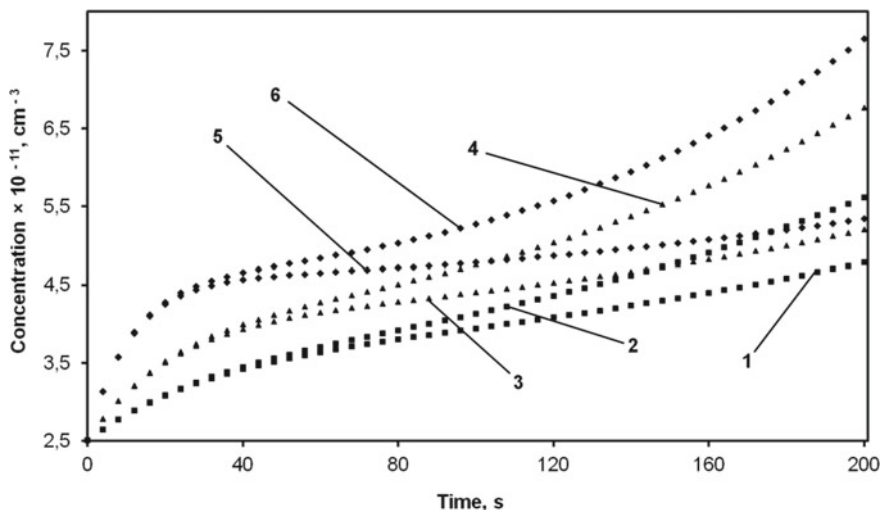


Fig. 1 Dependence of ozone concentration in the air on time with the different simultaneous growth of NO and NO₂ concentrations in comparison with the base case and cases of a fivefold increase in the concentrations of reduced carbon compounds: 1—basic variant (without increasing concentrations); 2—the concentration of reduced carbon compounds is increased by a factor of 5 in comparison with case 1; 3—simultaneous growth of NO and NO₂ concentrations by 2 times; 4—the concentration of reduced carbon compounds is increased by 5 times compared with case 3; 5—simultaneous growth of NO and NO₂ concentrations by 5 times; 6—the concentration of reduced carbon compounds increased 5 times compared to case 5

change in the ozone concentration is studied. In this chapter, a lot has been pointed out about its negative impact on health and an increase in mortality.

As follows from the figure, a significant increase in the O₃ content occurs in a time comparable to the time of its transfer from the region where the given chemical processes take place. The effect of reduced carbon compounds on the increase in the ozone content increases with the concomitant increase in the concentrations of nitrogen mono- and dioxide.

Then the effect of reduced carbon compounds is investigated in detail where it manifests itself most strongly, i.e. with a simultaneous increase in the concentrations of NO and NO₂ by 5 times. Numerical experiments are carried out for cases in comparison with this and include a change in the concentrations of reduced carbon compounds in a wide range from complete absence to an increase of 5 times, for each separately, and at the same number of times simultaneously (Fig. 2).

These numerical experiments reveal a significant effect of reduced carbon compounds on the growth of ozone content. At the same time, an increase in the CO concentration has a greater contribution than an increase in the CH₄ concentration, and their combined effect exceeds the individual effect for each of them. These patterns differ from those obtained for the cases of the impact of reduced carbon compounds on the base case [12, 13]. In the absence of reduced carbon compounds, the ozone concentration is lower (curve 2) than for their base content (curve 1).

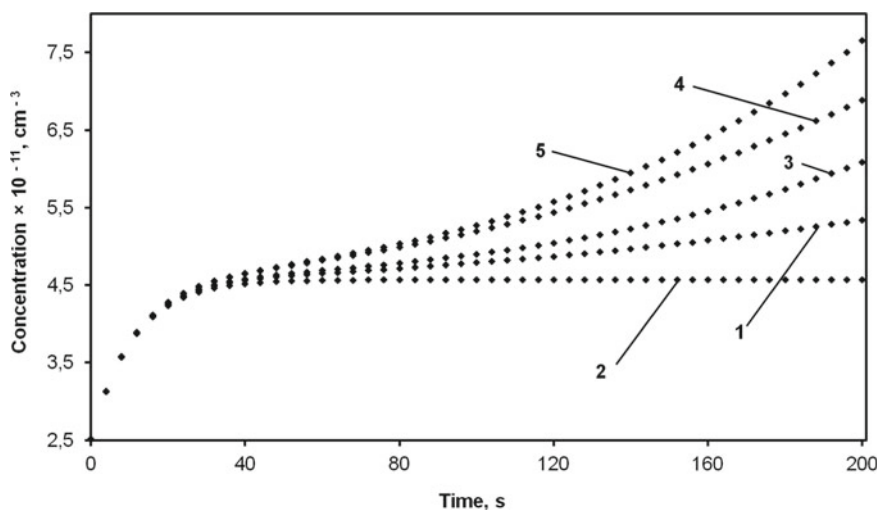


Fig. 2 Dependence of ozone concentration in the air on time at different concentrations of reduced carbon compounds in comparison with the option of the simultaneous increase in NO and NO₂ concentrations by 5 times: 1—the concentration of reduced carbon compounds is the same as in the base case; 2—the concentration of reduced carbon compounds is zero; 3—the concentration of CH₄ is increased by 5 times; 4—CO concentration increased by 5 times; 5—the concentration of reduced carbon compounds is increased 5 times

The article [14] considers a disproportionate change in the concentration of nitrogen oxides. It was found that the relative increase in the concentration of NO₂ compared to NO has the greatest contribution to the increase in the ozone content. However, the effect of the content of reduced carbon compounds under these conditions has not been studied. Meanwhile, it is nitrogen dioxide that is predominant in forest fires. And the transformation of nitrogen monoxide into dioxide as it is transferred from a relatively distant source (for example, a thermal power plant) can reach 80%. This is why this influence is important to consider. The results of the corresponding numerical experiments are shown in Fig. 3.

As follows from the figure, the effect of reduced carbon compounds on the increase in the ozone content increases with an increase in the concentration of nitrogen dioxide.

Further, by analogy with the above, the level of nitrogen dioxide concentrations (an increase in 2 times compared with the baseline version), in which reduced carbon compounds make the greatest contribution to secondary air pollution, is studied in detail.

These numerical experiments are shown in Fig. 4. They show approximately the same tendency as with the simultaneous growth of NO and NO₂ concentrations by 5 times.

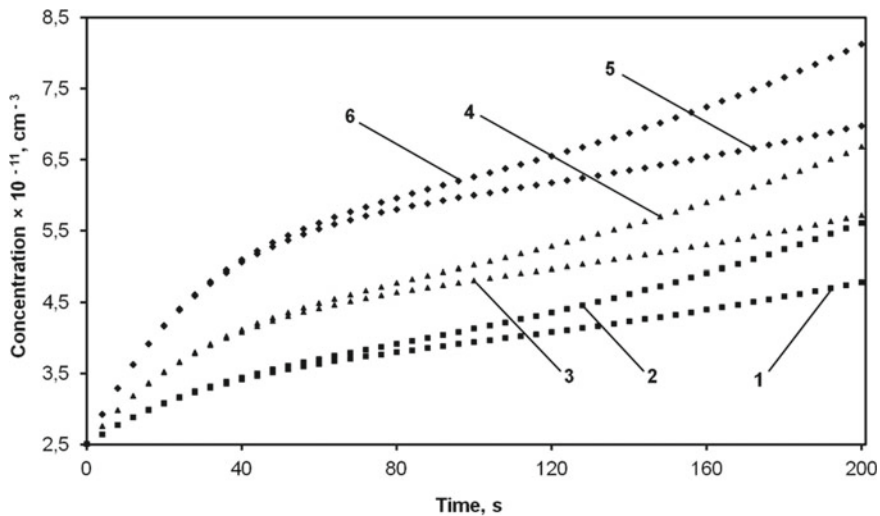


Fig. 3 Dependence of ozone concentration in the air on time with different growth of NO_2 concentration in comparison with the base case and cases of increase in the concentration of reduced carbon compounds by 5 times: 1—basic variant (without increasing concentrations); 2—the concentration of reduced carbon compounds is increased by a factor of 5 in comparison with case 1; 3— NO_2 concentration increased by 1.4 times; 4—the concentration of reduced carbon compounds is increased by 5 times compared with case 3; 5— NO_2 concentration is doubled; 6—the concentration of reduced carbon compounds is increased 5 times compared to case 5

The model was tested on the data of the State Budgetary Institution “Mosekomonitoring”. Numerical experiments show quite satisfactory agreement between the calculation results and the measured ozone concentration.

4 Conclusion

So, summarizing the above, we can draw the following conclusions:

1. The processes in the considered chemically active system occur rather quickly before the substances leave the region of the reaction.
2. With a simultaneous, proportional increase in the concentration of nitrogen mono- and dioxide, as well as the concentration of NO_2 without changing the concentration of NO , the effect of reduced carbon compounds on the increase in the ozone content increases.
3. The effect of reduced carbon compounds has been studied in detail where it manifests itself most strongly, i.e. with a simultaneous increase in the concentrations of NO and NO_2 by 5 times and NO_2 without NO by 2 times. A significant effect of reduced carbon compounds on the growth of the ozone content was found. At the same time, an increase in the CO concentration has a greater contribution

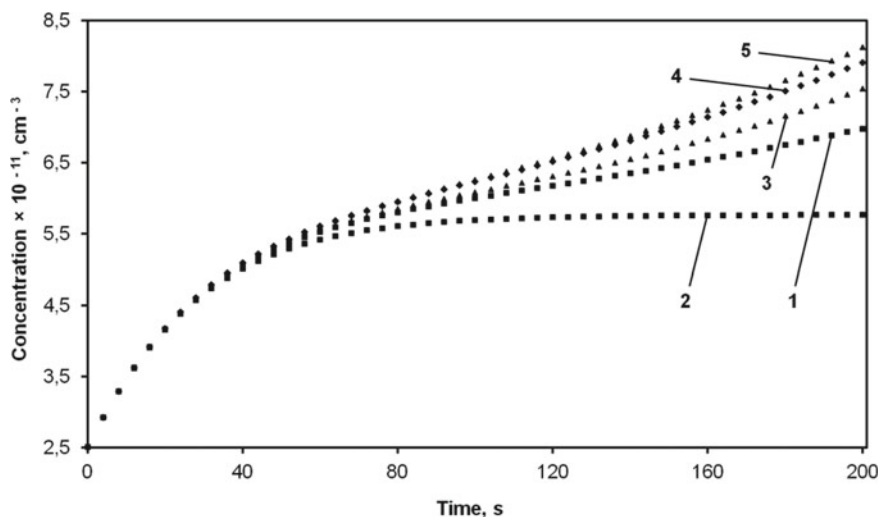


Fig. 4 Dependence of ozone concentration in the air on time at different concentrations of reduced carbon compounds in comparison with the option of doubling the NO_2 concentration: 1—the concentration of reduced carbon compounds is the same as in the base case; 2—the concentration of reduced carbon compounds is zero; 3—the concentration of CH_4 is increased by 5 times; 4—CO concentration increased by 5 times; 5—the concentration of reduced carbon compounds is increased 5 times

than an increase in the CH_4 concentration, and their combined effect exceeds the individual effect for each of them. These patterns differ from those obtained for the cases of the impact of reduced carbon compounds on the base case [12, 13].

4. For the most unfavorable of the considered cases, the ozone content increases by more than 3 times, which is approximately 2 times higher than its maximum permissible average daily concentration. However, this value is less than the maximum permissible maximum one-time concentration, the use of which is justified here to a greater extent.

Acknowledgements The research was supported by the Russian Foundation for Basic Research (grant 16-08-01252).

References

1. DeMore, W.B., Margitan, J.J., Molina, J.J., et al.: Chemical kinetics and photochemical data for use in stratospheric modeling. NASA panel for data evaluation. JPL Publ. **82-57**, 1–186 (1982)
2. Atkinson, R., Lloyd, A.C.: Evaluation of kinetic and mechanistic data for modeling of photochemical smog. J. Phys. Chem. Ref. Data **13**(2), 315–444 (1984)

3. Baulch, D.L., Cox, R.A., Hampson, R.F., et al.: Evaluated kinetic and photochemical data for atmospheric chemistry. *J. Phys. Chem. Ref. Data* **11**(2), 328–496 (1982)
4. Baulch, D.L., Cox, R.A., Hampson, R.F., et al.: Evaluated kinetic and photochemical data for atmospheric chemistry. *J. Phys. Chem. Ref. Data* **13**(4), 1259–1379 (1984)
5. Demerjan, K.L., Shere, K.L., Peterson, J.T.: Theoretical estimates of actinic (spherically integrated) flux and photolytic rate constants of atmospheric species in the lower troposphere. *Adv. Environ. Sci. Technol.* **10**, 369–459 (1980)
6. Anderson, L.C.: Atmospheric chemical kinetic data survey. *Rev. Geophys. Space Phys.* **10**, 369–459 (1976)
7. Konovalov, I.B., Elanskii, N.F., Belikov, I.B., Zvyagintsev, A.M., Beekmann, M.: Validation of chemistry transport model of the lower atmosphere in the Central European region of Russia using ground-based and satellite measurement data. *Russ. Meteorol. Hydrol.* **34**(4), 236–242 (2009). <https://doi.org/10.3103/S1068373909040062>
8. Zaripov, R.B., Kuznetsova, I.N., Konovalov, I.B., Belikov, I.B., Zvyagintsev, A.M.: WRF ARW and CHIMERE models for numerical forecasting of surface ozone concentration. *Russ. Meteorol. Hydrol.* **36**(4), 249–257 (2011). <https://doi.org/10.3103/S1068373911040054>
9. Konovalov, I.B., Beekmann, M., Kuznetsova, I.N., Glazkova, A.A., Zaripov, R.B., Vasil'eva, A.V.: Estimation of the influence that natural fires have on air pollution in the region of Moscow megalopolis based on the combined use of chemical transport model and measurement data. *Izvestiya. Atmos. Oceanic Phys.* **47**(4), 457–467 (2011). <https://doi.org/10.1134/S0001433811040062>
10. Zaripov, R.B., Konovalov, I.B., Glazkova, A.A.: Modeling the concentration of pollutants using the WRF-ARW atmospheric model and CHIMERE chemistry transport model. *Russ. Meteorol. Hydrol.* **38**(12), 828–839 (2013). <https://doi.org/10.3103/S1068373913120042>
11. Zaytsev, V.A.: Chemical transformations vehicle exhaust in cramped aerodynamic conditions of city. *Mathematical methods in engineering and technology. MMTT-29: XXIX International Scientific Conference: Proceedings*, vol. 10, pp. 24–28 (2016)
12. Zaytsev, V.A.: Quantitative characteristics of chemical transformations of pollutants in cramped aerodynamic conditions. *MMTT-31: XXXI International Scientific Conference: Proceedings*, vol. 4, pp. 86–90 (2018)
13. Zaytsev, V.A.: Some quantitative characteristics of the transformation of pollutant substances in the cramped aerodynamic conditions of the urban atmosphere. *Izvestiya Sankt-Peterburgskogo gosudarstvennogo tekhnicheskogo instituta (tekhnicheskogo universiteta)* **47**(73), 128–132 (2018)
14. Zaytsev, V.A.: Features of the kinetics of gas impurities in cramped aerodynamic conditions of an urban atmosphere. *MMTT-32: XXXII International Scientific Conference: Proceedings*, vol. 12–3, pp. 57–62 (2019)
15. Pokusaev, B.G., Zaitsev, A.A., Zaitsev, V.A.: Transfer processes under slug flow conditions in three-phase media. *Theor. Found. Chem. Eng.* **33**(6), 539–549 (1999)
16. Kotelnikov, S.N., Stepanov, E.V., Ivashkin, V.T.: Ozone concentration in the ground atmosphere and morbidity during extreme heat in the summer of 2010. *Doklady Biol. Sci.* **473**(1), 64–68 (2017). <https://doi.org/10.1134/S0012496617020107>
17. Fang, Y., Naik, V., Horowitz, L.W., Mauzerall, D.L.: Air pollution and associated human mortality: the role of air pollutant emissions, climate change and methane concentration increases from the preindustrial period to present. *Atmos. Chem. Phys.* **13**, 1377–1394 (2013)
18. Bates, D.V.: Ambient ozone and mortality. *Epidemiology* **16**(4), 427–429 (2005)
19. Hollingsworth, J.W., Kleeberger, S.R., Foster, W.M.: Ozone and pulmonary innate immunity. *Proc. Amer. Thorac. Soc.* **4**, 240–246 (2007)
20. Longphre, M., Zhang, L.-Y., Harkema, J.R., Kleeberger, S.R.: Ozone-induced pulmonary inflammation and epithelial proliferation are partially mediated by PAF. *J. Appl. Physiol.* **86**(1), 341–349 (1999)
21. Gryparis, A., Forsberg, B., Katsouyanni, K., Analitis, A., Touloumi, G., Schwartz, J.: Acute effects of ozone on mortality from the “Air pollution and health: A European project approach”. *Amer. J. Respir. Crit. Care Med.* **170**, 1080–1087 (2004)

22. Bell, M., Dominici, F.: Analysis of threshold effects for short-term exposure to ozone and increased risk of mortality. *Epidemiology* **17**(6), 223 (2006)
23. Elansky, N.F.: Russian studies of atmospheric ozone in 2007–2011. *Izvestiya. Atmos. Oceanic Phys.* **48**(3), 281–298 (2012). <https://doi.org/10.1134/S0001433812030024>
24. Elansky, N.F.: Russian studies of atmospheric ozone in 2011–2014. *Izvestiya. Atmos. Oceanic Phys.* **52**(2), 132–146 (2016). <https://doi.org/10.1134/S0001433816020031>

Numerical Modeling of Pneumatic Conveying in the Mode of the Inhibited Dense Layer



Asia. G. Mukhametzyanova, Andrei O. Pankov,
and Alina A. Abdrakhmanova

Abstract Using the methods of computational fluid dynamics, we studied the process of pneumatic conveying in the mode of the inhibited dense layer. The results of numerical modeling correlate well with other scientists' data. The impact of the particle movement mode and shape of the arrestor is assessed on the stability of the process of pneumatic conveying in the mode of the inhibited dense layer.

Keywords Pneumatic conveying · Inhibited dense layer · Numerical modeling · Particle motion mode · Reliable conveying velocity

1 Introduction

Pneumatic conveying in the mode of the inhibited dense layer (IDL) is a perspective method of transporting bulk materials for short distances. These conditions allow to preserve the dense layer along the whole conveyor tube, and to get a high concentration of the solid phase in the flow, bringing its value to concentration in the stationary layer. To create an inhibited dense layer at the end of the conveyor tube conic narrowing parts with the conic angle of $1\text{--}2^\circ$ are created, or arresters (diaphragm, cone head, etc.) are installed [1].

The properties that define pneumatic conveying in the mode of the inhibited dense layer are low consumption of the transported gas, the low conveying velocity with preservation of, system capacity in the solid phase, preservation of high concentration along the conveying tube, easiness of regulation of velocity of the bulk material flow

Asia. G. Mukhametzyanova (✉) · A. O. Pankov · A. A. Abdrakhmanova
Kazan National Research Technological University, Kazan 420015, Russia
e-mail: asia@kstu.ru

A. O. Pankov
e-mail: pankov.andrey@gmail.com

A. A. Abdrakhmanova
e-mail: abdrakhmanova.alina93@yandex.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_14

over a broad area, etc. [2]. One of the main advantages of pneumatic conveying in the mode of inhibited dense layer is the opportunity to combine with other technological processes, such as cleaning from light impurities with pneumatic separators, drying with the stream of the warmed gas, catalysis, pyrolysis, etc. [3].

Application of pneumatic conveying in the mode of the inhibited dense layer could be quite useful in many cases for solving a whole range of technological problems. The main problem for the performance of pneumatic conveying in the mode of inhibited dense layer is lack of studies, lack of impression of the structure of the inhibited dense layer, and impact on it of various technological, structural, and mode's parameters.

The reliability and economic feasibility of pneumatic conveying depend on the right choice of the airflow velocity. The increase of this velocity reduces the operational life of pneumatic pipes, leads to an increase of pressure loss in the network, and electrical power consumption, and its reduction causes fallout of the transported material in the pneumatic pipe and blockage of the latter. The right choice of the velocity of the transporting gas is extremely important both for sustainable conveying mode and for energy efficiency.

Controlling influence over speed of the material in pneumatic conveying in the mode of the inhibited dense layer has a hydrodynamic environment created in the braking area. Moreover, the concentration of the material transferred in the airflow also affects transportation speed [4]. Work [5] shows that the speed of the bulk material in the mode of the inhibited dense layer depends on the physical-mechanical properties of the granular material, gas speed, dimensions, and form of the narrowing arrester and does not depend on the length and unevenness of the transport pipe.

2 Background and Problem Statement

The reliable transporting speed should be higher than the engulfment velocity. This velocity may be in large excess over the first velocity of pseudo-fluidization in usual modes of pneumatic conveying and may be almost the same as the terminal velocity of the particles (the second critical velocity of pseudo-fluidization).

In course of pneumatic conveying in the mode of inhibited dense layer movement of the particles is restricted so much that terminal velocity becomes much lower and engulfment velocity becomes almost the same as the first velocity of pseudo-fluidization. The extension of the layer during pneumatic conveying is not large, its structure is almost the same as of the solid phase of the fluidized state. Therefore, gas speed should be a few times higher than the velocity at the beginning of the pseudo-fluidization of the transported material. In this case, we mean gas velocity in the lower part of the conveying tube. While moving up in the transport pipe the gas velocity increases due to its extension, bearing capacity increases, and engulfment hazard is excluded.

The process of the established pneumatic conveying of solid spherical particles along the smooth vertical part of the pipe is studied herein with methods of computational fluid dynamics. STAR-CCM + computational complex was used as the studies instrument, that allows obtaining solutions for a wide range of physical problems, including for transient circuits, coupled heat transfer, exposure, laminar, and turbulent frictional flow, etc.

The approach to the modeling of movement of two-phase flows is majorly determined by the intensity of mutual interaction and depends on the volume ratio of the solid phase in a flow [6–8].

If the volume ratio of solid particles is less than 10–6, the particles do not affect the movement of the gas flow. If the volume ratio of solid particles is increased to 10–3, the interaction between the gas and the particles is observed, however, the movement of separate particles is regarded to be independent of each other [6]. In the case of higher concentrations of solid particles interaction between the particles in a flow is observed, as well as between the particles and the device walls that lead to changes of movement of the gas phase.

In course of studies of fluid dynamics of two-phase flows, a few model approaches that differ from each other by the method of review of the interacting phases and link between them are formulated and clarified.

The work [9] provides a classification of imitation models depending on the area of their application and description of the interaction of the phases.

According to the authors [8–17], the most practical bearing for modeling of the weighted layer has the model of discrete elements and the continuous model, the principal difference between which is the description of the movement of the discrete particles, and the common problem is the preparation of the closing relations, and, particularly, determination of the mechanical interaction of the phases.

Movement of the gas phase for discrete element model or Eulern-Lagrangian model is solved in Euler representation, as a continuum in an immovable coordinate system, and movement of particles is represented with Lagrangian equations in the movable coordinate system that are integrated along their trajectory [6, 8].

The model of discrete elements, which is solved by integration of the particle movement equations within the velocity field of the bearing gas, allows obtaining statistical data on the movement of the separate particles with regard to their density, dimensions, shape, and polydispersity of the particles [6].

However, frequent application of the model of discrete particles in full for large concentrations of particles is restricted due to complexity and long term of computations made in a few iterations [13, 14].

3 Numerical Modeling

For the purposes of studies of two-phase flows, particularly, of the gas-solid particles flow STAR-CCM + computational complex provides for the Lagrangian multiphasicity model and Discrete Element Method—DEM.

For numerical modeling of two-phase gas-solid particles flow Discrete Element Method—DEM was applied, that allowed to design clashes of particles of any shape and to consider gyration of the particles affected by the forces and their moments.

Discrete Element Method DEM is a set of numerical methods for the design of a large number of particles of solid materials. The key provision of the method is, that the material is comprised of separate discrete particles, that may have various properties and surface shapes, elastic properties, density, etc.

Modeling starts with moving of all of the particles to a specific position and assignment to them of the initial speed. Then the forces affecting every particle are calculated by the initial data and the relevant physical laws. For every particle dynamic problem is solved that includes determination of the active forces and movement trajectory [17].

For calculation of the terminal velocity of solid spherical particles in a vertical pneumatic conveying system of three-dimensional Navier–Stokes equations, averaged by Reynolds (Formulas 1–7) is solved, supplemented with discrete methods based on the Lagrangian approach.

Airflow is described as follows:

- mass preservation equation

$$\bar{\nabla}(\vec{V}) = 0 \quad (1)$$

- impulse preservation equation

$$\frac{\partial \vec{V}}{\partial t} + \bar{V}(\vec{V}\vec{V}) = -\frac{\nabla p}{\rho} + \frac{\nabla(\bar{\tau} + \bar{\tau}_t)}{\rho}; \quad (2)$$

viscous tensor $\bar{\tau}$ is determined with Newton rheological law:

$$\bar{\tau} = \mu * (\nabla \vec{V} + [\nabla \vec{V}]^T) - \frac{2}{3} * \mu \nabla * \vec{V} * \bar{I}, \quad (3)$$

and Reynolds stress tensor $\bar{\tau}_t$ —in accordance with the generalized hypothesis by Boussinesq:

$$\bar{\tau}_t = \mu_t * (\nabla \vec{V} + [\nabla \vec{V}]^T) - \frac{2}{3} * \mu_t \nabla * \vec{V} * \bar{I} - \frac{2}{3} * \rho * k * \bar{I}. \quad (4)$$

To determine turbulence specifications – ε turbulence model was applied:

$$\frac{\partial}{\partial t}(\rho * k) + \nabla[\rho * \vec{V} * k] - (\mu + \frac{\mu_t}{\sigma_k} * \nabla) = \mu_t * P - \rho * \varepsilon, \quad (5)$$

$$\frac{\partial}{\partial t}(\rho * \varepsilon) + \nabla[\rho * \vec{V} * \varepsilon - (\mu + \frac{\mu_t}{\sigma_\varepsilon} * \nabla \varepsilon)] = C_{\varepsilon 1} * \frac{\varepsilon}{k} * -\mu_t * P - C_{\varepsilon 2} * \frac{\varepsilon^2}{k}; \quad (6)$$

Generation component in transfer equations

$$P = \mu * (\Delta \vec{V} + [\Delta \vec{V}]^T) * \Delta \vec{V}, \quad (7)$$

where —turbulence kinetic power; ε —turbulence kinetic power velocity; μ_t —turbulence viscosity; $C_{\varepsilon 1}$, $C_{\varepsilon 2}$ —semiempirical ratios of turbulence model.

General equation of impulse preservation for the material particle as follows

$$m_p = \frac{dv_p}{dt} = F_s + F_b; \quad (8)$$

where F_s —forces affecting the particle surface; F_b —mass forces affecting the particle. These forces, in their turn, divide into

$$F_s = F_d + F_p; \quad (9)$$

$$F_b = F_g + F_c; \quad (10)$$

where F_d —resistance force; F_p —tensile force; $F_g = m_p g$ —gravitational force; F_c —additional forces of the body that are represented by the interaction of the particles with other particles and boundaries of the area. At that, F_c is represented as follows

$$F_c = \sum_{\text{neighbor.particles}} F_{\text{contact}} + \sum_{\text{neighbor.boundaries}} F_{\text{contact}}; \quad (11)$$

In addition to impulse preservation equation (8), the discrete element model includes an equation for computation of the particles turn:

$$\frac{d}{dt} L_p = \frac{d}{dt} (I_p \omega_p) = \sum_{\text{neighbor.particles}} T_c + \sum_{\text{neighbor.particles}} T_A; \quad (12)$$

where the torque is calculated as follows

$$T_c = r_c F_c - \mu_r |r_c| |F_c| \frac{\omega_p}{|\omega_p|}; \quad (13)$$

where L_p —angular moment of the particle; I_p —inertia moment of the particle; T_c —torque of the separate particle that occurs as a result of the operation of the exchange forces, applied to the particle in the point other than the particle center of mass; r_c —vector from the center of mass to contact point; μ_r —coefficient of rolling friction; ω_p —the angular spin rate of the particle. The coefficient of rolling friction

is determined during verification with the experimental model and is set as 0.3. Force in (9) and (10) are modeled as follows.

The equation for tensile force F_d

$$F_d = \frac{1}{2} C_d \rho A_p |v_s| v_s; \quad (14)$$

where C_d —coefficient of rolling friction; ρ —air density; v_s —slide speed of the particle; A_p —projected area of the particle.

The equation for tensile force F_p

$$F_p = -V_p \nabla p_{static}; \quad (15)$$

where V_p —particle volume; ∇p_{static} —static pressure gradient of the basic phase.

For a description of the interaction between $F_{constant}$ particles Hertz-Mindlin interaction model was used that is based on the Hertz-Mindlin interaction theory.

Pursuant to this model the forces affecting between these two spheres A and B are described by the following group of the equations.

$$F_{contact} = F_n + F_t; \quad (16)$$

where F_n —normal and F_t —tangential forces.

Normal force:

$$F_n = -K_n d_n - N_n v_n; \quad (17)$$

where $K_n = \frac{3}{4} E_{eq} \sqrt{d_n + R_{eq}}$ —normal spring tension; $N_n = \sqrt{(5K_n M_{eq})} N_{n \text{ damp}}$ —normal damping; $N_{n \text{ damp}}$ —normal decay coefficient, determined by the Eq. (20).

Tangential force is determined as follows

$$F_t = -K_t d_t - N_t v_t; \quad (18)$$

if $|K_t d_t| < |K_n d_n| C_{fs}$;

$$F_t = \frac{|K_n d_n| C_{fs} d_t}{|d_t|}; \quad (19)$$

$|K_t d_t| \geq |K_n d_n| C_{fs}$,

where C_{fs} —coefficient of static friction, tangential spring tension $K_t = 8G_{eq} \sqrt{d_n R_{eq}}$, tangential damping; $N_t = \sqrt{(5K_t M_{eq})} N_{t \text{ damp}}$, $N_{t \text{ damp}}$ —tangential decay coefficient, determined by Eq. (21)

$$N_{n \text{ damp}} = \frac{-\ln(C_{n \text{ rest}})}{\sqrt{\pi^2 + \ln(C_{n \text{ rest}})^2}}; \quad (20)$$

$$N_{t \text{ damp}} = \frac{-\ln(C_{t \text{ rest}})}{\sqrt{\pi^2 + \ln(C_{t \text{ rest}})^2}}; \tag{21}$$

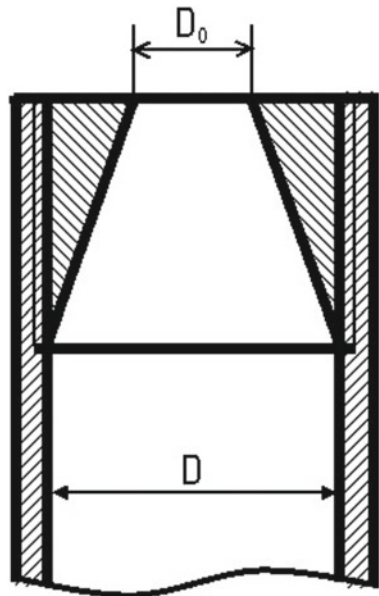
where $C_{n \text{ rest}}$ and $C_{t \text{ rest}}$ —normal and tangential elasticity coefficient; R_{eq} , M_{eq} , E_{eq} , G_{eq} —equivalent radius of particles, Young’s modulus, shear modulus correspondingly.

As initial data were taken physical parameters of the air under normal conditions, input, and output velocity. Configurations are set for the interaction of the particles with each other and with walls of the computational region.

The problem for setting the movement of the traffic flow in a rectangular channel with a diameter of 0.02 m and length of 2.2 m. Spherical particles with a diameter of 0.003 are taken. Change of modes (transfer from the inhibited dense layer to piston flow) was determined according to output pressure pulsation—in the mode of an inhibited dense layer, this parameter does not exceed 10% [18].

Numerical modeling of the movement process of two-phase flow along the conveying tube allowed to detect velocity fields of both of the phases, pressure, and other values. The results of numerical modeling correlate well with other scientists’ data [2]. The impact of the particle movement mode and shape of the arrestor is assessed (Fig. 1) on the stability of the process of pneumatic conveying in the mode of an inhibited dense layer, minimum gas velocities for transportation of the bulk materials in various conditions are determined.

Fig. 1 Example of the controller (arrestor)



4 Results and Discussion

The following values varied in the numerical experiments: compression rate D_0/D ; conveying tube diameter D ; density of the transported materials (in Table 1 it is implicitly expressed by the first velocity of pseudo-fluidization).

Analysis of the results of numerical modeling shows that the value of the minimum gas velocity during pneumatic conveying of bulk materials in the mode of inhibited dense layer is 30–70% higher than the velocity at the beginning of pseudo-fluidization. The velocity at the beginning of pseudo-fluidization was determined in accordance with Todes formula [19].

An increase of simplex D_0/D (that is the reduction of the compression level) leads to the growth of the minimum gas velocity. A maximum increase occurs upon achievement of the critical values D_0/D , which are the values when the mode of inhibited dense layers transfers to piston pneumatic transport. The maximum value of this increase amounts to 50–100%, and the upper limit may be regarded as reliable conveying gas velocity.

The results of numerical modeling show that the increase of conveying tube diameter causes a reduction of the minimum velocity value. Apparently, this can be explained by the fact that tubes of smaller diameter are more prone to piston formation. Therefore, the tubes of larger diameter ensure a more stable process, and a stationary mode of transportation is established earlier.

Table 1 Results of numerical modeling of pneumatic conveying of bulk materials in the mode of the inhibited dense layer in various conditions

Pseudofluidization velocity (m/s)	Conveying tube diameter, $D * 10^2$ m	Minimum conveying velocity (m/s)		Pseudofluidization number		Reliable conveying velocity (m/s) (comp.)
		Maximum inhibition	Minimum inhibition	Maximum inhibition	Minimum inhibition	
0.215	15	0.29	0.41	1.35	1.90	0.46
0.334	15	0.51	0.60	1.52	1.80	0.71
0.104	16	0.17	0.21	1.61	2.04	0.23
0.104	28	0.17	0.19	1.66	1.84	0.19
0.118	16	0.16	0.18	1.39	1.54	0.23
0.118	28	0.17	0.18	1.42	1.50	0.27
0.118	40	0.17	0.22	1.47	1.87	0.26
0.373	28	0.50	0.59	1.34	1.58	0.73
0.370	40	0.59	0.66	1.57	1.77	0.73
0.531	28	0.74	0.90	1.40	1.70	0.97
0.531	40	0.71	0.77	1.34	1.45	0.98

5 Conclusion

Thus, we can conclude that for the purposes of determination of the minimum allowable velocity of pneumatic conveying of the bulk materials it is necessary to carry out numerical experiments for specific terms, and unless it is possible to carry out numerical computations, this velocity shall be equal to the doubled value of the first velocity of pseudo-fluidization.

References

1. Li, Hongzhong; Mooson Kwauk. *Chem. Eng. Sci.* **44**(2), 261–271 (1989)
2. Razinov Y.I.: Study of vertical pneumatic conveying of bulk materials with inhibited dense layer. Dissertation Thesis for Ph.D. in Engineering (Candidate of Engineering Sciences), p. 212. Kazan (1974) (in Russian)
3. Mills, D., Jones, M.G., Agarwal, V.K.: *Handbook of Pneumatic Conveying Engineering*, p 720. CRC Press (2004)
4. Yu, M.A.: Numerical modeling of impact of the air flow on spherical particles in the round-section air duct. In: *Modern Issues of Science and Education*. No. 2 (part 2), pp. 107–113 (2015)
5. Yu, M.A., Kolosnitsin, A.N.: Development of numerical methods of vacuum cleaning system design. *Herald of Civil Engineers* **6**(59), pp. 151–155 (2016)
6. Klinzing, G.E., Rizk, F., Marcus, R., Leung, L.S.: *Pneumatic Conveying of Solids. A Theoretical and Practical Approach*, p. 435. Springer (2010)
7. Yu, V.A.: *Turbulentnye techeniya gaza s tverdyimi chasticami [Turbulent Flows of Gas with Solid Particles]*, 192 p. Moscow, Fizmatlit (2003) (in Russian)
8. Ostrovsky G.M.: *Pnevmaticheskiyi transport sypuchih materialov v himicheskoyi promyshlennosti [Pneumatic Transport of Loose Materials in the Chemical Industry]*, 104 p. Lenin-grad, Khimiia (1984) (in Russian)
9. Schreiber, M.: *Modellierung von Hydrodynamik und Wärmeübertragung in blasenbildenden Wirbelschichten zur Trocknung von Braunkohle: akademischen Grades eines Doktor-Ingenieurs genehmigte Dissertation. Brandenburgischen Technischen Universität Cottbus*, 167 p (2013)
10. van der Hoef, M.A., van Sint Annaland, M., Deen, N.G., Kuipers, J.A.M.: Numerical Simulation of Dense Gas-Solid Fluidized Beds: A Multiscale Modeling Strategy. *Annual Review of Fluid Mechanics*, vol. 40, pp. 47–70 (2008)
11. van der Hoef, M.A., Ye, M., van Sint Annaland, M., Andrews, A.T., Sundaresan, S., Kuipers, J.A.M.: Multiscale modeling of gas-fluidized beds. *Adv. Chem. Eng.* **31**, 65–149 (2006)
12. Elghobashi, S.: Particle-Laden turbulent flows. Direct simulation and closure models. *Appl. Sci. Res.* **48**, 301–314 (1991)
13. *Flow Simulation 2014 Technical Reference*. Dassault Systems, 204 p (2014)
14. Gui, N., Fan, J.R., Luo, K.: DEM-LES study of 3-D bubbling fluidized bed with immersed tubes. *Chem. Eng. Sci.* **63**(14), 3654–3663 (2008)
15. Chiesa, M., Mathiesen, V., Melheim, J.A., Halvorsen, B.: Numerical simulation of particulate flow by the Eulerian-Lagrangian and the Eulerian-Eulerian approach with application to a fluidized bed. *Comput. Chem. Eng.* **29**(2), 291–304 (2005)
16. Gidaspow, D., Ettehadieh, B.: Fluidization in two-dimensional beds with a jet—2 hydrodynamic modeling. *Ind. Eng. Chem. Fundam.* **22**(2), 193–201 (1983)
17. Gidaspow, D.I.: *Multiphase flow and fluidization continuum and kinetic theory descriptions*, p. 467. Academic Press Harcourt Brace & Company, San Diego (1994)

18. Mukhametzhanova, A.G., Pankov, A.O., Abdrakhmanova, A.A.: Selection of Criteria of Pneumatic Conveying in the Mode of Inhibited Dense Layer, vol. 22, No. 10, pp. 72–76. Herald of Engineering University (2019)
19. Yu, V.V., Belova, O.V., Skibin, A.P., Zhuravlev, O.N.: Determination of Fluid Dynamic Performance of Restrictor with Labyrinth Sealing Using Computational Fluid Dynamics, pp. 55–64. Herald of the Bauman Moscow State Technical University (2012)

Intelligent System for Determining the Presence of Falsification in Meat Products Based on Histological Methods



Alexander Bolshakov , Marina Nikitina , and Renata Kalimullina

Abstract An intelligent system is proposed to support decision-making to determine the presence of falsification in meat products based on histological studies. The tasks that must be solved for its development are formulated. The system architecture for the implementation of the formulated tasks is proposed. It includes, in particular, the expert subsystem and the decision support subsystem. Formalization of knowledge for making a decision about the presence of counterfeit is carried out on the basis of production rules. They are generated based on information contained in morphological tables. The development of a prototype ES is carried out in the programming language of artificial intelligence Prolog. The results of optimizing the approximation of a polychrome image of slices of meat products are described. A method for solving it based on a genetic algorithm is proposed. A program was developed in the C++ programming language, using which a complex of computational experiments was carried out to study the nature of the dependence of the convergence of the optimization process. Various parameters varied: the maximum number of iterations, the choice of the initial population. Further studies in this direction are related to the implementation of all the necessary image processing algorithms to automatically identify the required values of the characteristics of the image fragments for use in the cyber-physical system of automation of the process of identifying falsified meat products.

A. Bolshakov (✉) · R. Kalimullina
Peter the Great St. Petersburg Polytechnic University, 29, Polytechnicheskaya, St. Petersburg
195251, Russia
e-mail: aabolshakov57@gmail.com

R. Kalimullina
e-mail: krr12.04.98@mail.ru

M. Nikitina
FSBIU “Federal Scientific Center for Food Systems named after V.M. Gorbатов” RAS, 26, st.
Talalikhina, Moscow 109316, Russia
e-mail: nikitinama@yandex.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced
Human-Centered Society*, Studies in Systems, Decision and Control 333,
https://doi.org/10.1007/978-3-030-63563-3_15

Keywords Falsification of meat products · Polychrome image approximation · Optimization · Intelligent system · Genetic algorithm

1 Introduction

There is a direct link between nutritional quality and emerging diseases. Various diseases are initiated by the quality of nutrition. These include, but are not limited to, obesity, cardiovascular disease, diabetes mellitus, cancer, and others. This time dependence is evaluated as a geometric progression. Poultry and livestock products are carried out in accordance with technologies that use artificial metabolic disorders. As a result, an increase in live weight is observed. The value of these products is doubtful, and the breakdown of drugs significantly affects the safety and quality of the products [1–3].

Instead of meat, offal, such as head meat, internal organs, was usually added to meat products. Now, additives that are of plant origin with a carbohydrate and protein nature are mainly used [4, 5]. Existing technologies for the production of sausages and semi-finished products involve the differentiation of components of plant and animal origin, which have little in common with human needs, such as phosphates, carrageenas, confectionery proteins, and other substitutes or imitators of meat [6–8].

The unlimited authority for import substitution of goods provided by the state to domestic producers enables them to supply products with cheap, but not safe for human health substitutes, preservatives, and food additives to the agri-food market [9, 10].

The degree of deterioration in the quality of meat products due to freezing and further storage in this state depends on various aspects. These include conditions, as well as the duration of storage in a state of freezing, as well as the magnitude of biochemical changes that existed before freezing, as well as its speed. The duration of storage of frozen meat depends on the initial quality indicators of meat, type, and fatness of livestock and poultry, as well as storage technology, such as temperature, packing density, stack sizes, etc. [11, 12].

Determining the quality of meat products is very important. The degree of objectivity, as well as the efficiency of monitoring the characteristics of meat products significantly determines the quality of the products that are provided to the consumer. At the same time, the methods used to control the quality of meat as raw materials remain significantly time-consuming, focused on the use of expensive equipment [13–18].

The widespread distribution of counterfeit products in the consumer market is very negative. Relevant is the identification of actually used for the production of meat raw materials and meat products with the determination of the composition and the allocation of the corresponding components that are of animal and vegetable origin [19].

This identification is required to establish the true form, as well as the name of the product, its compliance with the regulatory documentation for the product,

as well as labeling and invoices, etc. To identify the falsification of meat products, microstructural methods for identifying the composition of the product and raw materials are used, which were proposed by the laboratory of microstructural analysis of meat products of the All-Russian Research Institute of Meat Industry. The proposed methods are GOST R when conducting relevant studies [20].

An analysis of the proposed methods shows the presence and necessity of using electronic microscopes, as well as expert technologists for processing the obtained images of slices of meat and meat products during their implementation. The complexity of the procedure associated with image processing, identification of the type of counterfeit determines the relevance of the work aimed at their automation.

In this regard, the urgent task is to automate the process of identifying differentiated counterfeit products in raw meat and meat products. Currently, this process is carried out manually by technologists. Since this is a rather time-consuming task due to the large analytical work, it is advisable to reduce the role of the “human factor” in the process of analysis and decision-making on the presence of counterfeit meat products. To do this, it is necessary to automate the procedure in which the presence of falsification is determined by the input image of slices of meat products according to a number of its characteristics, such as shape, color, size. Further, based on the analysis of these characteristics, the decision support system forms a conclusion. As a similar system, it is proposed to use one of the varieties of intelligent systems—a hybrid expert system. Thus, the original task is divided into two: (1) obtaining and processing a color image of a slice of meat products; (2) the decision on the presence of counterfeit according to the results of processing the image of the slice.

The proposed system is advisable to use in the corresponding cyber-physical automated system for the identification of falsified meat products based on the principle of “digital double.”

2 Description of the Process for Determining the Presence of Meat Falsification

The process of determining the presence of counterfeit meat products is regulated in the Federal laws of the Russian Federation, as well as the relevant State standards (GOSTs), for example. One of the most common methods is based on the so-called histological analysis [20]. Below is a diagram of the most popular histological method for determining the presence of falsification in meat products (see Fig. 1).

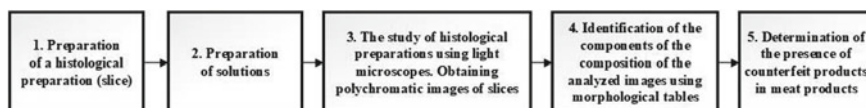


Fig. 1 Stages of the process of determining the presence of falsified meat in meat-based on histological analysis

Determination of the presence and type of counterfeit in meat products is carried out by technologists of sufficiently high qualification on the basis of the study of signs constructed as a result of the previous stages, which are of both quantitative and qualitative nature. Depending on the presence of these signs, their values, technologists determine the presence and type of falsified meat products.

Thus, the initial information for the construction and subsequent analysis of morphological tables are color images of slices of meat products, which are obtained in laboratories from electron microscopes. Therefore, the use of computer tools requires the application of appropriate methods and processing algorithms in the general case of color images.

3 Formulation of the Tasks

It is advisable to automate a rather laborious operation related to the identification and determination of the presence of counterfeit using morphological tables, which are given in the relevant GOSTs. In turn, automation is associated with the processing of images that are obtained by examining slices of meat products using electron microscopes. Thus, the functions of technologists in identification and decision-making must be transferred to the appropriate system to support decision-making on the presence and type of counterfeit meat.

Methods and algorithms have been developed that are used for segmentation and subsequent processing of selected fragments in the original images [21–27]. These include operations related to the selection of the boundaries of regions in the image under study and the determination of their structural characteristics. It should be noted that our analysis of morphological tables showed that in the general case, 8 colors are enough to describe the existing methods for identifying falsified by color. Therefore, it makes sense to optimize the original polychrome image. This will facilitate the procedure for identifying fragments in the studied image, as well as significantly reduce the amount of information for storing and transmitting data about the studied images of meat slices.

One of the conclusions of various existing scientific and practical studies is the recognition of the fact that the selection and development of a new method for solving problems related to digital image processing require quite a laborious testing on a test set of images in the studied subject area. This approach allows us to provide acceptable results based on the characteristics of the studied class of images.

In connection with the foregoing, we can formulate the purpose and objectives of the study. The goal is to increase the effectiveness of the process of determining the presence of counterfeit meat products based on the automation of image processing of sliced meat.

To achieve the goal it is necessary to solve the following tasks: (1) develop an image processing algorithm for automated detection of meat falsification; (2) develop a decision support system to determine if the meat is falsified.

4 Development of an Image Processing Algorithm for Automated Detection of Meat Falsification

4.1 Statement of the Task of Image Processing to Automate the Determination of Meat Falsification

The process of determining the presence of meat falsification is presented in Fig. 1. The proposed automation concerns the last two stages of this process. We carry out a systematic analysis of the decision-making process that is performed at these stages. To do this, consider the circuit in Fig. 2.

The input variable X , representing the image in JPEG format contains information about the pixels, can be represented as a matrix of numbers. When translating, for example, from JPEG to RGB, each pixel is assigned 3 numbers from the range [0; 255].

The output variable Y represents the result of the procedure for identifying the presence of counterfeit on the basis of morphological tables containing a description of fragments of the image structure and their characteristics. The variable Y has a symbolic value that corresponds to the identification result, i.e. “Absence of counterfeit” or the name of the counterfeit.

The variable U includes the data necessary for the “adjustment” of the procedure for identifying falsified meat products related to image processing, segmentation, isolation and determination of structural elements, values of their characteristics, etc.

The process is as follows. The input is an image of a slice of meat X , which is then processed. Image processing X , in General, includes the following steps: preliminary processing of images of slices (noise removal, palette optimization, etc.); color segmentation based on minimizing the palette; approximation of the boundaries of the areas selected in the image; determining the size of areas; determination of

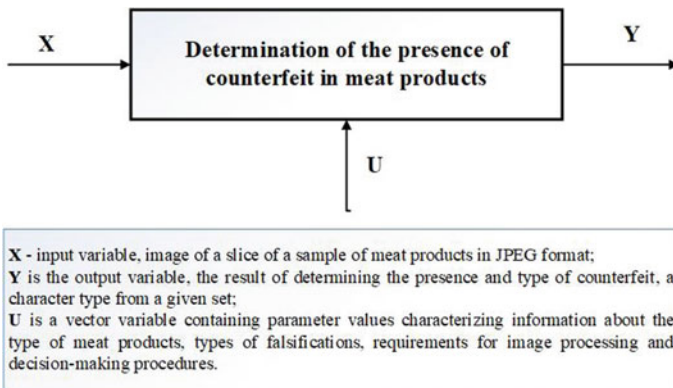


Fig. 2 Determination of the presence of counterfeit in meat products

the shape of particles; determination of the color of particles; identification of the presence of falsifications; conclusion of the results of determining the presence of falsified meat.

The consistent implementation of the above steps allows you to get an informed decision about the presence or absence of certain types of falsifications of meat products.

The next section proposes a generalized algorithm for processing images of slices of meat products, which can be implemented in the corresponding automated system.

4.2 Development of an Image Processing Algorithm to Automate the Detection of Meat Falsification

In accordance with the scheme in Fig. 2 in the previous subsection and the required stages of processing images of samples of meat products, the following algorithm for their implementation is proposed, which is presented in Fig. 3.

The functioning of the algorithm in accordance with the proposed block diagram in Fig. 3 is as follows.

Block “Cycle while” study of the samples is completed. A cycle is organized until the planned samples of meat products are examined.

Block “Image X input, procedure settings”. A slice image of meat products in JPEG format is entered, as well as parameter values that are necessary for the appropriate adjustment of the image processing procedures, segmentation, etc.

Block “Pre-processing images of slices.” The image is being converted from the JPEG format to the RGB format, which is more convenient for further digital processing. Then, the color palette of the polychrome image of the sample slice is minimized [26].

Block “Color segmentation based on minimizing the palette.” Implements the coloring of the pixels of the source image in the colors of the optimized palette, based on the proximity condition of the source and selected from the optimized color palette.

Block “Approximation of the boundaries of the selected areas.” The identification of lines and corresponding areas in the image of the slice of meat products using known algorithms and methods [21–23].

Block “Sizing of areas”. Based on the results of the previous block, the corresponding sizes are calculated for the detected areas.

Block “Determining the shape of particles.” Particles are detected in the image of the slice and their shape is determined using well-known algorithms and methods [24, 25].

Block “Particle color determination”. For the particles detected and the shape of the particles identified by the results of the work, their color is determined.

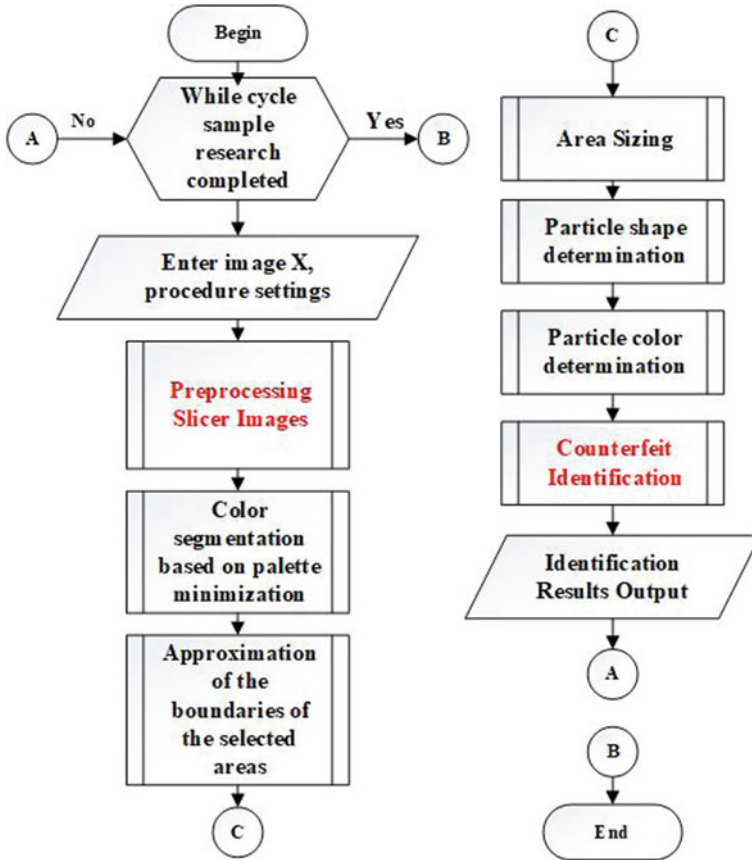


Fig. 3 Block diagram of the image processing algorithm for automating the detection of meat falsification

Block “Identification of the presence of falsifications.” This block identifies the presence or absence of counterfeit meat products. In the first case, the type of falsification is also determined using morphological tables. The data are of a qualitative nature, which requires the use of appropriate methods, primarily intellectual. Intelligent methods allow formalizing high-quality information, which is necessary for implementation in automated control systems.

Block “Output identification results.” The output obtained from the identification of the presence of falsified meat products is printed and/or screened, as well as in the database for storage and subsequent use.

5 Development of a Decision Support System for Determining the Presence of Meat Falsification

5.1 Statement of the Task of Developing a Decision Support System for Determining the Presence of Meat Falsification

The developed decision support system should have the necessary functions to automate the process of identifying counterfeit products in meat products.

These functions include the following: input of the necessary information (image slice samples in JPEG format, the values of the settings' characteristics of internal procedures); conclusion of identification results on the presence of counterfeit; image processing of slices to remove noise, optimize palettes, etc.; color segmentation based on the optimization of the color palette; approximation of the boundaries of the areas selected in the image; determining the size of areas; determination of the shape of particles; determination of the color of particles; identification of falsifications based on formalized knowledge.

Morphological tables contain a description of the signs of falsification and a description of their characteristics and values. In the general case, these characteristics are of a qualitative nature. Therefore, to formalize this information, it is advisable to use appropriate intellectual methods to formalize qualitative data and synthesize solutions based on them [26].

From the above list, it follows that the automation task is very complex, therefore, its solution should be divided into a number of stages. Of greatest interest to us are the tasks of preliminary processing of images of slices of samples of meat products, as well as the identification of the presence of falsifications based on formalized knowledge.

The solution to the first problem allows you to prepare information about the slice in the form that is most suitable for the remaining tasks, to ensure their successful implementation, because this eliminates noise, redundant information. The resulting image contains information about the most characteristic elements of the desired elements, areas associated with the possible presence of counterfeit. To determine a suitable truncated color palette, it is proposed to solve the corresponding optimization problem. Moreover, as a criterion for the solution, it is proposed to formulate a target function based on the degree of proximity of pixels in the RGB format.

The solution to the second problem is connected, as indicated above, with the formalization using the production rules of the information contained in the morphological tables described in the relevant regulatory documents. The decision-making procedure uses information about values that accept signs of fraud. With the proposed method, this information is introduced into the decision support system (DSS) by experts (technologists) based on the results of solving the first problem. It is also possible, if necessary when the original image is used. The decision result is displayed

by the DSS for the user in an understandable form for him: the wording about the absence or presence of a certain falsification in the sample of meat products.

The above two problems are solved in this qualification work. The remaining tasks are the subject of further research. It should be noted that there is a fairly large arsenal of methods and algorithms for their solutions. A lot of resources are needed, first of all, temporary for comparative testing of existing approaches and identifying the most suitable. It also requires substantiating the feasibility of developing our methods and algorithms.

5.2 Building the Architecture of a Decision Support System for Determining the Presence of Meat Falsification

The proposed architecture of the decision support system for identifying the presence of falsification in samples of meat products (see Fig. 4) ensures the fulfillment of the functions that the DSS should fulfill. These functions are described above.

The proposed option includes a control module that controls the DSS subsystems, organizes interaction using the appropriate interface, communication with the administrator and authorized user. The system includes an information subsystem in the form of ES, as well as subsystems for PPR.

Here, ES is implied in the classical form and includes standard blocks: A knowledge base of production type; Inference mechanism; Working memory; Explanatory component.

The subsystem for supporting decision-making about the presence of falsification in samples of meat products for processing sliced images includes the following modules (subprograms): image processing of slices; color segmentation based on the optimization of the color palette; approximation of the boundaries of the areas selected in the image; determining the size of areas; determination of the shape of particles; particle color determination.

The functioning of the DSS is carried out, as indicated, under the control of the control module, which organizes the necessary interaction of the subsystems, information support, and two groups of users on the basis of the corresponding support. These include the administrator and operator (user, technologist, knowledge engineer).

The administrator manages the password system to differentiate access rights for different categories of users, enters, and corrects data in the database.

The user carries out the process of identifying the presence of counterfeit by inputting images in the iterative mode. In the process of this work, an appeal is made to the modules (subprograms) of the preliminary processing of the initial image of the slice, segmentation, etc.

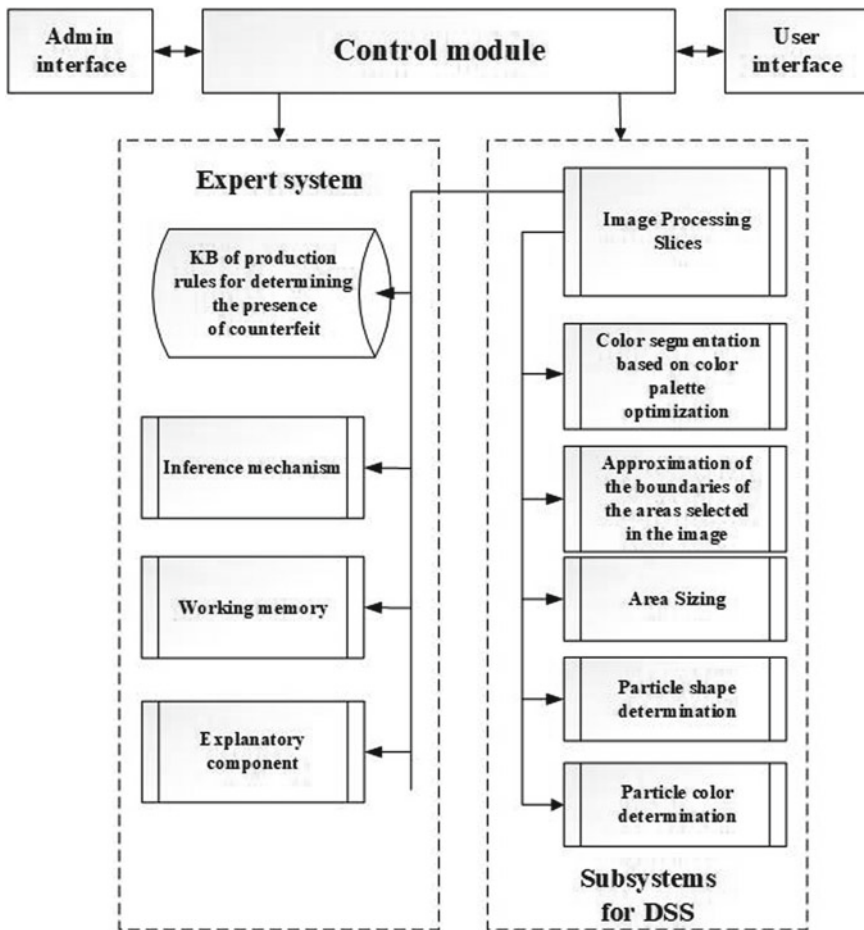


Fig. 4 The architecture of the decision support system for determining the presence of meat falsification

5.3 Development of a Prototype Expert System for Determining the Presence of Meat Falsification in Prolog

The methodology is proposed below and the results of developing a prototype expert system for DSS are described by the scheme shown in Fig. 4.

To develop a prototype, it is advisable to use the methodology of building expert systems [28–32]. According to this methodology, ES synthesis is carried out by the following steps.

1. Statement of the problem. The purpose of the construction of ES is to increase the effectiveness of determining the presence of falsification in samples of meat products.
2. Identification of the problem. This process is carried out by technologists (experts) based on a comparison of the data obtained after analysis of the samples prepared according to the instructive materials using electronic microscopes and morphological tables containing information on the correspondence of the revealed signs to the presence of certain falsifications. To reduce time as well as financial costs, it is proposed to develop an ES in which the knowledge presented in the relevant instructional materials, as well as, possibly, from expert technologists, is formalized.
3. Extraction of knowledge. As sources of knowledge, instructive documents, scientific publications [11–20], and experts were used. Systematization of this material made it possible to present information in the form of tables called morphological. The rows of these tables correspond to the characteristics, the content of which is established on the basis of visual analysis of images of slices of meat products prepared in special solutions, which are examined using electronic microscopes. Pillars—to certain types of counterfeit.
4. Conceptualization. At this stage, key concepts, relationships, and characteristics are identified that are necessary to describe the process of identifying the presence of a certain type of counterfeit. As such concepts, for example, are the identification features of plant components of protein nature: particle shape; particle size; the presence of fragments of the shell of soybean; tinctorial properties (with hematoxylin and eosin stain).

Depending on the detected values of the above characteristics, the following counterfeit variants are established: soy protein isolated; soybean concentrate; textured soy protein product.

Thus, information characterizes a limited set of attributes of a qualitative nature: the shape and size of particles, the presence of fragments, the coloring of particles and elements, etc. Given a given set of values for these characteristics, it is concluded that there is a certain counterfeit. If they are absent, a conclusion is drawn about the absence of counterfeit.

5. Formalization. In accordance with the above, we introduce the following values of the variables shown in Table 1.

The characteristics indicated in Table 1, corresponding to key concepts, can take on certain syntactic values, which are considered as constants. If the sign takes this value, then it is considered that the corresponding variable is equal to this value.

For example, according to Table 2, if it is revealed by the “Particle Shape” (A1) that it is observed that the shape is in the form of a rolled cord or bean, or rounded from a dark point in the center (C1), then $A1 = C1$ is considered.

Further, the assignment of the conditions for the identification of a certain counterfeit is carried out based on the rules of the form “IF...TO.” In our case:

IF $A1 = C1$ and $A2 = B1$ and $A3 = D1$ and $A4 = E1$ THEN $A5 = F1$

Table 1 Description of key concepts

Key concept	Designation	Characteristic
Particle shape	A1	Introduces a user or slicer image processing system
Size	A2	Introduces a user or slicer image processing system
Tinctorial properties (when stained with Lugol's solution)	A3	Introduces a user or slicer image processing system
Tinctorial properties (when stained with hematoxylin and eosin)	A4	Introduces a user or slicer image processing system
Identification of the presence of counterfeit	A5	ES displayed

Table 2 Description of the values adopted by key concepts

A1	C1: {(a) folded harness (b) bean (c) rounded with a dark dot in the center}	C2: {Rounded particles combined into large aggregates}	C3: {Particles are heterogeneous and irregular in shape}	C4: {Particles are irregular, more uniform}	C5: {Separate plant cells or groups of cells. Each cell is surrounded by a clear, non-stained cellulose membrane}
...
A5	F1: {Starch}	F2: {Flour}	F3: {Carrageenan Semi-Purified}	F4: {Carrageenan Peeled}	F5: {Gums of guar, packaging and carob}

IF A1 = C2 and A2 = B2 and A3 = D1 and A4 = E1 THEN A5 = F2

IF A1 = C3 and A2 = B3 and A3 = D3 and A4 = E2 THEN A5 = F3

IF A1 = C4 and A2 = B3 and A3 = D2 and A4 = E2 THEN A5 = F4

IF A1 = C5 and A2 = B4 and A3 = D3 and A4 = E3 THEN A5 = F5

Thus, the information contained in the morphological table is presented in the form of production rules.

6. Execution. For a software implementation, the Prolog language of artificial intelligence was chosen, which allows the construction of production rules to be implemented and is widely used to create ES of relatively small sizes. Further studies on the development of the prototype and the creation of a full-fledged ES to determine the presence of counterfeit suggest formalization of knowledge of their various types according to the method described above. Further, tests involving experts are required.

6 Development of a Method for Optimizing the Approximation of a Polychrome Image

6.1 *Statement of the Problem of Optimizing a Polychrome Image of a Slice of Meat Products*

To make a decision about the presence of falsification in a meat product, it is necessary to identify the corresponding area in the image of the slice. To reduce the amount of memory occupied by the image and simplify its processing, it is advisable to reduce the number of colors in which this image is filled. Further, a set of such colors is called an image palette.

The image of the histological preparation arriving for processing has the JPEG format. The color model used is RGB. Thus, the color of each pixel of the image X_{ij} is determined by a tuple of integers $[R_{ij}, G_{ij}, B_{ij}]$ in the range $[0 \dots 255]$, then these 3 numbers are called defining.

According to the National Standards for Determining the Presence of Fake, soy protein products and pea flour are stained with a solution of hematoxylin and eosin. Signs of falsification of meat products are: rounded pink particles; cylindrical or rounded particles in shades of red; red or purple bundles of fibers; unpainted particles.

Thus, to solve the recognition problem associated with determining the presence of a counterfeit, enough information about the presence in the image of fragments of several colors. Therefore, to reduce the amount of information when transmitting and storing images, it is necessary to determine the corresponding color palette, in which the number of colors is relatively small.

6.2 *Formalization of the Task Statement*

It is required to optimize the number of colors in a polychrome image to a given N , i.e. determine the value of the vector P_N , where N is the size of the palette. It is advisable to consider the following set of values of the N variable $C = \{2, 4, 8\}$ as possible values of the size of the palette, which is associated with the analysis of a priori information about the number of falsified flowers and, in fact, meat after appropriate processing of the slices with chemical reagents. Thus, based on the original polychrome image, it is required to obtain an optimized image that uses a given number of colors. In this case, all the colors of the original image are replaced with colors from the palette, based on the criterion of the proximity of the original and optimized images.

Input data: image of the histological preparation in the form of a matrix $X_{ij} = [R_{ij}, G_{ij}, B_{ij}]$, $i = 1, \dots, W$, $j = 1, \dots, H$, where W and H are the width and height of the image in pixels. We call such a matrix defining; many RGB color values that could

potentially be present as fake; the number of colors N used in the resulting image; the maximum allowable time of the algorithm (number of iterations).

Output: vector P_N containing the resulting image palette.

The task is optimization, the essence of which is to build a palette that maximally approximates the original image. Determining the optimality of the palette, i.e. the degree of its approximation in accordance with the criterion for assessing the quality of the approximation, a description of the choice of which is given below.

6.3 Criteria for Assessing the Quality of Polychrome Image Approximation

The choice of a criterion for numerical quality assessment is very important since it is the main characteristic of evaluating the effectiveness of the algorithm. Since the selection of the resulting image, which should consist of a limited number of colors, i.e. 4 is a search and optimization task, then the criterion is necessary to determine the correct search direction and optimize the corresponding deviation according to the selected quality assessment criterion. The choice is limited to a modular criteria-based assessment.

Thus, to assess the quality, the operation of calculating the total (absolute) deviation modulus was used:

$$\Delta Q = \sum_{i=0}^W \sum_{j=0}^H |R_{ij}^X - R_{ij}^P| + |G_{ij}^X - G_{ij}^P| + |B_{ij}^X - B_{ij}^P|,$$

where $R_{ij}^X, G_{ij}^X, B_{ij}^X$ —determining the pixel numbers of the original image, and $R_{ij}^P, G_{ij}^P, B_{ij}^P$ —defining numbers of one of the values of the vector P .

The criterion contains the functions of taking the module; therefore, its differentiability in the Frechet sense cannot be guaranteed, as well as continuity. Therefore, it is not possible to use gradient extremum search methods. The use of exhaustive search methods and procedures for minimizing the zero level does not guarantee the optimal value, since the problem belongs to the number of NP -complete, i.e. with an exponentially increasing level of difficulty.

Therefore, it is advisable to use heuristic methods, which, in particular, include genetic algorithms.

In the next paragraph, based on an analytical review, a reasonable choice of simple genetic algorithms from among the considered methods is made.

6.4 The Choice of a Method for Solving the Problem of Minimizing the Polychrome Image Palette

The task of reducing the polychrome image palette is reduced in the previous subsection to the optimization one. As a criterion of optimization, a modular assessment is selected. The proposed objective function is not differentiable in the sense of Frechet; therefore, the use of gradient methods to solve the formulated search and optimization problem is impossible. In general, the finite set of solutions, that is, all possible combinations of $256 * 3$ colors from the original palette, is too large to use the direct enumeration method. Therefore, this method of solution is very costly and does not guarantee the determination of the optimum. Therefore, to solve the proposed optimization problem, it is advisable to choose a heuristic algorithm.

A simple genetic algorithm was selected for the following reasons: (1) evolutionary programming does not use the recombination operator, which is required in the problem being solved; (2) swarm intelligence is mainly used in finding the optimal path, for calculating routes using graphs and for determining approximate solutions in the “traveling salesman problems”; (3) for the task of combinatorial type, the basic mechanisms of the genetic algorithm are sufficient, without additional information about the colony, as in swarm intelligence.

6.5 Construction of an Algorithm for Solving the Problem of Minimizing the Palette of a Polychrome Image of a Slice of Meat Products

The algorithm solves the optimization problem of search and allows you to form the optimal individual (palette).

An individual is the basic unit of the evolutionary process; in the context of this task, an individual is an image palette.

Each individual consists of two chromosomes. Moreover, the Alpha chromosome corresponds to the initial approximation of falsified colors specified by the user. Here, the Beta chromosome corresponds to the colors obtained during the algorithm without taking into account the initial approximation.

Each chromosome consists of genes that determine the genotype of an individual. A gene is a unit of hereditary information and an internal representation of an alternative solution. In the proposed algorithm, each gene contains a three-dimensional vector containing numbers in the range from 0 to 255, which corresponds to the color components in the RGB format.

The result of each iteration of the genetic algorithm is a population that is a collection of individuals.

The main stages of the method based on a genetic algorithm.

1. The formation of the initial population. The first stage of the algorithm is the formation of an initial population containing N_I individuals. To form the alpha chromosome, N_{alpha} , the color of the initial falsification approximation is randomly selected once, i.e. one of the colors that is specified by the user as the alleged color of falsification. For the formation of beta chromosomes, we can use two approaches. In the first case, N_{beta} randomly selects an element of the defining matrix, i.e. the color of the corresponding pixel, and its value is assigned to one of the genes of the beta chromosome. In the second case, a random color is selected for each beta chromosome gene.
2. Breeding. The idea of the genetic algorithm is to preserve the genotype of the best individuals in each of the newly formed generations and to transfer N_S of the most optimal individuals to the next generation at each iteration. To do this, it is proposed to introduce the following selection operator:

$$O_s(I) = \sum_{i=1}^h \sum_{j=1}^w (p_{i,j} - c(p_{i,j})), \quad c(p_{i,j}) = \min(p_{i,j} - c_k),$$

where I is the individual, h is the height of the defining matrix, w is the width of the defining matrix, $p_{i,j}$ is the element of the defining matrix, $c(p_{i,j})$ is the color closest to $p_{i,j}$ from the palette, c_k —color of the palette.

Thus, the individual I_i is “more optimal” than the individual I_k if $O_s(I_i) < O_s(I_k)$.

The remaining space in the new generation equal to $N_I - N_S$ is filled with individuals obtained as a result of crossing and mutation.

3. Crossbreeding. The result of the operation of crossing two individuals $O_C(I_1, I_2)$ is their offspring containing the characteristics of both parents. The crossing process is as follows: at the first parent I_1 random sections of the alpha chromosome and beta chromosome are selected. This site passes into the descendant, and the remaining space is filled with the corresponding genes of the second parent I_2 .
4. Mutation. For the convergence of the algorithm and the prevention of premature stabilization, it is very important that the diversity of individuals in each new generation is maintained. For this, it is proposed to introduce a mutation operator $O_M(I)$. This operator replaces the gene values in the alpha chromosome N_{Malpha} with others from the initial falsification approximation and in the N_{Mbeta} beta chromosome gene values with another randomly selected color.
5. Checking the stop condition. The condition for stopping the algorithm is to pass a given number of iterations $N_{i_{max}}$. The result of the algorithm is the most optimal individual obtained during all iterations of the algorithm.

6.6 Testing the Algorithm for Solving the Problem of Minimizing the Palette of a Polychrome Image of a Slice of Meat Products. Results Analysis

The program for minimizing the palette of a polychrome image of a slice of meat products is implemented in the C++ programming language. The program is generally intended to implement the procedure of minimizing the palette of a polychrome image specified in RGB format. The source data is presented in the form of a matrix of a given size with cells containing information about the image pixels. In this case, the color is represented using the RGB model. To assess the quality, a criterion is used based on the calculation of the total absolute module of deviations of the three-dimensional vector RGB of the original image from the approximating one. The program implements a genetic algorithm in which the genome is a pixel. The pixel is encoded in RGB format, i.e. is determined by three positive numbers in the range from 0 to 255. The result of the program is to determine the quantitative values of the approximating RGB palette that minimizes the initial polychrome image palette specified in the RGB format.

The proposed method for approximating a polychrome image of slices of meat products based on the use of a limited palette is reduced to solving the optimization problem. As shown above, the use of a genetic algorithm is most preferred. The advantage of GA is non-criticality to the properties of the objective function: continuity, differentiability, unimodality. In addition, it is relatively easy to take into account the constraints that form the region of feasible decision values. However, this requires studies related to the choice of the parameters of the GA, the study of the rate of convergence depending on the initial conditions and these parameters, as well as the presence of local extrema.

In connection with this, a technique is proposed that allows one to determine the most suitable values of the parameters of the proposed optimization method based on GA.

The initial information for the methodology contains test images with a given type of counterfeit and without its presence, a set of acceptable values of the algorithm parameters.

The output contains systematic data on the results of optimization according to the following criteria: values of the target (fitness) function, operating time (number of iterations).

It is proposed to vary the values of the following characteristics.

1. The variant of formation of the initial population:
 - (a) Random selection of elements of the defining matrix
 - (b) Random color generation
2. The number of selected individuals for selection:
 - (a) $N_1 \div 2$
 - (b) $N_1 \div 4$

Table 3 Optimization results for approximating a polychrome image based on GA

tmax	topt	penalty	iteration
------	------	---------	-----------

3. The number of mutated genes $N_{M_{\alpha}}$ и $N_{M_{\beta}}$:
 - (a) $N_{M_{\alpha}} = N_{\alpha} \div 2$, $N_{M_{\beta}} = N_{\beta} \div 2$
 - (b) $N_{M_{\alpha}} = N_{\alpha} \div 4$, $N_{M_{\beta}} = N_{\beta} \div 4$
4. The number of iterations.
 - (a) $N_{i_{\max}} = 1000$
 - (b) $N_{i_{\max}} = 2000$
 - (c) $N_{i_{\max}} = 4000$

Thus, the technique represents an iterative procedure, each nested loop of which is represented by the steps described above.

The results are summarized in tables of the following form (see Table 3).

Here t_{\max} is the specified maximum time to search for the optimal value; t_{opt} —time for which the optimal value is determined, sec.; *iteration*—the number of iterations that needed to be performed to find the optimal value for the formation of the best individual GA; *penalty*—a fitness function for assessing the quality of polychrome image approximation.

The proposed polychrome image approximation algorithm is aimed at reducing the color palette used. For computational experiments, the value of the size of the resulting image palette is further selected $C = 4$.

Moreover, the values of other parameters of the genetic algorithm are fixed and the following values are selected: the number of individuals in the population of 20; the number of selected individuals for selection $n_i = 2$; the number of mutated genes $N_{M_{\beta}} = 2$.

In the study of the genetic algorithm, the formation of the initial population changes, the following parameters vary:

1. Alpha chromosome used $N_{\alpha} = 1$, those each individual has the color of one counterfeit, which could potentially be contained in a slice. The beta chromosome uses only elements of the defining matrix.
2. Only the beta chromosome is used, the gene values of which are initialized by the elements of the determining matrix.
3. Only the beta chromosome is used, the gene values of which are randomly selected.
4. Only the beta chromosome is used, the gene values of which are “central values” ($R = G = B = 128$).
5. Only the beta chromosome is used, the gene values of which are the extreme values of the three components of the RGB model, namely $R = G = B = 0$.
6. Only the beta chromosome is used, the gene values of which are the extreme values of the three components of the RGB model, namely $R = G = B = 255$.

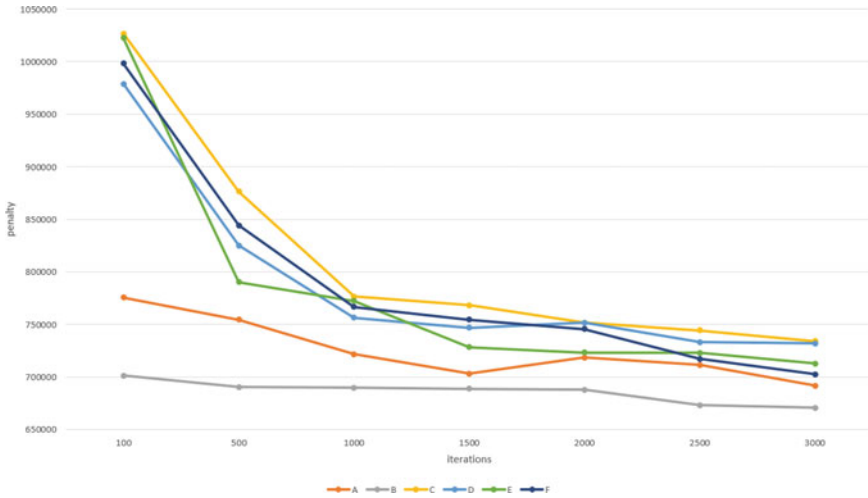


Fig. 5 Dependence of the value of the objective function on the number of iterations

For each formation of the initial population, a graph is constructed of the dependence of the objective function on the maximum allowable time of the algorithm (number of iterations). The graphs are presented in Fig. 5, 10 experiments were performed for each number of iterations, and the average value was calculated.

Based on the results obtained, the following conclusions can be drawn.

After the first thousand iterations, the convergence rate of the algorithm drops significantly, which makes further work of the algorithm impractical.

The best approximation was obtained by initializing the initial population with the colors present in the image (option B), however, the convergence rate is extremely low, which is explained by the small palette of the original image, i.e. the initial population always consists of a limited number of flowers. Because of the mutation, changes occur in the genes of an individual, however, the value of the objective function of a given individual is a priori greater than the initial ones, because the colors of the original palette are closer.

The best convergence rate is shown with a random selection of colors for the formation of the initial population (option C). However, this option shows the worst approximation, which seems logical enough.

With the above parameters, for each item in the list of initial population formation, the results of the algorithm for the number of iterations $N_i = 3000$ are visually demonstrated below (see Figs. 6 and 7).

Figure 7 shows the best result of the algorithm for other histological preparations.

Since the graph of the dependence of the objective function value on the number of iterations showed that when choosing the formation of the initial population of variant B, the *penalty* parameter is the smallest, this initialization is selected. Also, according to the results of the study, the number of iterations $N_i = 1000$ was selected.

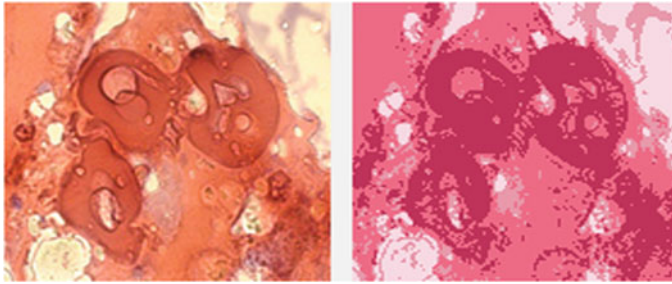


Fig. 6 Formation of the initial population option C ($penalty = 732030$)

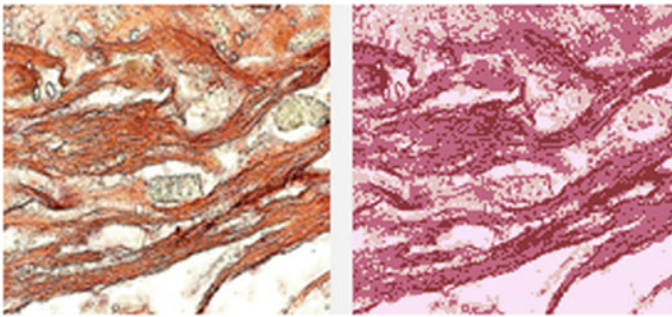


Fig. 7 Formation of the initial population option F ($penalty = 830288$)

7 Conclusions

A systematic analysis of the process of identifying the presence of counterfeit products in meat products was carried out, the last two stages that are expedient to automate to achieve this goal are highlighted. The purpose of the work was to increase the effectiveness of the process of determining the presence of meat falsification through automation. A generalized algorithm is developed for determining the presence of counterfeit based on the analysis of images of sections of product samples. The following tasks were selected: preliminary processing of images of slices of samples of meat products, as well as identification of the presence of falsifications based on formalized knowledge.

To automate the determination of counterfeit availability, a generalized algorithm and architecture of a decision support system are proposed, which includes two subsystems: expert and decision support. Formalization of knowledge for making a decision about the presence of counterfeit is carried out on the basis of production rules, which are generated on the basis of information contained in morphological tables. The development of a prototype ES is made in the programming language of artificial intelligence Prolog. To build an ES for identifying the presence

of counterfeit, a technique is proposed based on the methodology for building expert systems.

The second solved problem is to optimize the approximation of a polychrome image. The analysis of the features of this problem is carried out, which made it possible to propose a method for solving it based on the genetic algorithm. A program was developed in the C++ programming language, using which a complex of computational experiments was carried out to study the nature of the dependence of the convergence of the optimization process. Different parameters and characteristics varied. Among them: the maximum number of iterations, the choice of the initial population, etc.

In the future, to determine the most significant parameters, it is advisable to develop an appropriate technique using the statistical hypothesis, methods of multidimensional processing of experimental data.

Further research in this direction is associated with the full implementation of all the necessary image processing algorithms for the automatic detection of the required values of the characteristics of fragments of images that are counterfeit. Another area of work is the formalization of knowledge in the corresponding knowledge base of the product type, which will allow the appropriate decision-making support system to be formed, as well as to explain the conclusions on the presence of detected falsifications based on the processing of images of sliced meat products. It is advisable to use the results obtained when constructing typical automated systems for identifying the presence of falsification in samples of meat products for use in promising cyber-physical systems in the studied subject area. In addition, it is recommended that algorithms and programs be used to study the tissues of living creatures and humans when conducting studies aimed at identifying abnormal phenomena.

The chapter was prepared based on the results of the project with the support of the RFBR grant No. 18-08-01178\20.

References

1. Surkov, I.V., Kantere, V.M., Motovilov, K.Y., Renzyaeva, T.V.: The development of an integrated management system to ensure the quality stability and food safety. *Foods Raw Mater.* **3**(1), 111–119 (2015)
2. Xu, C., Tang, X., Shao, H., Wang, H.: Salinity tolerance mechanism of economic halophytes from physiological to molecular hierarchy for improving food quality. *Curr. Genomics* **17**(3), 207–214 (2016)
3. Tseng, S.-Y., Li, S.-Y., Yi, S.-Y., Sun, A.Y., Gao, D.-Y., Wan, D.: Food quality monitor: paper-based plasmonic sensors prepared through reversal nanoimprinting for rapid detection of biogenic amine odorants. *ACS Appl. Mater. Interfaces.* **9**(20), 17306–17316 (2017)
4. Kanareykina, S.G., Kanareykin, V.I., Ganieva, E.S., Burakovskaya, N.V., Shadrin, M.A., Halepo, O., Babaeva, M.V., Nikolaeva, N.V., Voskanyan, O.S.: The structure development of yogurt with vegetable ingredients. *Inter. J. Recent Technol. Eng.* **8**(2), 1587–1592 (2019)
5. Gupta, A.J., Wierenga, P.A., Gruppen, H., Boots, J.-W.: Influence of protein and carbohydrate contents of soy protein hydrolysates on cell density and igg production in animal cell cultures. *Biotechnol. Prog.* **31**(5), 1396–1405 (2015)

6. Wang, Q., Zhang, J.: Research status, opportunities and challenges of high moisture extrusion technology. *J. Chin. Inst. Food Sci. Technol.* **18**(7), 1–9 (2018)
7. Pateiro, M., Domínguez, R., Gómez, B., Lorenzo, J.M., Barba, F.J., Sant’Ana, A.S., Mousavi Khaneghah, A., Gavahian, M.: Essential oils as natural additives to prevent oxidation reactions in meat and meat products: a review. *Food Res. Int.* **113**, 156–166 (2018)
8. Hao, J., Liang, G., Li, A., Man, Y., Jin, X., Pan, L.: Review on sensing detection progress of “lean meat agent” based on functional nanomaterials. *Nongye Gongcheng Xuebao* **35**(18), 255–266 (2019)
9. Kancheva, V.D., Angelova, S.E.: Synergistic effects of antioxidant compositions during inhibited lipid autoxidation. *Lipid Peroxidation: Inhibition, Effects and Mechanisms* (2016)
10. Loutfi, A., Coradeschi, S., Mani, G.K., Shankar, P., Rayappan, J.B.B.: Electronic noses for food quality: a review. *J. Food Eng.* **144**, 103–111 (2015)
11. Faridnia, F., Bremer, P.J., Oey, I., Ma, Q.L., Hamid, N., Burritt, D.J.: Effect of freezing as pre-treatment prior to pulsed electric field processing on quality traits of beef muscles. *Innovative Food Sci. Emerg. Technol.* **29**, 31–40 (2015)
12. Shenoy, P., Ahrné, L., Fitzpatrick, J., Viau, M., Tammel, K., Innings, F.: Effect of powder densities, particle size and shape on mixture quality of binary food powder mixtures. *Powder Technol.* **272**, 165–172 (2015)
13. Mayer-Scholl, A., Gayda, J., Thaben, N., Bahn, P., Nöckler, K., Pozio, E.: Magnetic stirrer method for the detection of trichinella larvae in muscle samples. *J. Visualized Exp.* **121**, e55354 (2017)
14. Okulakrishnan, P., Kumar, R.R., Sharma, B.D., Mendiratta, S.K., Malav, O., Sharma, D.: Determination of sex origin of meat and meat products on the dna basis: a review. *Crit. Rev. Food Sci. Nutr.* **55**(10), 1303–1314 (2015)
15. Tian, Y., Zhang, J., Chen, Y., Li, X., Cheng, H.: Applications of mass spectrometry-based proteomics in food authentication and quality identification. *Se pu* **36**(7), 588–598 (2018)
16. Duan, X.-Y., Feng, X.-S., Zhang, Y., Yan, J.-Q., Zhou, Y., Li, G.-H.: Progress in pretreatment and analysis of cephalosporins: an update since 2005. *Critical Reviews in Analytical Chemistry* (2019)
17. Chernukha, I.M., Vostrikova, N.L., Khvostov, D.V., Zvereva, E.A., Taranova, N.A., Zherdev, A.V.: Methods of identification of muscle tissue in meat products. Prerequisites for creating a multi-level control system. *Theory Pract. Meat Process.* **4**(3), 32–40 (2019)
18. Tedtova, V.V., Temiraev, R.B., Kononenko, S.I., Tukfatulin, G.S., Kozyrev, AKh, Gazzaeva, M.S.: Effect of different doses of non-genetically modified soybean on biological and productive properties of pigs and consumer characteristics of pork. *J. Pharm. Sci. Res.* **9**(12), 2405–2409 (2017)
19. Tamakhina, A.Y., Kozhokov, M.K.: Biosecurity and methods of falsification of meat products. [Izvestiya Kabardino-Balkarskogo gosudarstvennogo agrarnogo universiteta im. V.M. Kokova] **2**(16), 53–58 (2017) (In Russian)
20. Nikitina, M.A., Chernukha, I.M., Pchelkina, V.A.: Artificial neural network technologies as a tool to histological preparation analysis. In: *IOP Conference Series: Earth and Environmental Science 60. “60th International Meat Industry Conference, MEATCON 2019”*, p. 012087 (2019)
21. Kong, Z., Li, T., Xu, S., Luo, J.: Automatic tissue image segmentation based on image processing and deep learning. *J. Healthc. Eng.* **2019**, 2912458 (2019)
22. Sadhana, B., Nayak, R.S., Shilpa, B.: Comparison of image restoration and segmentation of the image using neural network. *Adv. Intell. Syst. Comput.* **436**, 951–963 (2016)
23. Javanmardi, M., Tasdizen, T.: Domain adaptation for biomedical image segmentation using adversarial training. In: *Proceedings—International Symposium on Biomedical Imaging*, pp. 554–558 (2018)
24. Jiang, X., Yang, X., Ying, Z., Zhan, g L., Pan, J., Chen, S.: Segmentation of shallow scratches image using an improved multi-scale line detection approach. *Multimedia Tools Appl.* **78**(1), 1053–1066 (2019)

25. Dhal, K.G., Das, A., Ray, S., Gálvez, J., Das, S.: Nature-inspired optimization algorithms and their application in multi-thresholding image segmentation. In: *Archives of Computational Methods in Engineering* (2019)
26. Gaiduk, A.R., Neydorf, R.A., Kudinov, N.V.: Application of cut-glué approximation in analytical solution of the problem of nonlinear control design. In: *Cyber-Physical Systems: Industry 4.0 Challenges. Studies in Systems, Decision and Control*, vol 260, pp. 117–132. Springer Nature Switzerland AG 2020. ISSN 2198-4182, ISSN 2198-4190 (electronic). ISBN 978-3-030-32647-0, ISBN 978-3-030-32648-7 (eBook). https://doi.org/10.1007/978-3-030-32648-7_19
27. Dykin, V.S., Musatov, V.Y., Varezchnikov, A.S., Bolshakov, A.A., Sysyoev, V.V.: Application of genetic algorithm to configure artificial neural network for processing a vector multisensor array signal. In: *International Siberian Conference on Control and Communications, SIBCON*, pp. 719–722. <https://doi.org/10.1109/sibcon.2015.7147049>. ISBN: 978-147997102-2 (2015)
28. Algorithmic Intelligence. Towards an Algorithmic Foundation for Artificial Intelligence In: *Series: Artificial Intelligence: Foundations, Theory, and Algorithms* Edelkamp, Stefan. Spinger. ISSN 2365-3051 (2020)
29. Rybina, G.V., Rybin, V.M., Blokhin, Y.M., Sergienko, E.S.: Intelligent technology for integrated expert systems construction. *Adv. Intell. Syst. Comput.* **451**, 187–197 (2016)
30. Bolshakov, A.A., Veshneva, I.V., Chistyakova, T.B.: The architecture of intellectual system for monitoring of university students competences formation process. In: *2016 International Conference on Actual Problems of Electron Devices Engineering (APEDE 2016): Conference*, vol. 2, pp. 30–37. <https://doi.org/10.1109/apede.2016.7878971>. ISBN: 978-150901712-6 (2016)
31. Rybina, G.V., Blokhin, Y.M., Tarakhyan, L.S.: Some approaches to implementation of intelligent planning and control of the prototyping of integrated expert systems. *Commun. Comput. Inf. Sci.* **934**, 145–151 (2018)
32. Bolshakov, A., Kulik, A., Sergushov, I., Scripal E.: Decision support algorithm for parrying the threat of an accident. In: *Cyber-Physical Systems: Industry 4.0 Challenges. Studies in Systems, Decision and Control*, vol. 260, pp. 237–247. Springer Nature Switzerland AG 2020. ISSN 2198-4182, ISSN 2198-4190 (electronic). ISBN 978-3-030-32647-0, ISBN 978-3-030-32648-7 (eBook). https://doi.org/10.1007/978-3-030-32648-7_19 (2020)

Society 5.0: Healthcare Smart Technology

Probability-Entropy Model of Multidimensional Risk as a Tool for Population Health Research



Alexander N. Tyrsin , Dmitriy A. Yashin, and Alfiya A. Surina 

Abstract The development of medicine leads to more complex tasks. Such tasks include issues of complex multi-factor health assessment. In this chapter, we have a set of interrelated heterogeneous risk factors that can affect the state of human health, both one- and multidirectional. In such situations, it is also necessary to take the system patterns into account. For this purpose, a new tool for probability-entropy modeling of multidimensional risk is proposed.

Keywords Risk · Differential entropy · Multidimensional random variable · Vector · Randomness · Self-organization · Health · Population

1 Introduction

Comprehensive health assessment across multiple risk factors is one of the current and poorly studied problems in medicine. Risk factors are heterogeneous, interrelated, and can affect a person's state of health both one- and multidirectional. Therefore, it is not clear how to account for the contribution of each risk factor to the overall health assessment.

A. N. Tyrsin (✉)

Ural Federal University named after the first President of Russia B.N.Yeltsin, 19 Mira str., Ekaterinburg 620002, Russia
e-mail: at2001@yandex.ru

D. A. Yashin

South Ural State Medical University, 64 Vorovsky str., Chelyabinsk 454092, Russia
e-mail: yashid.chel@mail.ru

A. N. Tyrsin · A. A. Surina

South Ural State University (National Research University), 76 Lenin Av., Chelyabinsk 454080, Russia
e-mail: dallila87@mail.ru

Along with multidimensionality, another problem in assessing health is the complexity of a person as a biological system. It is well known that it is not possible to fully simulate the behavior of complex systems. Complex systems are multi-faceted, defined by many different indicators that make it difficult to choose a single criterion for management efficiency, and are characterized by the fact that the interaction of their elements cannot or is extremely difficult to represent explicitly [1]. The stochastic nature of biological systems is accepted, for example, in medical research in the statistical processing of the observed results of patients [2].

At present, we have accumulated a huge experience in the field of modeling biological systems. Many mathematical models have been developed that adequately describe certain aspects of the behavior and state of biological systems. Human research as a biological system had been studied by a separate science, medicine. On the other hand, there is a theory of systems—a knowledge unit, a scientific and methodological concept of studying objects that are systems. One of the basic problems of system theory is the study of complex systems. At the same time, the analysis of well-known results in the field of mathematical modeling of biological systems suggests that the models of humans and other representatives of wildlife as complex systems developed in medicine and biology are insufficiently based on the theory of systems. The lack of systematic mathematical models of biological systems is manifested in the fact that they cannot always explain the reasons for the change in the state of biological systems, in particular the appearance and development of diseases.

This leads to the need in ambiguous situations from formalized methods to expert assessments, which cannot be considered fully objectively [3–5].

One way to solve this problem may be to represent the population as a complex, multidimensional, stochastic, open, and self-organizing system. The use of entropy modeling made it possible to understand the main systemic causes of health deterioration [6]. But to make specific recommendations for improving health, both population, and individual patients, it is necessary to supplement system-wide representation with a quantitative assessment of the contribution to its deterioration of each of the risk factors. Risk assessment [7] is complicated by the fact that it is determined at the level of an individual population element without taking into account the relationship with other individuals, risk factors, and their mutual influence. This significantly reduces the timely diagnosis of diseases' early stages. To eliminate this disadvantage, a multidimensional risk model has been proposed in which the system is represented as a random vector with mutually correlated components [8]. It is interesting to combine the multidimensional risk model and the entropy model.

The obtained probability-entropy model of multidimensional risk, on the one hand, will take into account the multidimensionality and interconnectedness of elements of the stochastic system (risk factors), and on the other hand, it will allow for a systematic view of both the problems of medical diagnosis and prediction and the assessment of the effectiveness of medical and preventive measures. Thus, we will consider multidimensional risk as a systemic pattern, which will certainly lead to a deeper understanding of the risk problems and their impact on health.

2 Research Objective

Non-communicable diseases (NCDs), mainly cardiovascular diseases, oncological diseases, diabetes mellitus, and chronic bronchopulmonary diseases, are the main cause of death and disability worldwide. Every year, 41 million people die from these diseases worldwide, which is 71% of all deaths in the world [9]. The Director-General of the World Health Organization, Dr. Margaret Chen (2014), rightly said: “The world has reached a critical point in the history of the fight against NCDs and now has an unprecedented opportunity to change its development” [10]. However, the emerging progress towards achieving the UN’s (2015) «Sustainable Development Goals» on NCD mortality [11] has been identified as insufficient to achieve the ultimate goal [12]. One way to improve the situation in the fight against NCDs is to strengthen epidemiological surveillance and monitoring of the population’s health at the national level. The World Health Organization determines that the basic basis for such monitoring consists of three components, the first component is the impact monitoring to risk factors [13].

Such monitoring is carried out mainly at the population level, which in clinical epidemiology is understood as a population, a group of persons, united by territorial basis. Another definition is the set of individuals from which the sample is selected and to which the results obtained for this sample can be extended [14]. In the fight against non-communicable diseases, the population is a preventive medical and biosocial system (PMBS-system), which, as a complex system, is characterized by:

- complexity (at the same time there are several significant interacting/interdependent subsystems);
- multi-directional changes of analyzed parameters;
- these parameters may have different distributions and accordingly define different data types, averages, and other characteristics;
- during dynamic observation, the characteristics of the system can change significantly.

In the presence of such properties, the implementation of preventive programs to combat non-communicable diseases and their risk factors is associated with great difficulties, since it is difficult to assess and analyze population changes. A comprehensive quantifiable assessment of population-based changes is needed. This encourages the search and application of new mathematical methods for the analysis of PMBS systems in preventive medicine [15].

The purpose of this work is to describe a mathematical approach that combines the methods of multidimensional risk analysis and entropy modeling and its testing for the analysis of population changes in monitoring NCDs risk factors.

3 Mathematical Models and Methods

We assume that the state of population health as some complex multidimensional stochastic system, and distinguish risk factors X_1, X_2, \dots, X_m in that system. As a result, we get a representation of the system as a random vector with a certain probability density $p_{\mathbf{X}}(\mathbf{x})$.

3.1 Multidimensional Risk Model

Instead of the generally conventional selection of concrete dangerous situations, we will define the geometric area D of adverse outcomes. It may look arbitrarily depending on a specific objective.

The concept of dangerous states as larger and improbable deviations of a conception of dangerous conditions as large and improbable deviations of random variables x_{ij} from some best provision Θ is most distributed. In this case, D represents an external area of an m -axis ellipsoid [16].

$$D = \left\{ \mathbf{x} = (x_1, x_2, \dots, x_m) : \sum_{j=1}^m \frac{(x_j - \theta_j)^2}{b_j^2} \geq 1 \right\}$$

Setting the function of consequences from dangerous situations (risk function) in the form of $g(\mathbf{x})$, we will receive a model for the quantitative assessment of risk

$$r(\mathbf{X}) = \iiint_{\mathbf{R}^m} \dots \int g(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \tag{1}$$

If $g(\mathbf{x}) = 1 \forall \mathbf{x} \in D$ and $g(\mathbf{x}) = 0 \forall \mathbf{x} \notin D$, that $r(\mathbf{X}) = P(\mathbf{X} \in D)$, i.e. the risk is estimated as probability of an unfavorable outcome. Defining a function $g(\mathbf{x})$ requires a quantitative assessment of consequences for the studied system depending on the values of risk factors. It demands to carry out separate research. The risk function can be set, for example, as

$$g(\mathbf{x}) = \begin{cases} \sum_{j=1}^m \alpha_j (x_j - \theta_j)^2, & \mathbf{x} \in D, \\ 0, & \mathbf{x} \notin D, \end{cases} \tag{2}$$

where $\alpha_j = (D_j^- - \theta_j)^{-2}$ for $x_j < \theta_j$ and $\alpha_j = (D_j^+ - \theta_j)^{-2}$ for $x_j \geq \theta_j$. We believe that for each risk factor there is information about at least one of the individual values D_j^- (to the left θ_j) and D_j^+ (to the right θ_j), upon reaching which, as a result,

they become almost uncontrollable or irreversible. If D_j^- (D_j^+) is not present, we consider $D_j^- = -\infty$ ($D_j^+ = +\infty$).

In the problems of risk monitoring, along with risk assessment $r(\mathbf{X})$ on all risk factors of X_1, X_2, \dots, X_m of the multidimensional system is expedient to estimate the contribution of each factor to total risk. We introduce a random vector $\mathbf{X}_k^- = (X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_m)$. Then the absolute and relative changes of risk of the multidimensional system due to the addition of factor X_k are equal

$$\Delta r(X_k) = r(\mathbf{X}) - r(\mathbf{X}_k^-) \tag{3}$$

$$\delta r(X_k) = \Delta r(X_k) / r(\mathbf{X}_k^-) \tag{4}$$

Along with a contribution to the common risk of one factor, (3) and (4) allow us to estimate influence and groups of factors. As a result, the factors that had the greatest impact on risk growth are found.

3.2 Vector Entropy Model

The differential entropy of a multidimensional continuous random variable $\mathbf{X} = (X_1, X_2, \dots, X_m)$ was introduced by K. Shannon in 1948 [17]. It is defined as

$$H(\mathbf{X}) = - \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p_{\mathbf{X}}(x_1, \dots, x_m) \ln p_{\mathbf{X}}(x_1, \dots, x_m) dx_1 \dots dx_m, \tag{5}$$

where $p_{\mathbf{X}}(x_1, \dots, x_m)$ —the joint density function of random variables.

The vector entropy model consists of representing the differential entropy of the system as a two-dimensional vector [6]

$$\mathbf{h}(\mathbf{X}) = (h_V; h_R) = (H(\mathbf{X})_V; H(\mathbf{X})_R) \tag{6}$$

where $H(\mathbf{X})_V = \sum_{i=1}^m H(X_i) = \sum_{i=1}^m \ln \sigma_{Y_i} + \sum_{i=1}^m \kappa_i$ is the randomness entropy, $H(\mathbf{X})_R = \frac{1}{2} \sum_{k=2}^m \ln (1 - R_{X_k/X_1 X_2 \dots X_{k-1}}^2)$ is the self-organization entropy, $\kappa_i = H(X_i/\sigma_{Y_i})$ is an entropy indicator of the type of the distribution law of a random variable X_i ; $R_{X_k/X_1 X_2 \dots X_{k-1}}^2$ is the indices of determination of regression dependencies.

For the Gaussian random vector \mathbf{X} , we have

$$H(\mathbf{X})_V = \sum_{i=1}^m \ln \sigma_{X_i} + m \ln \sqrt{2\pi e}, \quad H(\mathbf{X})_R = \ln (|\mathbf{R}_X|) / 2,$$

where \mathbf{R}_X is the correlation matrix of the random vector \mathbf{X} .

If a random vector $X_m^- = (X_1, X_2, \dots, X_{m-1})$ is entered the changes in the entropies $H(\mathbf{X}), H(\mathbf{X})_V$ and $H(\mathbf{X})_R$ of the system due to the addition of the component X_m to it are equal

$$\Delta H(X_m) = \Delta H(\mathbf{X}) - \Delta H(\mathbf{X}_m^-), \tag{7}$$

$$\Delta H(X_m)_V = \ln \sigma_{X_m} + \kappa_m, \quad \Delta H(X_m)_R = \frac{1}{2} \ln (1 - R_{X_m/X_1 X_2 \dots X_{m-1}}^2). \tag{8}$$

For Gaussian random vector

$$\Delta H(X_m)_V = \ln (\sqrt{2\pi} e \sigma_{X_m}), \quad \Delta H(X_m)_R = \frac{1}{2} \ln \frac{|\mathbf{R}_X|}{|\mathbf{R}_{X_m^-}|}.$$

Along with contributing to the entropy of a single component system, similarly to Formulas (6), (7), one can estimate influence and groups of factors. As a result, the components that have had the greatest impact on the entropy of the system are found.

4 Results and Discussion

The prevention of chronic NCDs is the main resource in improving health, reducing mortality, and increasing life expectancy in Russia. It was carried out at both individual and population levels. At the same time, the population-based prevention strategy is the most effective, including from an economic point of view [18].

As an object of research, we need to consider the main biological factors of the risk of NCDs [18]. The analysis of dynamic changes of the PMBS system will be carried out on the materials of a comprehensive continuous survey of the organized male population (a team of employees of one of the metallurgical industry enterprises). The analysis includes survey data of 253 people in 2010 and 136 people in 2015 (the average age ($M \pm m$) is 46.0 ± 0.7 years and 46.4 ± 0.98 years, respectively) at the range of 22 to 64 years. Populations can be compared by age and nature of work activity. The analysis was performed in age groups: 18–44—young age (group I) and 45–64—middle age (group II) according to the recommendations of ICD-10 for grouping by age for general purposes.

As for indicators that characterize the main biological measurable risk factors for NCDs, we will consider X_1 —systolic blood pressure (SBP), mmHg; X_2 —body mass index (BMI; $\text{BMI} = \text{body weight (kg)}/\text{growth (m)}^2, \text{kg}/\text{m}^2$); X_3 —total blood cholesterol (Chol), mmol/L.; X_4 —blood glucose level (Glc), mmol/L. We consider the vector X to be Gaussian [19].

The analysis was carried out with the following comparisons: different age groups—one year of the survey (analysis of age changes); one age group—different survey years (chronological dynamics).

Table 1 Table of optimal, borderline and critical values of biological parameters that characterize the main risk factors of non-communicable diseases for conducting a risk analysis of medical and biosocial systems in the prevention of NCDs at the population level

Parameter	Lower level		Optimal level θ_j	Upper level	
	Critical D_j^-	Borderline d_j^-		Borderline d_j^+	Critical D_j^+
SBP	90	100	120	140	180
Chol	2.2	3.1	4.05	5.0	8.0
BMI	16.0	18.5	21.75	25.0	40.0
Glc	2.2	3.3	4.4	5.5	6.1

To determine the optimal, borderline and critical values of biological parameters which characterize the main risk factors for non-communicable diseases, when conducting a risk analysis of medical and biosocial systems (see Table 1) in the prevention of NCDs at the population level. The following methods were used:

- expert assessments;
- literature data (based on national clinical recommendations);
- preliminary survey and reference population survey data;
- determination of parameters relative to mathematical expectation (boundary values—deviations of two SD, critical levels—deviations of three SD).

It should be noted that the determination of parameters for risk analysis is a crucial stage and, in the last analysis, can be carried out only based on expert consensus using research data of the highest level of evidence and persuasiveness. The data in Table 1 are preliminary and can be changed in the future and used primarily to demonstrate the feasibility of using the probability-entropy model of population change analysis.

Table 2 presents a population-based scorecard of the probability-entropy model of multidimensional risk in the male preventive medical and biosocial system.

Tables 3 and 4 show the absolute and relative changes in the probability of an unfavorable outcome (the probability of getting into the NCDs occurrence zone for four main biological risk factors) and the risk of NCDs occurrence due to the addition of risk factors for different age groups. Based on the analysis of this information, the contribution of each subsystem (risk factors in a given case) to the change in the overall score can be estimated, which allows the identification of priority risk factors for prevention.

Table 5 presents the results of the analysis of changes in the male population across the four major risk factors for NCDs using a probability-entropy model. The left half of the table shows the analysis between age groups of 18–44 and 45–64 in the same year of survey, in the right half shows the comparison between single age groups in the 2010–2015 survey.

Several additional possibilities of using a probability-entropy multidimensional risk model in preventive medicine should be noted.

Table 2 Table of probability-entropy model indicators of multidimensional risk of the male preventive medical and biosocial system

Indicator	2010 survey year		2015 survey year	
	18-44	45-64	18-44	45-64
BMI level (kg/m ²) ($M \pm m(\delta)$)	24.63 ± 0.37 (3.6)	26.32 ± 0.29 ^a (3.64)	24.9 ± 0.48 (3.68)	27.0 ± 0.48 ^a (4.18)
SBP level (mmHg) ($M \pm m(\delta)$)	130.69 ± 1.39 (13.61)	140.31 ± 1.52^{a*} (19.05)	133.02 ± 2.5 (19.26)	150.14 ± 2.71^{a*} (23.76)
Chol level (mmol/l) ($M \pm m(\delta)$)	4.03 ± 0.09[*] (0.87)	4.3 ± 0.07 ^a (0.92)	4.36 ± 0.08[*] (0.6)	4.44 ± 0.1 (0.84)
Glc level (mmol/l) ($M \pm m(\delta)$)	3.87 ± 0.04[*] (0.4)	3.99 ± 0.06 (0.79)	4.12 ± 0.07[*] (0.57)	4.02 ± 0.06 (0.55)
H_V (randomness entropy)	8.496859	9.581841	8.821329	9.466392
H_R (self-organization entropy)	-0.09974	-0.0613	-0.01874	-0.0467
H (general entropy)	8.39712	9.520545	8.802591	9.419692
$P(D)$ (probability of transition to the NCDs risk zone)	0.602	0.874	0.749	0.911
$r(X)$ (overall risk of NCDs)	1.135	2.246	1.439	2.509

Note ^a—The differences between the averages characterize the underlying risk factors for NCDs in the age groups of 18-44 and 45-64 years of a single survey year are significant ($p < 0.05$); ^{*}—the differences between the averages indicators that characterize the main risk factors for NCDs in 2010 and 2015 in the same age groups are significant ($p < 0.05$) (the corresponding pairs of values are highlighted in bold)

Table 3 Absolute and relative changes in the probability of an adverse outcome due to the addition of risk factors for different age groups

Risk factor	Indicator	2010		2015	
		18–44	45–64	18–44	45–64
BMI	$\Delta r(X_k)$	0.254	0.171	0.260	0.133
	$\delta r(X_k)$ (%)	72.90	24.35	53.04	17.15
SBP	$\Delta r(X_k)$	0.081	0.116	0.162	0.149
	$\delta r(X_k)$ (%)	15.55	15.35	27.55	19.61
Chol	$\Delta r(X_k)$	0.052	0.066	0.074	0.015
	$\delta r(X_k)$ (%)	9.42	8.19	10.93	1.68
Glc	$\Delta r(X_k)$	0.001	0.012	0.017	–0.006
	$\delta r(X_k)$ (%)	0.11	1.41	2.37	–0.66

Table 4 Absolute and relative changes in the risk of non-infectious diseases due to the addition of risk factors for different groups of people

Risk factor	Indicator	2010		2015	
		18–44	45–64	18–44	45–64
BMI	$\Delta r(X_k)$	0.645	0.822	0.645	0.855
	$\delta r(X_k)$ (%)	131.49	57.75	81.29	51.73
SBP	$\Delta r(X_k)$	0.361	0.762	0.609	1.081
	$\delta r(X_k)$ (%)	46.66	51.34	73.36	75.71
Chol	$\Delta r(X_k)$	0.195	0.340	0.154	0.229
	$\delta r(X_k)$ (%)	20.81	17.83	12.03	10.07
Glc	$\Delta r(X_k)$	0.198	0.375	0.156	0.222
	$\delta r(X_k)$ (%)	21.13	20.01	12.16	9.72

Methods and strategy for population-based prevention of cardiovascular and other major NCDs have been well studied and developed [20, 21]. However, each population, especially organized, has its characteristics. For example, the population of accountants will differ from the population of therapists while maintaining the general principles of exposure to risk factors and response to preventive intervention. Researching to identify these features that meet the requirements of evidence-based medicine is economically and technically difficult in each subpopulation. The probability-entropy model allows us to identify these features (assess population health, internal relationships, determine the contribution of risk factors, etc.) and adjust the preventive intervention accordingly—to select priority risk factors, the preventive dose for this particular subpopulation, i.e. to conduct targeted population prevention.

When identifying patients with risk factors, a significant part of them with values of parameters that characterize risk factors close to the border of the norm falls out of the active control group (small risk group). We point to the conclusion of

Table 5 Map of analysis of population changes for 2010–2015 in the male organized population in the prevention of non-communicable diseases

Year of survey	Comparison between age groups I and II (18–44 and 45–64)	Comparison between the same age groups in 2010–2015 surveys	Age group, year
2010	<ul style="list-style-type: none"> expressed negative changes ($\Delta P(D) = +0.272$); pronounced increase in instability ($\Delta H = +13.5\%$); increase in average BMI and SBP levels; low level of subsystem interaction; the major contribution to the growth of $P(D)$ BMI and SBP with the decreased role of BMI 	<ul style="list-style-type: none"> moderate negative changes ($\Delta P(D) = +0.147$); increase in instability ($\Delta H = +4.9\%$); Increasing average Chol and Glucose levels; extremely low level of interaction of subsystems with reduction of HR $c - 0.1$ to -0.02; the major contribution to the growth of $P(D)$ BMI and SBP with the decreased role of BMI and increased SBP 	I 18–44
2015	<ul style="list-style-type: none"> moderate negative changes ($\Delta P(D) = +0.162$); moderate increase in instability ($\Delta H = +7.05\%$); increase of average BMI and SBP levels; low level of subsystem interaction; major contribution to growth of $P(D)$ BMI and SBP with decreased role of BMI 	<ul style="list-style-type: none"> minor negative changes ($\Delta P(D) = +0.037$); the system is stable ($\Delta H = -1.05\%$); increase of mean levels of SBP; extremely low level of subsystem interaction is stable; a more even contribution of all risk factors to $P(D)$ compared to other groups 	II 45–64

the well-known epidemiologist Rose: “From a large number of people at low risk, significantly more cases of the disease may occur than of the small number of people at high risk” [22]. The probability-entropy model allows these “borderline” patients to be identified—by considering the internal correlation, such patients fall into the NCDs risk zone.

Thus, based on the analysis of the use of systemic-entropy and risk-analysis in preventive medicine, combined into a probability-entropy model of multidimensional risk, the following conclusions can be drawn:

- the probability-entropy model allows to obtain a single numerically expressed assessment of population health in the prevention NCDs based on analysis of all significant subsystems at the same time (the probability of getting into the risk zone of NCDs and entropy characteristics);
- the probability-entropy model allows to identify priority subsystems for prevention taking into account the necessary preventive dose of intervention by analysis

of internal relationships (through the self-organization entropy) and determining the contribution of subsystems to the overall assessment of population health (through risk analysis);

- the probability-entropy model applies to the implementation of a population-based strategy for the prevention of NCDs.

5 Conclusion

The probability of the occurrence of non-communicable diseases is significantly higher in the older age group. In dynamics, a significant increase in the probability of non-communicable diseases is observed in the group of 18–44 years and a slight increase in the group of 45–64 years. The level of total entropy is significantly higher in the older age group in 2010 and 2015, i.e., since the preventive medical and biosocial system becomes more unstable with age. In dynamics, entropy growth was observed only in the age group 18–44 years. In all studied groups, there is an extremely low level of self-organization entropy, which indicates that there is no mutual influence of risk factors on each other.

Analysis of population dynamics in 2010–2015 showed an increased risk of NCDs in both age groups. This corresponds to the real picture, as average levels of risk factors tend to rise. The higher risk level in the older group is also explained by the deterioration of health with increasing age. The change in entropy is not so pronounced. There is a tendency to increase the randomness entropy with increasing age. The self-organization entropy in both age groups was almost unchanged and close to zero. This indicates that the manifestations of risk factors are independent.

The work was supported by the Russian Foundation for Basic Research (project no. 20-51-00001 Bel_a).

References

1. Sayama, H.: Introduction to the Modeling and Analysis of Complex Systems. Binghamton University, SUNY (2015)
2. Gerking, Sh, Adamowicz, W., Dickie, M., Veronesi, M.: Baseline risk and marginal willingness to pay for health risk reduction. *J. Risk Uncertainty* **55**(2), 177–202 (2017). <https://doi.org/10.1007/s11166-017-9267-x>
3. Kiryanov, B.F., Tokmachev, M.S.: Mathematical models in healthcare, p. 279. NovSU named Yaroslav the Wise, Great Novgorod (2009)
4. Yu, T.M., Kiryakov, D.A., Kamaltdinov, M.R.: Application of the complex index of population health disorders to assess population health in the Perm region. In: Proceedings of the Samara Scientific Center of the Russian Academy of Sciences, vol. 15, No. 3(6), pp. 1988–1992 (2013)
5. Corvellec, H.: The practice of risk management: silence is not absence. *Risk Manag.* **11**(3–4), 285–304 (2009). <https://doi.org/10.1057/rm.2009.12>
6. Tyrsin, A.N., Gevorgyan, G.G.: Population health assessment based on entropy modeling of multidimensional stochastic systems. *Commun. Comput. Inf. Sci. (CCIS)* **794**, 92–105 (2019). https://doi.org/10.1007/978-3-030-35400-8_7

7. Mun, J.: Modeling Risk, 2nd edn. Wiley (2010)
8. Tyrsin, A.N., Surina, A.A.: Monitoring of risk of the multidimensional stochastic system as tools for research of sustainable development of regions. IOP Conference Series: Earth and Environmental Science, vol. 177, 012005, 8 p. DOI: 10.1088/1755-1315/177/1/012005 (2018)
9. World Health Organization https://www.who.int/health-topics/noncommunicable-diseases#tab=tab_1 (2020)
10. Global status report on noncommunicable diseases. World Health Organization, Geneva (2014)
11. Goal 3.4., Sustainable Development Goals. Transforming Our World: The 2030 Agenda for Sustainable Development. United Nations. <https://sustainabledevelopment.un.org/post2015/transformingourworld> (2015)
12. World health statistics overview 2019: monitoring health for the SDGs, sustainable development goals. World Health Organization, Geneva (2019)
13. Report on the situation of noncommunicable diseases in the world, 2010 Executive summary, 21 p. Geneva: who. https://www.who.int/nmh/publications/ncd_report_summary_ru.pdf (2011)
14. Fletcher, R., Fletcher, S., Wagner, E.: Clinical Epidemiology. The Essentials, 3rd edn. Williams & Wilkins, Baltimore (1996)
15. Kalev, O.F., Kaleva, N.G., Yashin, D.A.: Quality of Human Health. Modern Problems of Science and Education, no. 4, 11 p. <http://www.science-education.ru/ru/article/view?id=25004> (2016)
16. Surina, A.A., Tyrsin, A.N.: Risk management in Gaussian stochastic systems as an optimization problem. Commun. Comput. Inf. Sci. (CCIS) **1090**, 562–577 (2019). https://doi.org/10.1007/978-3-030-33394-2_43
17. Shannon, C.E.: A Mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423, 623–656 (1948)
18. Prevention of chronic non-communicable diseases. Recommendations, 128 p. Moscow. <http://www.webmed.irkutsk.ru/doc/pdf/prevent.pdf> (2013)
19. Nasonova, N.V.: Automated system of complex monitoring of risk factors of chronic non-infectious diseases. Med. Inf. **1**(13), 56–67 (2007)
20. Rose, G.: The strategy of preventive medicine. Oxford University Press, Oxford (1992)
21. Tackling NCDs. “Best buys” and other recommended interventions for the prevention and control of noncommunicable diseases. World Health Organization, Geneva. <https://apps.who.int/iris/bitstream/handle/10665/259232/WHO-NMH-NVI-17.9-eng.pdf?sequence=1&isAllowed=y> (2017)
22. Rose, G.: Sick individuals and sick populations. Int. J. Epidemiol. **14**, 32–38 (1985)

Generative Models Based on VAE and GAN for New Medical Data Synthesis



Vladislav V. Laptev, Olga M. Gerget, and Nataliia A. Markova

Abstract The chapter deals with the construction of generative models using Variational Autoencoder (VAE) and Generative Adversarial Neural Networks to synthesize new medical data. VAE is a synthesis of two complete neural networks: an encoder E and a generator G, as well as the latent space connecting them and enabling them to carry out random transformation and interpolation. Generative Adversarial Nets (GAN) in their turn are built on the principle of interaction between a generative model (generator G) and a discriminating model (discriminator D). When creating generator G (both VAE and GAN), its architecture of a neural network based on convolutional layers, with the application of the new deep learning framework Tensorflow-addons is used. As E and D encoders, respectively, the models of transfer learning, problem domain-image feature vector are used in the work. The comparison between them is made in the chapter and the most optimal model for solving the proposed problem is selected. The chapter presents the results of the research obtained on the basis of VAE and GAN implementation.

Keywords Generative adversarial nets · Variational autoencoder · Convolution neural network · Generation · Transfer learning · Discrimination · Distribution · Probability · Training

V. V. Laptev (✉) · O. M. Gerget · N. A. Markova
National Research Tomsk Polytechnic University, Tomsk 634000, Russia
e-mail: vv139@tpu.ru

O. M. Gerget
e-mail: olgagerget@mail.ru

N. A. Markova
e-mail: markovana@tpu.ru

1 Introduction

Nowadays, generative models based on neural networks of deep learning are used to solve various problems: generation of handwritten digits [1], generation of faces [2], poetry writing [3], etc.

In modern medical practice, including the field of cardiology, there continues to be steady growth in the use of methods of automatic data graphic processing techniques. The most popular are algorithms for processing anatomical structures based on magnetic resonance imaging (MRI) and computed tomography (CT) data. However, in some cases, the use of the presented modalities is impossible, since, for example, the main restriction of computed tomography is the absence of real-time mode. Transfer learning aims to solve this problem. It is necessary to develop algorithms for data tracking and visualization based on intellectual analysis. VAE and GAN [4] implementation allow one to synthesize unique data for training the tracking model.

Popular generative models like a restricted Boltzmann machine and its many variants [5–7] have been successfully used in restricted conditions such as layer-by-layer pre-training. However, the implementation of generative models as a separate tool is difficult due to the problems in assessing maximum probability distribution.

2 Research Study

2.1 Variational Autoencoder

VAE consists of two neural networks: an encoder and a generator. The encoder receives the data at the input and performs some data transformations, which in the architecture of the neural network are presented in the form of convolutions, bringing thus the input data to a compact and compressed form of the feature vector. Then there is a special hidden space layer separating the feature vector obtained from E into two vectors: the mean value vector μ and the standard deviation vector σ . The mean value vector determines the point in the neighborhood of which the vertex will be located, while the values of the standard deviation vector determine how far the vertex can be located from this mean value vector. Therefore, the input object is not one point in the latent space, but a continuous area corresponds to it.

VAE forms parameters of the vector of length n out of random X_i values. These values, by means of concatenation, form an n -dimensional random vector that comes to the input of the generator to restore the data to the initial layout. This process allows different results to be obtained from one input sample due to the random selection of the coding vector. Figure 1 shows a classical scheme of Variational Autoencoder.

When training the Variational Autoencoder, the loss feature list was based on 2 metrics: mean squared error (MSE); Kullback–Leibler divergence (KLD) which shows the information divergence (relative entropy) of two probability distributions.

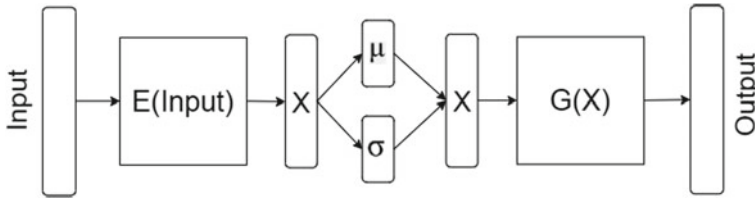


Fig. 1 VAE classical scheme

As a result, the loss function L was as follows:

$$L = K.mean(Reconstruction\ los + KLD\ loss) \tag{1}$$

$$Reconstruction\ loss = Original\ dim * MSE \tag{2}$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \tag{3}$$

$$KLD = \sum_{i=0}^N p(x_i) \cdot \log\left(\frac{p(x_i)}{q(x_i)}\right) \tag{4}$$

where y_i is the observation value, \hat{y}_i is the predicted observation value, N is the number of observations in the sample, $p(x_i)$ is the initial distribution, $q(x_i)$ is the approximation distribution.

This feature list allows areas to be located in the latent space as close to each other as possible, but at the same time differentiate them as separate components. In this case, it will be possible to perform smooth interpolation and generate unique data.

To solve this problem, it was agreed to implement the deep architecture of the generator network based on trained convolution and deconvolution layers with the use of parallel convolution and subsequent concatenation to amplify the features.

It should be noted that the deep convolutional generator has 5 levels of dimensional increase, and the total number of trained parameters was 26,690,336, more details can be found in Table 3. The description of the increase level structure (1 level, 1 = 4) is given in Table 1.

The following models are considered as the encoder (see Table 2). The analysis showed that the Xception model has the best Accuracy/Parameters ratio and therefore, it will be used for the further solution of the problem (Table 3).

The results of the loss function and the metrics under study are shown in Fig. 2. Based on the graphs, we can see that the model converges uniformly throughout the whole training.

Table 1 Description of VAE generator architecture layers

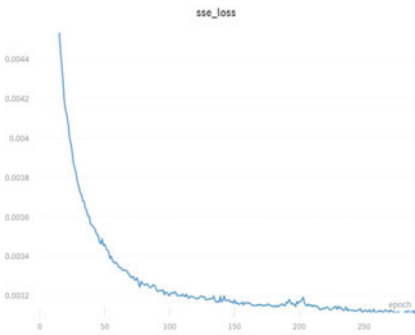
Name_layer	Kernel_size	Pool_size	Padding	Filters	Strides
Conv2d	(3, 3)	None	Same	512 (2**1*8)	None
Up_Sampling2d	None	(2, 2)	None	None	None
Conv2d_1	(1, 1)	None	Same	1024 (2**(1 + 1)*8)	None
Conv2dTranspose_1	(3, 3)	None	Same	512 (2**1*8)	(2, 2)
Conv2d_2	(1, 1)	None	Same	1024 (2**(1 + 1)*8)	None
Conv2dTranspose_2	(1, 1)	None	Same	512 (2**1*8)	(2, 2)

Table 2 Comparison of transfer learning models

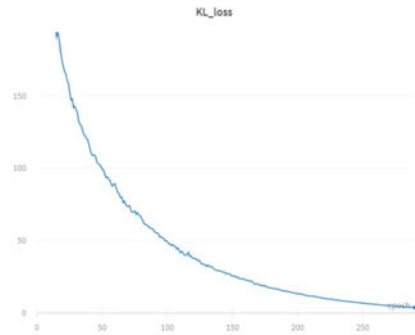
Model	Size (MB)	Top-1 accuracy	Top-5 accuracy	Parameters
ResNet50	98	0.749	0.921	25,636,712
ResNet101	171	0.764	0.928	44,707,176
ResNet50V2	98	0.760	0.930	25,613,800
InceptionV3	92	0.779	0.937	23,851,784
MobileNet	16	0.704	0.895	4,253,864
MobileNetV2	14	0.713	0.901	3,538,984
NASNetMobile	23	0.744	0.919	5,326,716
Xception	88	0.790	0.945	22,910,480

Table 3 VAE model weights

	Encoder	Generator	VAE
Total params	27,164,200	20,449,632	47,613,832
Trainable params	6,298,624	20,391,712	26,690,336
Non-trainable params	20,865,576	57,920	20,923,496



a) Track records of mean squared error



b) Track records of Kullback-Leibler divergence

Fig. 2 The result of the VAE loss function

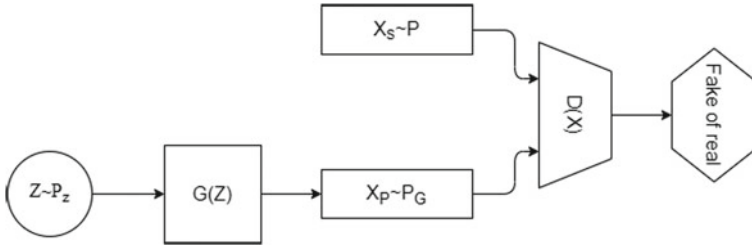


Fig. 3 Classical GAN architecture

2.2 Generative Adversarial Nets (GAN)

GANs consist of two neural networks, in this case functioning as a generator and a discriminator. The generator samples random numbers of some distribution $P(Z)$ and the input noise, for example, $N(0,1)$, and generates objects $X_p = G(Z; \theta_g)$ from them. The discriminator, getting to the input samples of the initial sample X_s and the generator X_p , learns to predict the properties of the sample (where it came from), giving the output scalar $D(X; \theta_d)$. Figure 3 presents the classical GAN structure.

The GAN training process can be represented as the following algorithm (the generator and the discriminator are trained separately but within the same network):

Step 1. Specify the random input $G(Z)$.

Step 2. Teach the discriminator D : parameters θ_d are updated to reduce binary cross-entropy (BCE):

$$BCE = -\frac{1}{N} \sum_{i=0}^N y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i) \tag{5}$$

$$\theta_d = \theta_d - \nabla_{\theta_d} (\log(D(X_s)) + \log(1 - D(G(Z)))) \tag{6}$$

Step 3. Teach the generator: parameters of the generator θ_g are updated to increase the logarithm probability that the discriminator assigns an actual label to the generated object:

$$\theta_g = \theta_g + \nabla_{\theta_g} (\log(1 - D(G(Z)))) \tag{7}$$

The GAN training process is reduced to solving the task:

$$\min_G \max_D E_{X \sim P} [\log(D(X))] + E_{Z \sim P_Z} [\log(1 - D(G(Z)))] \tag{8}$$

With the given generator, the optimal discriminator produces the probability:

$$D(X) = \frac{P(X)}{P_g(X) + P(X)} \tag{9}$$

Ian J. Goodfellow shows in his work [8] that the maximum capacity of both networks forms a scope in which the generator can generate the distribution of $P_g(X)$ coinciding with $P(X)$, and the discriminator on any X gives the probability of 0.5. Figure 4 illustrates the training process of the Generative Adversarial Nets [8].

The black dot curve in Fig. 5 is the real distribution of $P(X)$, the green curve is the distribution of the generator $P_g(X)$, the blue one is the distribution of the discriminator probability $D(X; \theta_d)$ to predict the object belonging. As a result of repeating steps a, b, c many times, $P_g(X)$ coincided with $P(X)$ and the discriminator is unable to distinguish a generated picture from a real one.

To solve the problem of the generator, it was settled to implement the generator network architecture (see Fig. 2). The distinctive feature in comparison with the VAE architecture is the input random vector of size $100 * 1$.

The Xception model of transfer learning shown in Table 2 is used as the discriminator. It was trained on Imagenet images. The total number of model weights is shown in Table 4.

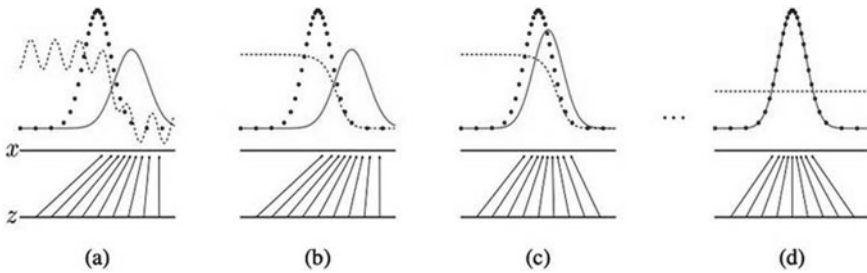
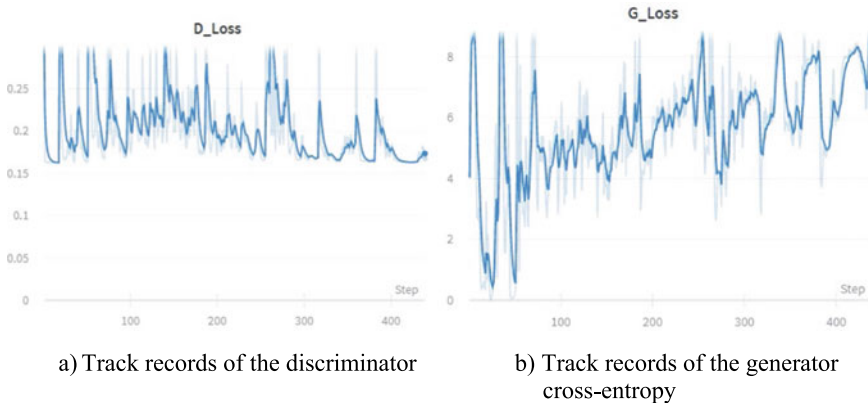


Fig. 4 GAN training process demonstration



a) Track records of the discriminator

b) Track records of the generator cross-entropy

Fig. 5 The results of the GAN loss function

Table 4 GAN model weights

	Generator	Discriminator	GAN
Total params	17,074,528	20,867,624	37,942,152
Trainable params	17,016,608	2,048	17,018,656
Non-trainable params	57,920	20,865,576	20,923,496

The results of loss functions are shown in Fig. 5. From the graphs, we can see that the model tends to reduce the loss to an average value. Thus, each step of the generator loss results in the reduction of the discriminator loss and vice versa.

3 Results

The training process has been tested on medical data to solve the problem of stenosis segmentation when diagnosing coronary arteries by angiography. The process of image reconstruction through the Variational Autoencoder is shown in Table 5. It shows examples of the images generated by the network.

The images in Table 5 correspond to the following sequence:

- (a) the 1st, the 100th, and the 200th epochs of learning. The analysis of the generated images has shown that their quality improves in the process of training by the 200th epoch.
- (b) the 200th, the 250th, and the 300th epochs of learning. The quality of the generated images improves, but with less progress than at the stage of training presented in line (a).
- (c)–(d) the 300th, the 350th, the 400th, the 500th, and the 600th epochs of learning. It should be noted that the quality of images generated by the network is not improving; therefore the network is no longer trained.

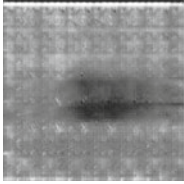
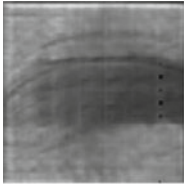
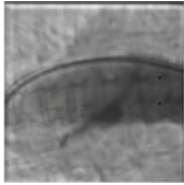
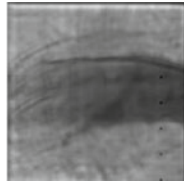
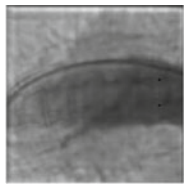
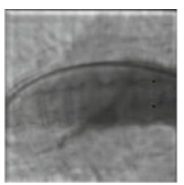
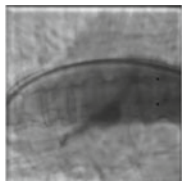
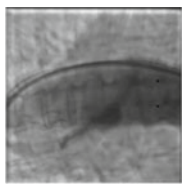
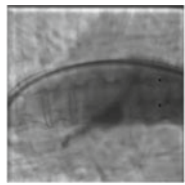
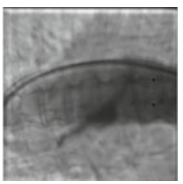
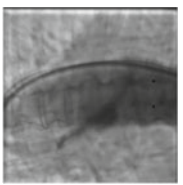
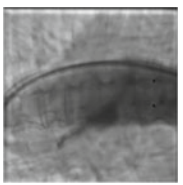
The process of image reconstruction through the Generative Adversarial Net models is shown in Table 6. The images correspond to the following sequence:

- (a) the 1st, the 100th, and the 200th epochs of learning. It is possible to notice that the result becomes better through the training process.
- (b) the 200th, the 250th, and the 300th epochs. The tendency of improving the quality of images is still present.

The images generated by the network and given in line (c) demonstrate a further informativity enhancement and achievement of the required quality. The analysis of the results presented in line (d) allows one to say that high accuracy of images has been achieved by the end of training (by the 500th epoch). Therefore, the network is trained fully and properly; it is ready to be used.

The given results show that the Generative Adversarial model generates more detailed images even though in the 200th epoch and continues to increase the accuracy throughout all the training. VAE shows similar results in the first 100 epochs, but

Table 5 VAE image reconstruction process

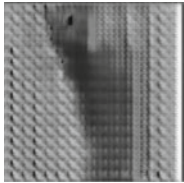
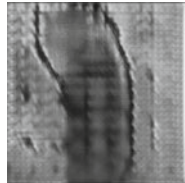
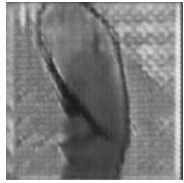
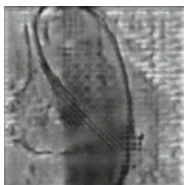
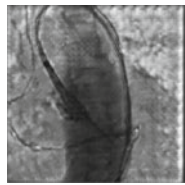
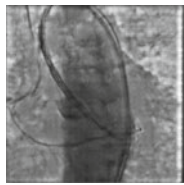
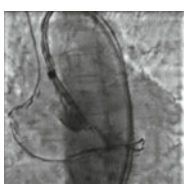
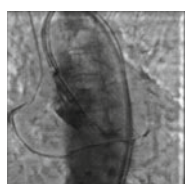
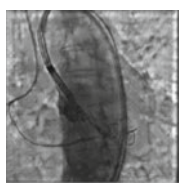



(a) Data visualization: epochs 1–200

(b) Data visualization: epochs 200–300

(c) Data visualization: epochs 300–400

(d) Data visualization: epochs 400–600

Table 6 GAN image reconstruction process

(a) Data visualization: epochs 1–200			
(b) Data visualization: epochs 200–300			
(c) Data visualization: epochs 300–400			
(d) Data visualization: epochs 400–500			

further, the images stop improving despite the positive track records of loss function reduction (see Fig. 3).

4 Conclusions

As a result of the research, the proper neural network architecture is developed. It is applicable both for the Convolutional Variational Autoencoder and for the Generative Adversarial model on the basis of which medical data generation is performed. The use of GAN, in contrast to VAE, is justified by reducing the time spent on searching for optimal model parameter values and increasing the accuracy of generated images.

It should be noted that the use of the Variational Autoencoder is reasonable in the case of closer scope data synthesis, as it inputs the real images to be changed. GAN is capable of synthesizing data with higher accuracy, which allows expanding

the training samples considerably. This method is used as an extension of the initial samples followed by its subsequent use in the task of keypoint tracking of anatomical structures and medical instruments.

Acknowledgements This work is performed within RFBR 19-07-00351\19 and State Assignment "Science" №FFSWW-2020-0014.

References

1. Introducing Variational Autoencoders (in Prose and Code). <https://blog.keras.io/building-autoencoders-in-keras.html> . Accessed 20 Nov 2019
2. Facial Surface and Texture Synthesis via GAN. <https://neurohive.io/en/computer-vision/facial-surface-and-texture-synthesis-via-gan> . Accessed 9 Jan 2020
3. How to teach your neural network to generate poetry. <https://habr.com/ru/post/334046> . Accessed 2 Apr 2020
4. Danilov, V.V., Skirnevskiy, I.P., Gerget, O.M., Shelomentcev, E.E., Kolpashchikov, D.Y., Vasilyev, N.V.: Efficient workflow for automatic segmentation of the right heart based on 2D echocardiography. *Int. J. Cardiovasc. Imaging* **34**(7), 1041–1055 (2018)
5. Bengio, Y., Mesnil, G., Dauphin, Y., Rifai, S.: Better Mixing via Deep Representations. *arXiv:1207.4404* (2012)
6. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. *Science* **313**(5786), 504–507 (2006)
7. Salakhutdinov, R., Hinton, G. E.: Deep Boltzmann machines. In: *International Conference on Artificial Intelligence and Statistics*, pp. 448–455 (2009)
8. Ian J. Goodfellow, Pouget-Abadie J., Mehdi Mirza, Bing Xu, Warde-Farley, D., Ozair S., Courville, A., Bengio, Y.: Generative Adversarial Nets. <https://arxiv.org/abs/1406.2661> (2014). Accessed 10 Feb 2020

Inhibitors Selection to Influenza Virus A by Method of Blocking Intermolecular Interaction



L. I. Zharkikh, Yu. A. Smirnova, I. M. Azhmukhamedov, E. V. Golubkina,
and M. N. Trizno

Abstract The application of a technique for blocking intermolecular interactions for the selection of inhibitors for influenza A virus has been proposed. The computational method included in the technique allows choosing drugs that can block the receptors of a specific virus. The technique was previously tested on the example of the action of drugs on a living cell, as well as on the selection of antidotes for hydrogen sulfide and is confirmed by experimental and literature data. Based on the proposed method, a software package for was developed for a targeted search for inhibitors of influenza A virus. The main goal of the method is to block, on the one hand, virus receptors and, on the other, living cell membrane components that are involved in the penetration of the virus into the cell. The stages that include the process of searching for virus blockers are described. Signature blocking schemes of interaction centers of substances involved in the formation of molar complexes are constructed. The chapter presents the studies and calculation results for the influenza A virus, conclusions are drawn, their coincidence with experimental data is noted.

L. I. Zharkikh (✉)

Caspian Institute of Maritime and River Transport Branch of Volga State Academy of Water Transport, St. Nicolskaya, 6, Astrakhan 414000, Russia
e-mail: lesy_g@mail.ru

Yu. A. Smirnova · I. M. Azhmukhamedov

Astrakhan State University, St. Tatischeva, 20A, Astrakhan 414000, Russia
e-mail: 2013qwer22@gmail.com

I. M. Azhmukhamedov

e-mail: aim_agtu@mail.ru

E. V. Golubkina · M. N. Trizno

Astrakhan State Medical University, st. Bakinskaya, 121, Astrakhan 414000, Russia
e-mail: neiron-2010@mail.ru

M. N. Trizno

e-mail: pakotm@yandex.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced
Human-Centered Society*, Studies in Systems, Decision and Control 333,
https://doi.org/10.1007/978-3-030-63563-3_18

Keywords Intermolecular interaction · Antidotes · Influenza A virus · Computational method · Information support · Blocking the action of the virus · Drugs

1 Introduction

For more than two decades, the molecular mechanisms and structure of molecules have been studied using quantum chemical programs by specialists in the field of structural, analytical, and other types of chemistry [7, 20, 23], as well as physicists [9, 10]. Using computer simulation methods and biological activity were analyzed to create new drugs [2, 4, 11, 17]. Earlier Mathematical modeling of intermolecular interactions helped us to study the pharmacological properties of the ephedrine and how it acts on the cell membrane and explain fat-burning and energetically stimulating agent. In the early works [8], the approach used by the authors was repeatedly tested on various drugs. An algorithm was developed for determining the centers of intermolecular interaction and software was created for conducting computational experiments. It was also proved that the created software products were used to search for inhibitors to hydrogen sulfide and experimental confirmation was obtained for one of the substances.

Thus, studies in this direction will helps to understand molecular interactions, select and arrange antidotes and confirm their resistance to toxicants. It becomes interesting the possibility of applying mathematical modeling of complex multicomponent molecular systems. A computer and a quantum-chemical simulation search for prediction of inhibitors to viral components, such as proteins of influenza A.

2 Problem Analysis

The infectious particle of the influenza virus (Fig. 1), called the virion, 80–200 nm in size, spherical shape with a shell. The envelope of the virus is a lipid membrane with built-in glycoproteins and matrix proteins that form ion channels. Under the lipid membrane is a matrix protein that forms the inner layer of the virus virion shell and also increase stability and rigidity to the outer lipid layer [6]. Glycoproteins, hemagglutinin and neuraminidase are the key proteins for the propagation of type A virus. Hemagglutinin (HA) acts in the moment of penetration of the virus into the cell, neuraminidase—upon living from it. Viruses attach to the sialic acids on the surface of the target cells by the external part of HA and the virions penetrate the cell through endocytosis. The specificity of the hemagglutinin receptor determines the type of association of N-acetylneuraminic (sialic) acid with galactose [1, 18].

Influenza A viruses are divided into serotypes, which depends on the combination of hemagglutinin, neuraminidase (NA) and proteins on the surface of the virus. As of

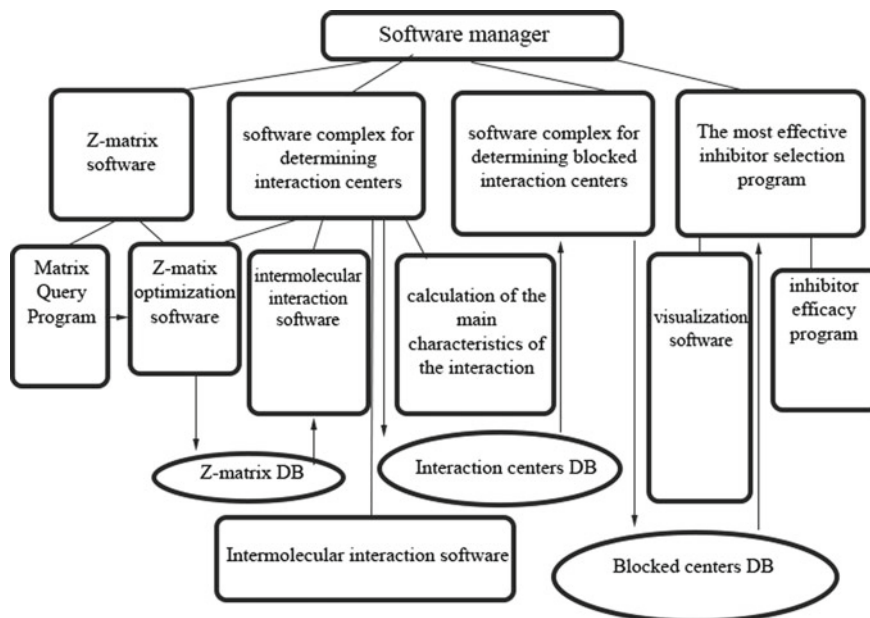


Fig. 1 The scheme of the software package for the targeted search for inhibitors of influenza A viruses

2016, scientists knew 18 subtypes of HA, 11 subtypes of NA, which together allows 198 variants of the influenza A virus to exist.

Alpha influenza virus virion contains 8 segments of viral RNA. Inside the virion is the virus genome, which carries genetic information about the envelope and internal proteins of the virus [5, 14–16, 25].

Six types of antiviral drugs used in the treatment of influenza: neuraminidase inhibitors, interferons, interferon inducers, cyclic amines, herbal antiviral drugs, and other antiviral drugs.. The effectiveness of some that inhibit the neuraminidase has been proven by studies (oseltamivir and zanamivir), while others continue to be tested (penamivir, laninamivir) [12, 22].

The main goal of the proposed method is to block, on the one hand, the influenza virus receptors and, on the other hand, the components of the membrane of a living cell that are involved in the penetration of the virus into the cell (Fig. 2).

3 Description of the Stages of the Decision

The search for influenza A virus blockers includes the following steps:

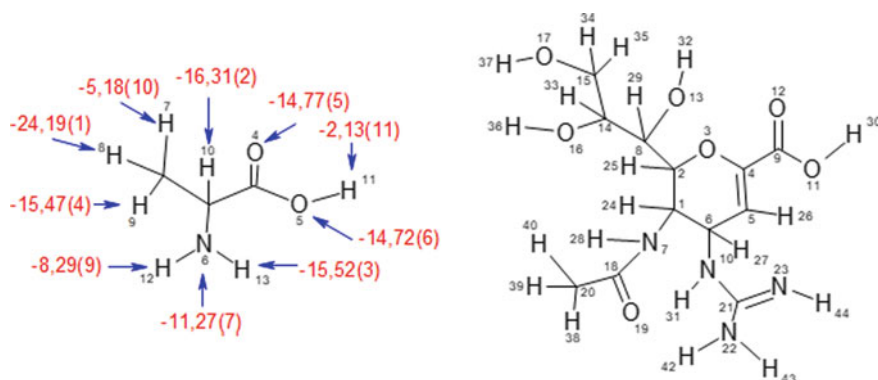


Fig. 2 Scheme of the signature of the active centers of the molecular complex: alanine + zanamivir

1. Highlight the main molecular structures that will form molecular complexes. Based on the selected molecular complexes, models will be built for a numerical experiment. Thus, at this stage, it is necessary to distinguish the structural components of the virus virion, which are most dangerous when it enters the cell of a living organism, and the structural components of the cell membrane of a living organism, which are involved in the formation of a connection with the virus.
2. Describe using a special code accepted for description in the Gamess program [24] in the form of z-matrix the composition and structure of each component of the molecular complex selected in step 1. To verify the constructed structures, a “preliminary” quantum-chemical calculation is performed for each studied molecules. All processed molecular structures in the form of z-matrices are stored in the Z-matrix database.
3. To apply the methodology for identifying “interaction centers” (atoms involved in the formation of a hydrogen bond between interacting molecules) of the interaction of two molecules, based on the analysis of information obtained as a result of systematic processing of data on the basic energy and geometric characteristics of the formation of the molecular complex. Each intermolecular interaction optimised (the process of analysis of the stability of intramolecular bonds) in a quantum-chemical program. Based on the identified interaction centers, a “signature of active centers” (a set of atoms involved in the formation of a hydrogen bond between interacting molecules) is compiled, and added to the database “Signatures of interaction centers”.
4. Visualize the results in the form of signature schemes of active centers for each intermolecular interaction. To improve the visibility and clarity of the obtained signatures, it is convenient to present them in the form of a scheme, an example of which is shown in Fig. 4. In this case, the direction of action of the atom of one molecule on the atom of another is displayed in the form of an arrow.
5. Apply the methodology for choosing the most effective drug. Intermolecular interactions are considered from two sides: the effect of the virus on the membrane

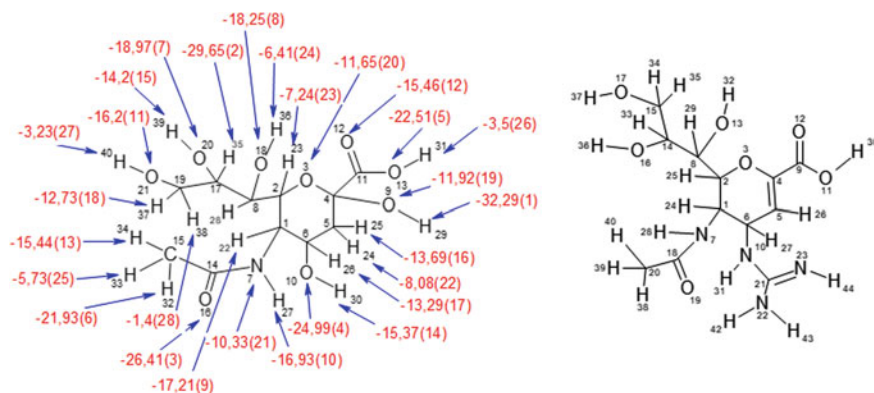


Fig. 3 Scheme of the signature of the active centers of the molecular complex: sialic acid + zanamivir

and the effect of the drug on the membrane. The obtained signatures of active centers for each membrane component are compared for two effects. Based on this comparison, a conclusion is made for each component. Information on blocking active centers is collected in the database “Blocked centers”. The effectiveness of the drug in relation to the action of the virus is calculated. The values of the effectiveness of the drugs are compared among themselves, and their ranked list is compiled.

The scheme of the software package that implements the methodology for the targeted search for inhibitors of influenza A virus is presented in Fig. 3.

4 Testing the Results

As a result of the implementation of the methodology for blocking intermolecular interactions for the selection of inhibitors for influenza A virus, 3 lists were compiled (Table 1) containing information about components of the cell membrane, the components of the influenza A virus-cell and the selected inhibitors comparison.

Using the software package, the structures of molecules in the form of z-matrix were simulated and calculated, a preliminary quantum-chemical calculation was performed for each molecule to form “Z-matrix” database. Based on these matrices, the following intermolecular interactions were simulated:

- each component of the influenza A virus + each component of the cell membrane;
- each drug + each component of the cell membrane.

Z-matrices of the structures of intermolecular interactions were compiled, in total $23 * 20 + 15 * 20 = 760$ z-matrices.

Table 1 Lists of molecules involved in interactions

Cell membrane components	Influenza A virus components	Drugs
Alanine	Sialic acid	Zanamivir
Arginine	Galactose	Oseltamivir
Asparagine	Tryptophan	Sialic acid
Aspartic acid	Alanine	Remdesivir
Valine	Arginine	Chloroquine
Histidine	Asparagine	Hydroxychloroquine
Glycine	Aspartic acid	Umifenovir
Glutamic acid	Valine	Blockavirmarboxyl
Isoleucine	Histidine	Rimantadine
Leucine	Glycine	Cycloferon
Lysine	Glutamic acid	Lopinavir + ritonavir
Methionine	Isoleucine	Routine
Proline	Leucine	Vitamin C
Serine	Lysine	Interferon b-b1
Tyrosine	Methionine	
Threonine	Proline	
Tryptophan	Serine	
Phenylalanine	Tyrosine	
Cysteine	Threonine	
	Tryptophan	
	Phenylalanine	
	Cysteine	

At this stage, algorithms were used that create z-matrices of the structures of intermolecular interactions based on data taken from the Z-matrix database. The optimization of the obtained molecular complexes was carried out in the Gamess quantum chemical program [24].

For each molecular complex, the following property was verified:

1. The distance ρ between the atoms for which the formation of a hydrogen bond is assumed must not exceed the allowable hydrogen bond for these atoms.
2. The adsorption energy must be strictly less than zero, i.e. when a bond is formed, energy should only be released. This value represents the difference between the total energy of the molecular complex and the sum of the total energies of the molecules that form it.
3. The sum of the atomic charges of each molecule in the molecular complex must be non-zero. A value of zero corresponds to a stable molecule or molecular complex. When a hydrogen bonds are formed between the molecules, the electron density is redistributed in the space between the molecules; this value for each molecule

in the complex shifts from zero to the side, corresponding to the specifics of the resulting bond.

If at least one of the conditions is not satisfied, then the active center will not form.

The calculation results of these characteristics for some molecular complexes are presented in Tables 2 and 3. Molecular complexes (MK) are numbered in order of increasing adsorption energy.

Based on the identified centers of intermolecular interaction, “signatures of active centers” were compiled and stored in the database “Signatures of interaction centers”.

Signature schemes for substances, the results of the interaction are shown in Tables 2 and 3 are presented in Figs. 2 and 3, respectively. The arrows indicate the direction of action of an external substance on the cell component. The numbers indicate the energy characteristics, and the values in brackets correspond to the MK number in the corresponding table of energy and geometric characteristics.

On the left is an alanine molecule interact with zanamivir. On the right is the structural formula of the zanamivir molecule. The numbers indicate the energy characteristics, and the values in parentheses correspond to the number.

On the left is a sialic acid molecule interact with zanamivir. On the right is the structural formula of the zanamivir molecule. The numbers indicate the energy characteristics, and the values in parentheses correspond to the number.

As we see, in the above example, the action of the zanamivir completely blocks the active atoms of alanine and the active atoms of sialic acid, thereby using it can prevent the effects of these substances on each other. Similarly, all possible options for interaction the components of the cell membrane with the active molecules of the virus virions are sorted out.

Similar schemes and tables of energy and geometric characteristics were obtained for all 760 molecular complexes as a result of calculations performed on the software package created by the authors.

5 Conclusions

Obviously, to analyze the whole picture as a whole, it is necessary to study the system in full, with all the structural features of molecular structures and consider how the complete molecular chains will be located and interact when modeling the general structure, but for modern computer powers this, unfortunately, remains impossible. However, the experiments and conclusions can be confirmed based on the literature of the authors who conducted experiments on these substances [3, 13, 19, 21]. Therefore, we accept and generalize our results to the entire system in aggregate and obtain the following conclusions:

- as a result of the application of the proposed methodology for the selection of influenza A receptor blockers, the possible effectiveness of the following drugs was studied and proved:

Table 2 The main energy and geometric characteristics of molecular complex alanine + zanamivir formation

MK	Bond	R (Å)	Δq (e)	E _{abs} (kJ/mole)	MK	Bond	R (Å)	Δq (e)	E _{abs} (kJ/mole)
1	N14...H8	2.78	-0.0065	-24.19	7	H25...N6	2.74	-0.0023	-11.27
2	O11...H10	1.85	0.0110	-16.31	8	O12...H8	1.86	0.0098	-9.12
3	O12...H13	1.86	0.0133	-15.52	9	O12...H12	1.86	0.0169	-8.29
4	O13...H9	1.87	0.0125	-15.47	10	O12...H7	1.88	0.0128	-5.18
5	H25...O4	1.89	-0.0367	-14.77	11	O11...H11	1.85	-0.0053	-2.13
6	H30...O5	1.81	-0.0212	-14.72					

Table 3 The main energy and geometric characteristics of molecular complex sialic acid + zanamivir formation

MK	Bond	R (Å)	Δq (e)	E _{abs} (kJ/mole)	MK	Bond	R (Å)	Δq (e)	E _{abs} (kJ/mole)
1	H29...O12	1.82	0.0425	-32.29	15	H39...O11	1.85	0.0134	-14.20
2	H35 ... O13	1.88	-0.0149	-29.65	16	H25...O12	1.85	-0.0018	-13.69
3	O16... H34	1.79	-0.0288	-26.41	17	H26...O19	1.86	0.0186	-13.29
4	O10... H28	1.84	-0.0225	-24.99	18	H37...O16	1.89	0.0077	-12.73
5	O13... H25	1.88	0.0275	-22.51	19	O9...H25	1.88	-0.0127	-11.92
6	H32 ...O12	1.84	-0.0143	-21.93	20	O3...H30	1.83	-0.0146	-11.65
7	O20... H25	1.87	-0.0103	-18.97	21	N7... H24	2.57	-0.0066	-10.33
8	O18... H30	1.80	-0.0216	-18.25	22	H24...O12	1.88	0.0105	-8.08
9	H22...O12	1.86	-0.0135	-17.21	23	H23...O12	1.86	0.014	-7.24
10	H27... O11	1.89	0.0075	-16.93	24	H36...O11	1.86	0.0063	-6.41
11	O21... H24	1.86	-0.0142	-16.20	25	H33...O12	1.85	0.01	-5.73
12	O12... H25	1.85	-0.0089	-15.46	26	H31...N7	1.86	0.0416	-3.50
13	H34...O16	1.87	0.0102	-15.44	27	H40...O11	1.86	0.0139	-3.23
14	H30...O12	1.82	0.0201	-15.37	28	H38...O19	1.86	0.0187	-1.40

- zanamivir;
- oseltamivir;
- sialic acid;
- remdesivir;
- chloroquine;
- hydroxychloroquine;
- umifenovir;
- blockavirmarboxyl;
- rimantadine;
- cycloferon;
- lopinavir + ritonavir;
- routine;
- vitamin C;
- interferon b-b1.

this technique can serve as the basis for subsequent experimental studies on living systems of already known drugs, and also allows for predictive calculations for new substances.

References

1. Arranz, R., Coloma, R., Chichón, F.J., Conesa, J.J., Carrascosa, J.L., Valpuesta, J.M., Ortín, J., Martín-Benito, J.: The structure of native influenza virion ribonucleoproteins. *Science* (New York, N.Y.), **338**(6114), 1634–1637 (2012). <https://doi.org/10.1126/science.1228172>
2. Aucar, M.G., Cavasotto, C.N.: Molecular docking using quantum mechanical-based methods. *Methods Mol Biol* (Clifton, N.J.) **2114**, 269–284 (2020). https://doi.org/10.1007/978-1-0716-0282-9_17
3. Bright, R.A., Medina, M.J., Xu, X., Perez-Oronoz, G., Wallis, T.R., Davis, X.M., Povinelli, L., Cox, N.J., Klimov, A.I.: Incidence of adamantane resistance among influenza A (H3N2) viruses isolated worldwide from 1994 to 2005: a cause for concern. *Lancet* (London, England) **366**(9492), 1175–1181 (2005). [https://doi.org/10.1016/S0140-6736\(05\)67338-2](https://doi.org/10.1016/S0140-6736(05)67338-2)
4. Bryce, R.A.: What next for quantum mechanics in structure-based drug discovery? *Methods Mol Biol* (Clifton, N.J.) **2114**, 339–353 (2020). https://doi.org/10.1007/978-1-0716-0282-9_20
5. Carrillo, B., Choi, J.M., Bornholdt, Z.A., Sankaran, B., Rice, A.P., Prasad, B.V.: The influenza A virus protein NS1 displays structural polymorphism. *J. Virol.* **88**(8), 4113–4122 (2014). <https://doi.org/10.1128/JVI.03692-13>
6. Dadonaite, B., Vijayakrishnan, S., Fodor, E., Bhella, D., Hutchinson, E.C.: Filamentous influenza viruses. *J. General Virol.* **97**(8), 1755–1764 (2016). <https://doi.org/10.1099/jgv.0.000535>
7. Das, G.K., Bhattacharyya, D., Burma, D.P.: A possible mechanism of peptide bond formation on ribosome without mediation of peptidyl transferase. *J. Theor. Biol.* **200**(2), 193–205 (1999). <https://doi.org/10.1006/jtbi.1999.0987>
8. Dias, D.A., Urban, S., Roessner, U.: A historical overview of natural products in drug discovery. *Metabolites* **2**(2), 303–336 (2012). <https://doi.org/10.3390/metabo2020303>
9. Él'Kin, P., Pulin, O., Dzhalmukhambetova, E.: Theoretical analysis of vibrational spectra of tautomeric purine forms. *J. Appl. Spectrosc.* **75**, 21–26 (2008). <https://doi.org/10.1007/s10812-008-9018-5>

10. Él'kin, P., Pulin, O., Dzhalmukhambetova, E.: Anharmonic analysis of vibrational states for five-membered heterocyclic compounds. *J. App. Spectrosc.* **74**, 169–173 (2007). <https://doi.org/10.1007/s10812-007-0026-7>
11. Gao, Z.G., Kim, S.K., Biadatti, T., Chen, W., Lee, K., Barak, D., Kim, S.G., Johnson, C.R., Jacobson, K.A.: Structural determinants of A(3) adenosine receptor activation: nucleoside ligands at the agonist/antagonist boundary. *J. Med. Chem.* **45**(20), 4471–4484 (2002)
12. Higashiguchi, M., Matsumoto, T., Fujii, T.: A meta-analysis of laninamivir octanoate for treatment and prophylaxis of influenza. *Antiviral Therapy* **23** (2017). <https://doi.org/10.3851/IMP3189>
13. Murphy, K.R., Eivindson, A., Pauksens, K., Stein, W.J., Tellier, G., Watts, R., Léophonte, P., Sharp, S.J., Loeschel, E.: Efficacy and safety of inhaled zanamivir for the treatment of influenza in patients with asthma or chronic obstructive pulmonary disease: a double-blind, randomised, placebo-controlled, multicentre study. *Clinical Drug Invest.* **20**, 337–349 (2000). <https://doi.org/10.2165/00044011-200020050-00005>
14. Lakdawala, S.S., Nair, N., Hutchinson, E.: Educational material about influenza viruses. *Viruses* **11**(3), 231 (2019). <https://doi.org/10.3390/v11030231>
15. Le Sage, V., Nanni, A.V., Bhagwat, A.R., Snyder, D.J., Cooper, V.S., Lakdawala, S.S., Lee, N.: Non-Uniform and Non-Random binding of nucleoprotein to Influenza A and B viral RNA. *Viruses* **10**(10), 522 (2018). <https://doi.org/10.3390/v10100522>
16. Lee, N., Le Sage, V., Nanni, A., Snyder, D., Cooper, V., Lakdawala, S.: Genome-wide analysis of influenza viral RNA and nucleoprotein association. *Nucleic Acids Res.* **45**, 8968–8977 (2017). <https://doi.org/10.1093/nar/gkx584>
17. Li, Z., Wan, H., Shi, Y., Ouyang, P.: Personal experience with four kinds of chemical structure drawing software: review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch. *J. Chem. Inf. Comput. Sci.* **44**, 1886–1890 (2004). <https://doi.org/10.1021/ci049794h>
18. Lyons, D.M., Lauring, A.S.: Mutation and epistasis in influenza virus evolution. *Viruses* **10**(8), 407. <https://doi.org/10.3390/v10080407>
19. Muthuri, S.G., Venkatesan, S., Myles, P.R., Leonardi-Bee, J., Lim, W.S., Al Mamun, A., Anovadiya, A.P., Araújo, W.N., Azziz-Baumgartner, E., Báez, C., Bantar, C., Barhoush, M.M., Bassetti, M., Beovic, B., Bingisser, R., Bonmarin, I., Borja-Aburto, V.H., Cao, B., Carratala, J., Cuezzone, M.R., et al.: Impact of neuraminidase inhibitors on influenza A(H1N1)pdm09-related pneumonia: an individual participant data meta-analysis. *Influenza Other Respiratory Viruses* **10**(3), 192–204. <https://doi.org/10.1111/irv.12363>
20. Nagy, G., Gyurcsik, B., Hoffmann, E.A., Körtvélyesi, T.: Theoretical design of a specific DNA-Zinc-finger protein interaction with semi-empirical quantum chemical methods. *J. Mol. Graph. Model.* **29**(7), 928–934 (2011). <https://doi.org/10.1016/j.jmglm.2011.03.002>
21. Penttinen, P., Catchpole, M.: ECDC expert opinion on efficacy and effectiveness of neuraminidase inhibitors published for public consultation. *Influenza Other Respir. Viruses* **10**(3), 152–153 (2016). <https://doi.org/10.1111/irv.12377>
22. Peramivir for Influenza: Australian Prescriber **42**(4), 143 (2019). <https://doi.org/10.18773/ausprescr.2019.047>
23. Rezakazemi, M., Albadarin, A.B., Walker, G.M., Shirazian, S.: Quantum chemical calculations and molecular modeling for methylene blue removal from water by a lignin-chitosan blend. *Int. J. Biol. Macromol.* **120**(Pt B), 2065–2075 (2018)
24. Schmidt, M.W., Baldrige, K.K., Boatz, J.A., Elbert, S.T., Gordon, M.S., Jensen, J.H., Koseki, S., Matsunaga, N., Nguyen, K.A., Su, S., Windus, T.L., Dupuis, M., Montgomery, J.A., Jr.: General atomic and molecular electronic structure system. *J. Comput. Chem.* **14**, 1347–1363 (1993). <https://doi.org/10.1002/jcc.540141112>
25. Xu, X., Zhu, X., Dwek, R.A., Stevens, J., Wilson, I.A.: Structural characterization of the 1918 influenza virus H1N1 neuraminidase. *J. Virol.* **82**(21), 10493–10501 (2008). <https://doi.org/10.1128/JVI.00959-08>

The Stochastic and Singular Analysis of Fractal Signals in Cyber-Physical Systems of Biomedicine



Vladimir Kulikov, Alexander Kulikov, and Alexander Ignatyev

Abstract This chapter deals with the consideration of methods for the analysis of experimental data in cyber-physical systems (CPS) for medical monitoring, the creation of stochastic components of expert systems for diagnostic and therapy processes. These methods are relevant for creating mathematical models of mechanisms for the functioning of human organs and systems with fractal stochastic properties as well as for managing complex subsystems of cyber-physical systems in Biomedicine. The results of the study are presented as related to the stochastic and singular analysis of the experimental implementation of a fractal random process—an EGG-signal in gastroenterology. The distribution laws for the local segment of the process have been identified based on algorithms and programs elaborated by the authors for diagnostic purposes, a modification of the singular spectral analysis (SSA-method) has been implemented for compressing signals and accelerating diagnosis; trajectory matrices of the EGG-signal have been investigated. It has been shown that segments of the EGG-signal and their first differences have polymodal distribution densities.

Keywords CPS · Biomedicine · EGG-signal · Fractal process · Polymodal densities · SSA-method · Monitoring · Diagnostics

V. Kulikov (✉) · A. Kulikov
Nizhny Novgorod State Technical University n. a. R. E. Alekseev, 24 Minin St., Nizhny Novgorod
603950, Russia
e-mail: vb.kulikov@yandex.ru

A. Kulikov
e-mail: akulikov@nntu.ru

A. Ignatyev
Medical Diagnostics, Ltd., 12B Novikov-Priboya St., Nizhny Novgorod 603064, Russia
e-mail: alarig@yandex.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced
Human-Centered Society*, Studies in Systems, Decision and Control 333,
https://doi.org/10.1007/978-3-030-63563-3_19

1 Introduction

This chapter is concerned with modern methods for the analysis of experimental data in cyber-physical medical monitoring systems to control treatment and diagnostic processes. These methods are based on algorithms for solving inverse incorrect problems and relevant for creating mathematical models of human body structures, organs, and systems that display stochastic and fractal properties.

Bioelectrical signals recorded during monitoring from organs of the gastrointestinal tract (GIT), the cardiovascular or respiratory systems are unsteady (in general) processes, which reflect nonlinear dynamic properties of the body structures.

The evolution of physiological processes in terms of random processes and values manifests itself as complex (non-Gaussian, polymodal) distributions of biosignal parameters to be identified for correct diagnosis and therapy management both in clinical and research problems.

The law of distribution of stochastic characteristics is necessary to form a mathematical probabilistic model of a local object for the biomedical CPS control. Based thereon it is possible to present an experimental-mathematical model of the human respiratory or digestive system functioning, to create systems for monitoring of critical parameters in the course of therapy, to predict future conditions, and to select effective medicines.

The chapter describes the results of the study as related to the stochastic and singular analysis of an electrogastroenterographic signal (EGEG signal) with its fractal nature. The laws of distribution of local segments in this fractal process are identified, the SSA method modification has been implemented to compress the signal, and to accelerate diagnostics; the trajectory matrices of the EGEG signal have been investigated.

Gastrointestinal electrogastroenterography is currently fast developing as an alternative to probe diagnostic methods [1–5]. Bioelectric signals of this kind may be referred to as the class of quasi-fractal random processes or, in other terms, to models of the fractional Brownian motion.

The fractional Brownian motion is a generalization of the classical Brownian motion. The probability distribution of the standard Brownian motion is based on the Gaussian statistics, i.e. the unimodal distribution density of characteristics of the stochastic structure, process. A rigorous description of such random processes is quite difficult (the identification of joint multidimensional distributions). And a conceptual analysis is only elaborated for stationary, Markov, Gaussian processes. Despite its many advantages, Fourier analysis does not track the “inclusion” points of new random harmonics.

While computational methods in the study of such processes in medicine are just started to be mastered, in physics the fractional Brownian motion has been studied for quite a long time; it is generally regarded as a random process non-complying with Gaussian statistics. For example, to study fluctuations in high-temperature plasma in modern torsatrons, unimodal Levy distributions are used [6, 7], which describe plasma turbulences by the model of the fractional Brownian motion.

Interest in the Levy motion is caused by anomalous diffusion problems, the motion of charged particles in turbulent electrostatic fields, and plasma superdiffusion in a magnetic field [8–10]. Characteristical and momental functions of this distribution are used to identify deviations of the fluctuation process distribution from Gaussian ones in the physical chemistry of polymers and statistical mechanics [9, 11–14].

The kinetic description of the Levy motion is based on fractional derivatives in the Fokker–Planck equation and is proposed as a phenomenological equation for PDF (Probability Density Functions) fluctuations of densities measured with a tokamak [15].

Papers [16–19] may be noted as fundamental theoretical works assigning the so-called dissipative nonlinear equations for calculating local fluctuations in “continua” in hydrodynamics, chemical kinetics, and physics of continua. In a situation with the system approaching critical parameters, the linear Gaussian approximation is not sufficient anymore; the nonlinear fluctuation theory is needed. Polymodal distributions are the reflection of nonlinear and fractal processes.

In [20] a fractal model of the EGEG signal as the fractional Brownian motion was first considered in the concept of polymodal distributions.

A characteristic feature of an electrogastroenterographic signal is a combination of probabilistic and persistent properties determined by chaotic patterns. The problem arises to verify the peculiarities of such processes in order to create effective algorithms for their stochastic analysis, including the use of the SSA method.

2 Solution Method and Computational Experiments

The problems set in the performed studies are solved by the sequential identification of distribution laws for local segments of the EGEG signal (sampling implementations) or by the moving sample method. It enables us to calculate from the reconstituted densities the numerical characteristics and moment functions of the EGEG signal segments on the basis of restored densities and to identify classification features.

The suggested approach is complementary to known methods with ample opportunities for the consideration of stochastic properties of fractal phenomena. The identification results obtained during the studies at various time scales will help to create models displaying the human system and organ behavior at the level of the Ito or Fokker–Planck equations with fractional derivatives. As compared to well-known forecasting methods the suggested algorithms may identify the evolution of distribution densities in near real-time and provide valuable information for verifying stochastic solutions.

Figure 1 shows an experimental graph of the total peripheral EGEG signal and its initial segment of 1200 units. It is seen that this is an essentially unsteady, walking random process. This graph demonstrates quite a frequent return of the process values to the neighborhood of the zero ordinate.

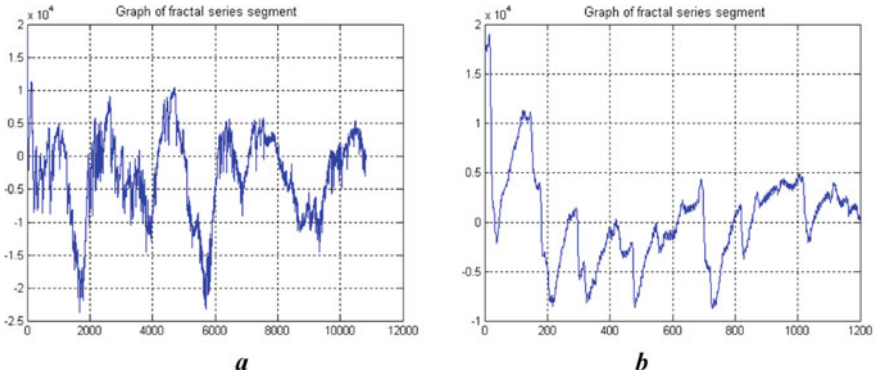


Fig. 1 Graph of the EGEG signal (the number of samples is over 10,000) (a); an initial segment of the process (the number of samples is 1200) (b); the Y-axis measures the EGEG-signal potential difference, mkV; the X-axis shows

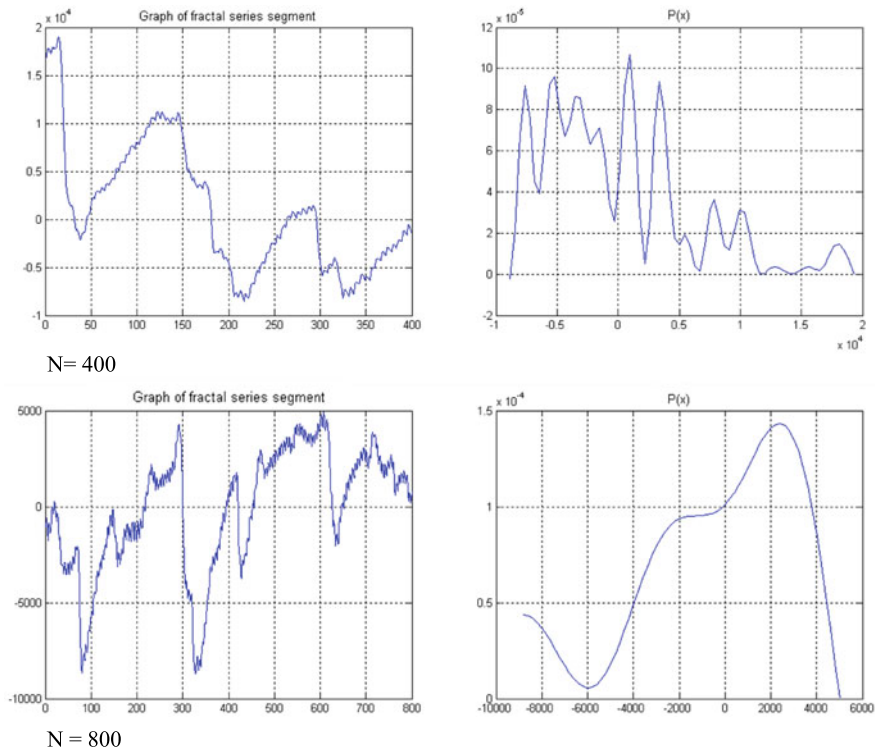


Fig. 2 Graphs of segment realizations (l-h series) and relevant identified densities of the EGEG signal distribution (400, 800 samples)—(r-h series). N is the sample length

Figure 2 shows the results of identifying probability distributions of samples (segments of time series) of various lengths.

The studies of segments with the different length of a sample have identified the multimodal distribution densities of the EGEG signal on different time scales. It enables us to formulate the following hypothetical statement: the polymodal stochastic characteristic of the distribution law corresponds to the fractal nature of the process.

Based on the generalization and correspondence principle the following working hypothesis is also made: the stable Levi-Khinchin distribution—the approximation of fractional Brownian motion—is only characteristic of certain types of anomalous diffusion.

But in general, fractal processes and fluctuations should be described by polymodal distributions. This conclusion is proved based on the analysis of publications on the statistical study of characteristics of fractal-like structures and processes in natural science and technology [21].

The identification of the distribution density of the EGEG signal implementation (10,818 counts) has been performed for the first time. The problem has been solved using elaborated algorithmic and software tools within 43 s. The result obtained is shown in Fig. 3a. The general distribution is bimodal, thus, confirming once again the distinctive stochastic property of fractal processes. Graph (b) in Fig. 3 shows the dynamics of such an important parameter of the identification algorithm as the condition number of systems of equations to be solved and reflects the degree of its stability.

This chapter is also the first to describe the results of identifying the distribution laws for the EGEG signal increments. Signals of the first and second increments (differences) of both local segments and full implementation have been studied. In general, the polymodality of laws for signal-increments has been revealed. For

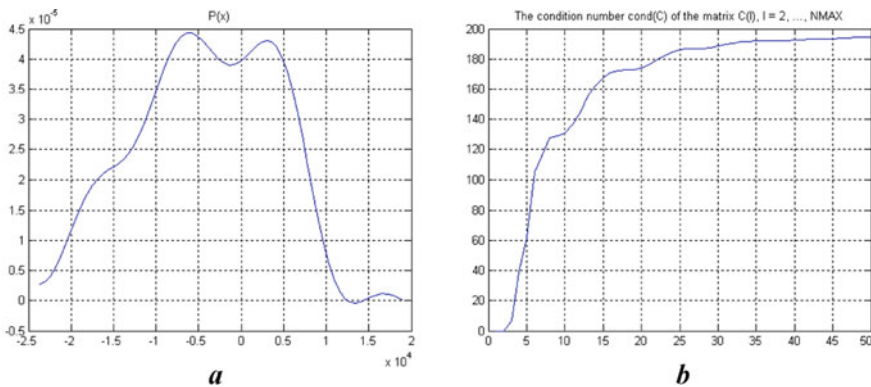


Fig. 3 Graph of the identified EGEG signal distribution density (10,818 samples) (a). Graph of condition number values for matrices to be solved by the identification algorithm $P(x)$. At the axis of abscissas there is the order of solvable SLAEs (b)

short-length segments, there are also unimodal densities but with increasing sample volumes higher modes begin to appear.

The distribution density of the EGEG signal differences, like the signal itself, is identified by the law with two evident modes. This result proves the correctness of the conclusions obtained in the study of fractals [20] and it is of great methodological significance for the theory of fractal processes in biomedicine and biology.

Figure 4 shows the identified distribution densities of the first and second increments for the EGEG signal (the number of samples is 10,817 and 10,816, respectively).

It will be recalled that the described results are based on the principle of the structural risk minimization and algorithmically consist of solving equivalent systems of linear equations by the regularization method [22] to find the expansion coefficients for the desired density $P(x)$. The expansion has a basis consisting of trigonometric functions.

The more complex is the kind of the distribution density, the more expansion members and the longer computation time are required. It is especially evident in polymodal distributions with sole modes. For the first and second increments we have

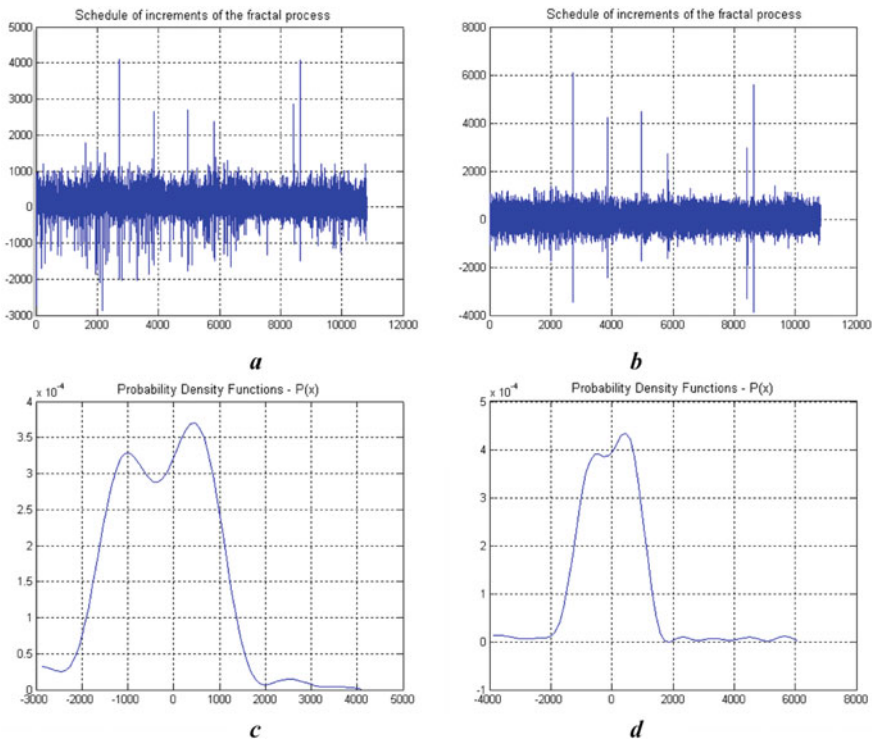


Fig. 4 Signals-increments for the EGEG-signal and their distribution laws: first increments (a, c); second increments (b, d)

Table 1 Identified parameters for the Gaussian process

Algorithm	H_p	m_x	k	t_{calc}, s
Gaussian process	1.34	-0.0079	2.058	3.09
First-order difference signal	1.67	-0.0029	2.062	3.09
Secondary-order difference signal	2.19	-0.0274	2.056	3.16

9 and 16 harmonics, Shannon entropy 8.30 and 8.19, and the entropy coefficient 1.97 and 1.70, respectively. The density identification time is about 45 s.

To verify the obtained results, we compare the fractal and Gaussian processes. The number of samples of the normal stationary process shall be 2401.

Table 1 and Fig. 5 illustrate the effectiveness of the used algorithms and the compliance of identified data with theoretical, classical results. The number of expansion coefficients is five in all three cases.

In Table 1 H_p is the Shannon entropy; m_x is a mathematical expectation; k is entropy coefficient; t_{calc}, s is the time for the distribution density calculation.

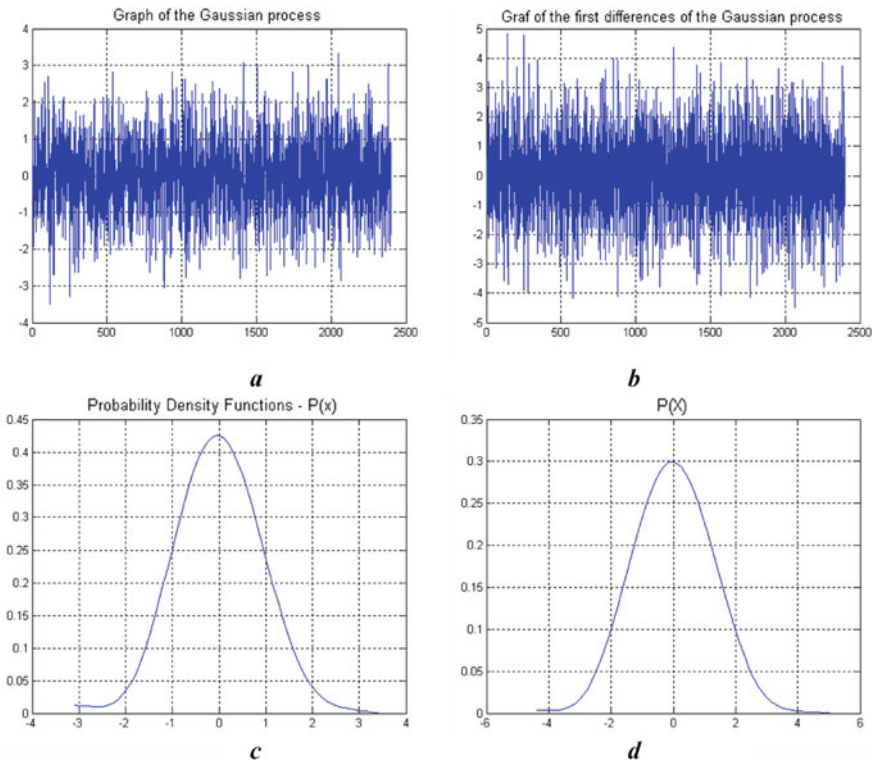


Fig. 5 Gaussian noise, signal of its first-order differences (2400 samples) and their distribution laws: Gaussian noise (a, c); first-order differences (b, d)

Thus, it may be concluded that: a random process with a Gaussian distribution (*randn* in MATLAB) is quite consistent with the theory. The process itself, its first- and second-order differences (increments of increments) have a normal distribution with the “zero” average and increasing dispersion, the entropy coefficient is close to the theoretical value $k = 2.066$.

3 SSA-Method Modification for Fractal Signals

The second aspect of the studies is related to the elaboration of algorithms for composing trajectory matrices of the EGEG signal for the further stochastic analysis of these matrix operators.

The EGEG signal represents a time series, to analyze thereof the SSA method may be used [23]. This method is based on the singular decomposition of trajectory matrices which columns are moving segments of a series of length L , where L is an integer—“the window length”. This method is effective for the analysis and forecasting of stationary series.

The SSA method is widely used for decomposing a time series into characteristic components: a slowly changing trend, periodic components, and noise. Nevertheless, there are quite many problems in the method, which do not have a regular and formalized solution. This is the distinguishing of a complex-nature trend (for example, superpositions of alternating ascending and descending segments), periodic components with close carrier frequencies, identification of noise of a significant level comparable to levels of useful components, etc.

All these problems predetermine the need to study series segments, trajectory matrices, embedding vectors, and noise components in the SSA method using stochastic analysis and the correct identification of distribution laws.

In the singular spectral analysis method, an arbitrary time series with real values forms the matrix X with elements taking the same values on “diagonals” $i + j = \text{const}$.

Furthermore, the SSA method used to the said series decomposes thereof into additive components (elementary restored series) and is based on the singular (SVD) decomposition of the trajectory matrix, i.e., provides information on components in the form of matrices of eigen and singular vectors and a set of eigenvalues [24].

We consider now some results of the trajectory matrix X analysis for the EGEG signal within the SSA method. The software algorithm implementation is made in the MATLAB package.

Because of limitation in work with high-order matrices in the SVD expansion the trajectory matrix is only calculated for half of the EGEG signal. It is assumed that: $N = 5409$ is the series length; $L = 2500$ —the number of the matrix lines; $K = 2910$ is the number of columns. The singular decomposition of the matrix X (2500, 2910) took 559 s. The result of calculating the singular spectrum is shown in Fig. 6.

For a segment of the EGEG signal with a smaller number of samples, the identification and singular decomposition time are far shorter. For example, we form a

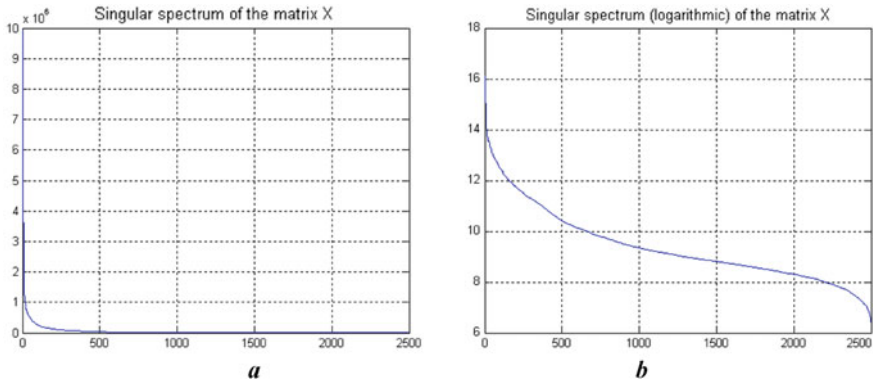


Fig. 6 The singular spectrum of the trajectory Hankel matrix X for the EGEG-signal (a); the same graph at logarithmic scale (b)

trajectory matrix for a series of 500 samples ($N = 500$; $L = 230$; $K = 271$). The condition number is thereat $\text{cond}(X) = 1.0394 \text{ e} + 04$; the rank is 230. The singular decomposition of the matrix X (230, 271) and the density function identification took half a second. The calculation of the singular spectrum showed similar logarithmic plots of singular spectra.

To check the local stability of the EGEG signal, a “violate” perturbation was performed above the given segment of a normally distributed random realization. As a result, one more property of the EGEG signal was identified, namely, its distribution density was relatively stable when mixed with other signals. Results and explanations are shown in Fig. 7.

The calculation of the singular spectrum and the identification of the distribution density for each of 230 singular values of the perturbed trajectory matrix X (230, 271) shows that the largest value is distributed with small dispersion and small singular

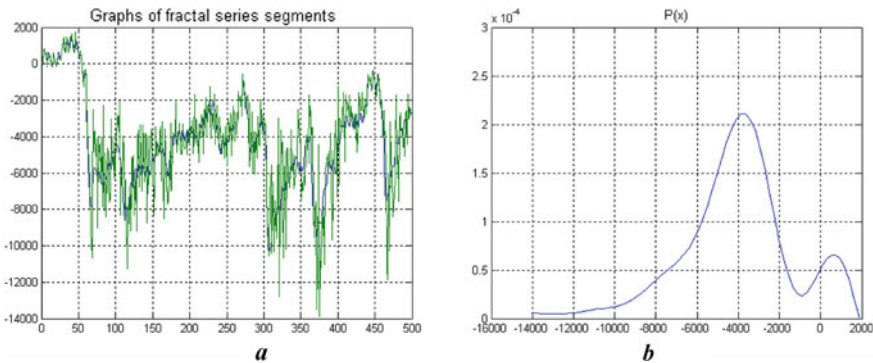


Fig. 7 Graph of a perturbed segment of fractal realization (500 samples) (a) and of identified distribution density of a perturbed segment (b)

values increase by about 30 times (normal distribution laws are identified for all singular values of the considered matrix). The condition number of the trajectory matrix decreases thereat proportionally with performing a certain regularization of the problem. Such a result is quite important for increasing the SSA method's stability.

The study of stochastic properties of embedding vectors for trajectory matrices was performed for the SSA method modification. The distribution densities were identified and the probabilistic evolution of 10 selected columns of the above trajectory matrix $X(230,271)$ was reconstructed.

Figure 8 (r-h series) shows distribution deformation examples and Table 2 contains some calculated parameters. It should be noted that the known papers on the SSA method lack such approaches.

In Table 2 N_{opt} is the optimal number of trigonometric functions achieved by minimizing the functional; H_p is the Shannon entropy; m_x is a mathematical expectation; k is entropy coefficient; t_{calc}, s is the time for distribution density calculation. The entropy column values are correctly correlated with the degree of polymodality taking into account that the last two lines characterize unimodal densities resembling mirror-like Levy distributions.

For comparison: the calculated Shannon entropy value of the total EGEG signal is 10.35, it illustrates its "wide" distribution form (Fig. 3).

Figure 9 shows examples of distribution deformation. The identified densities of the selected embedding vectors enable us to perform the Shannon, Renyi, or Kolmogorov entropy analysis. The data processing rate in tenths of a second.

The elaborated approach has a promising applied aspect since the quick identification of the embedding vector distribution will enable us to select reference components and reduce the trajectory matrix order, to compress a fractal signal without losing any important information. According to the evolution of distributions in the above example 10 columns may be left instead of 271, the matrix "henkelization" may be performed, and to obtain more than two times compression of the EGEG signal segment compression more than twofold may be obtained $[500/(230 + 10 - 1)]$. The segment will be reduced to 239 samples. The two- three-fold decrease in the EGEG signal length is essential, as it will reduce the time of diagnosing to make a proper decision and further emergency therapy. The measurement time of 10,818 points in the EEG signal is approximately 20 min. The gain in data compression may be several-fold increased by varying the window length L .

4 Conclusion

The suggested integrated approach enables us to carry out the more accurate and prompt comparative analysis of the EEG or other biosignals for patients in a normal state and with various-degree pathology.

The identified polymodal characteristics contain information on the dynamics in the state of the body organs and systems. The time for identifying one distribution

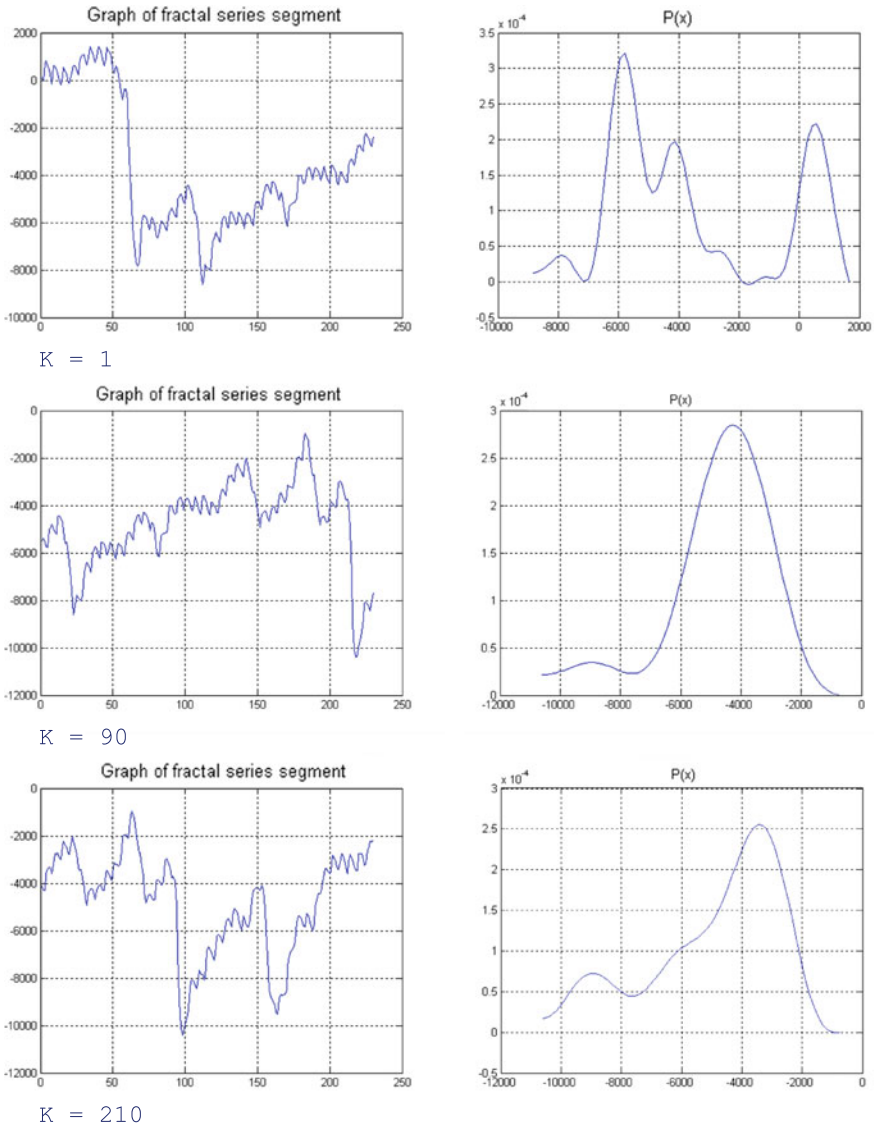


Fig. 8 The evolution of distributions of embedding vectors for the trajectory matrix $X(230,271)$. Graphs of realization segments ($K = 1, 90, 210$) of 230 samples (l-h series) and the relevant identified distribution densities of the EGEG signal—(r-h series). K is here the number of a trajectory matrix column

Table 2 Identified properties of embedding vectors for the trajectory matrix X

	N_{opt}	H_p	m_x	k	t_{calc}, s	Number of modes in the distribution
1	14	8.77	-3.61e+003	1.168	1.06	5
30	12	8.75	-4.33e+003	1.392	0.73	5
60	5	8.61	-4.71e+003	2.008	0.78	1
90	5	8.76	-4.73e+003	1.800	0.64	2
120	5	8.83	-4.80e+003	1.869	0.67	1
150	5	8.89	-4.97e+003	1.762	0.65	2
180	5	8.95	-5.08e+003	1.807	0.59	2
210	7	8.92	-4.89e+003	1.731	0.75	2
240	2	9.11	-5.05e+003	1.970	0.80	1
270	2	9.12	-5.13e+003	1.963	0.83	1

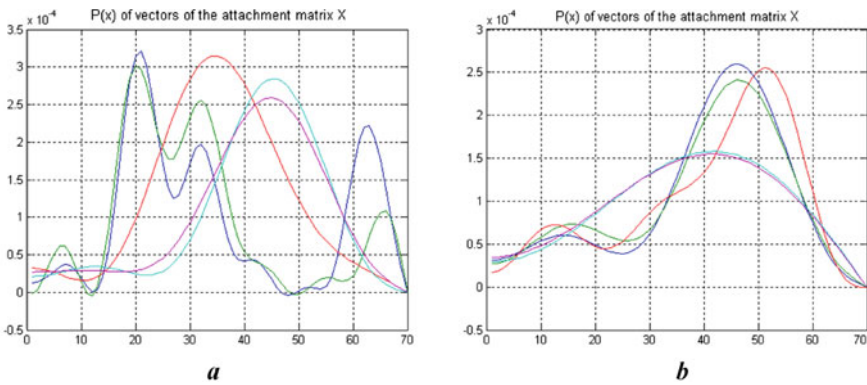


Fig. 9 Graphs of distribution densities for the first five embedding vectors ($K = 1, 30, 60, 90, 120$) (a); the following five embedding vectors ($K = 150, 180, 210, 240, 270$) (b); K is here the number of a trajectory matrix column. The sequence of densities is as follows: blue, green, red, turquoise, pink

depends on its complexity and varies from several tenths to 10–15 s with a sampling size of up to 5000 samples. It enables us to regard this method as an express-analysis.

Bioelectric signals recorded from the gastrointestinal tract organs, algorithms for processing, and programs for analyzing fractal properties form a set of a mathematical probabilistic model for the local biomedical CPS control system. A similar diagnostic set may be formed for the respiratory system of a patient.

The results of the analysis may be used for emergency diagnosis and therapy, in the creation of prospective expert systems for personalized medicine based on the big data systems.

The combination of SSA algorithms, wavelet analysis, the perturbation method, and independent components (ICA) will enable to identify the dominant and latent

factors in the course of studying any problems in the functioning of the GIT and other human organs and systems.

Acknowledgements The study was carried out with the financial support from the Russian Fund for Basic Research as part of a research project № 19-07-00926_a.

References

1. Parkman, H.P., Hasler, W.L., Barnett, J.L., Eaker, E.Y.: Electrogastrography: a document prepared by the gastric section of the American Motility Society Clinical GI Motility Testing Task Force. *Neurogastroenterol. Motil.* **15**(2), 89–102 (2005)
2. Gopu, G., Porkumaran, K., Neelaveni, R.: Investigation of digestive system disorders with cutaneous electrogastrogram (EGG) signal—an engineering approach. *Eur. J. Sci. Res.* **53**(2), 210–221 (2011)
3. Matsuura, Y., Takada, H.: Comparison of electrogastrograms in seated and supine positions using Wayland Algorithm. *Adv. Sci. Technol. Eng. Syst. J.* **4**(4), 42–46 (2019)
4. Koch, K.L., Stern, R.M.: *Handbook of Electrogastrography*. Oxford University Press, New York (2004)
5. Matsuura, Y., Takada, H., Yokoyama, K., Shimada, K.: Proposal for a new diagram to evaluate the form of the attractor reconstructed from electrogastrography. *Forma* **23**, 25–30 (2008)
6. Ohyabu, N., Narihara, K., Funaba, H.: Edge thermal transport barrier in LHD discharges. *Phys. Rev. Lett.* **84**, 103 (2000)
7. Metzler, R., Klafter, J.: The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Reports* **339**, 1 (2000)
8. Bouchaud, J.P., Georges, A.: *Phys. Reports* **195**, 127 (1990)
9. Shlesinger, M.F., Zaslavsky, G.M., Klafter, J.: Strange kinetics (Review Article). *Nature* **363**, 31–37 (1993)
10. Klafter, J., Shlesinger, M.F., Zumofen, G.: *Phys. Today* **49**(2), 33 (1996)
11. Montroll, E.W., Shlesinger, M.F.: On the wonderful world of Random Walks. In: Montroll, E.W., Lebowitz, J.L. (eds.) *Studies in Statistical Mechanics*, vol. XI, pp. 46–121. North-Holland PC, Amsterdam (1984)
12. West, B.J., Deering, W.: Fractal physiology for physicists: Lévy statistics. *Phys. Rep.* **246**, 1–100 (1994)
13. Ott, A., Bouchaud, J.P., Langevin, D., Urbach, W.: Anomalous diffusion in 'living polymers': a genuine Lévy flights? *Phys. Rev. Lett.* **65**, 2201 (1990)
14. Brockmann, D., Hufnagel, L.: Front propagation in reaction—superdiffusion dynamics: taming Lévy flights with fluctuations. *Phys. Rev. Lett.* **98**, 178301 (2007)
15. Zaslavsky, G.M., Edelman, M., Weitzner, H.: *Phys. Plasmas* **7**, 3691 (2000)
16. Risken, H.: *The Fokker Planck Equation*. Springer, Heidelberg (1984)
17. Haken, H.: *Advanced Synergetics*. Springer, Heidelberg (1983)
18. Gardiner, C.W.: *Handbook of Stochastic Methods in Physics, Chemistry and Natural Sciences*. Springer, Heidelberg (1983)
19. Kyprianou, A.: *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer, Berlin (2006)
20. Kulikov, V.: The identification of the distribution density in the realization of stochastic processes by the regularization method. *Appl. Mathem. Sciences* **9**(137), 6827–6834 (2015)
21. Klimontovich, Yu.L.: Nonlinear Brownian motion. *UFN* **164**(8), 811–844 (1994) ((in Russian))
22. Kulikov, V., Kulikov, A.: Regularization methods for the stable identification of probabilistic characteristics of stochastic structures. In: Kravets, A., Bolshakov, A., Shcherbakov, M. (eds.) *Cyber-Physical Systems: Advances in Design & Modelling*. Studies in Systems, Decision and Control, vol. 259. Springer, Cham (2020)

23. Golyandina, N., Nekrutkin, V., Zhigljavsky, A.: *Analysis of Time Series Structure: SSA and Related Techniques*. Chapman & Hall/CRC, London (2001)
24. Elsner, J.B., Tsonis, A.A.: *Singular Spectrum Analysis. A New Tool in Time Series Analysis*. Plenum Press, New York (1996)

Quality Research of the Interval Cetlin Method as a Component of the Cyber-Physical System of Continuous Monitoring of the Human-Operator State by ECG Signals



Alexey Khalaydzhii 

Abstract Chapter deals with the problem of quality assessment of the Cetlin method as a representative of the family of interval methods of automatic arrhythmia detection by ECG records from open MIT-BIH database using annotations on R-peaks. The author suggested new quality metrics, algorithms for their calculation in real-time mode, and formulated 5 approaches for solving the given problem with their use. The chapter shows a detailed analysis of the Cetlin method from the standpoint of each approach and reveals its advantages and limitations. Numerical values of the proposed quality metrics are obtained for all records. The author revealed linear correlations between anomalous R-peaks classes and the classes of Cetlin intervals and estimated the importance of each of them for the method as a classifier. The chapter shows parameters of the Cetlin method, variation of which allowed improving the quality of the method on the record, for which the worst results were initially obtained. The chapter suggests directions for further research and possible modifications of the Cetlin method for further improving its quality from the perspective of the described approaches and describes the possible architecture of the two-level model of the cyber-physical continuous monitoring system, using Cetlin method as one of the base components.

Keywords RR-interval · Arrhythmia · ECG record · Interval method · Cetlin method · Real-time mode · Quality metric · R-peak · MIT-BIH · Regression analysis · Cyber-physical system · Continuous monitoring · Human-operator state

A. Khalaydzhii (✉)

Bauman Moscow State Technical University, 5/1, 2-nd Baumanskaya street, Moscow, Russia
e-mail: aleksei_halaidzh@mail.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced Human-Centered Society*, Studies in Systems, Decision and Control 333, https://doi.org/10.1007/978-3-030-63563-3_20

1 Introduction

Due to the constant increase in the complexity of systems, processes that require increased concentration or are dangerous to humans are automated, while humans are assigned to the role of the operator—the managing unit of the cyber-physical system. The high risk associated with operator errors requires continuous monitoring of the operator’s state. To solve this problem, modelling using biosignals is used more and more [1], since they have greater predictive power. Moreover, it’s shown in [2–5] that instead of the signals their models can be analyzed with the same accuracy, and their parameters can be used as a “unique” code of a human, such as his signature. In practice, the ECG signal is most often used, since it is well studied and can be effectively obtained using special sensors and smart devices.

There are many methods of automatic analysis of ECG signal: representation of the signal as a superposition of elementary functions [2]; using of models of physical processes, forming the signal [5, 6]; analysis of information and energy features of the signal [7, 8]; application of the principles of synergetics and nonlinear dynamic systems [9]; use of statistics and machine learning methods [10–12]; application of neural networks architectures [13, 14]; use of hidden Markov [15, 16], etc.

One of the most important tasks solved by these methods is automatic arrhythmia detection [11]. Within this task, RR-intervals durations, heart rate variability (HRV) [10, 17], as well as other morphological features of the signal are primarily studied. These parameters are often used as features of a classification model, whose task is to assign R-peak or RR-interval to one of the previously known classes.

An important feature of the task is the need to work in real-time mode—in the process of receiving new values of the analyzed signals from the sensors or smart devices. This restricts the variety of methods and models, which can be used. Thus, the use of deep neural networks is inefficient due to high computational and time complexity, when very simple methods such as linear regression are much less accurate.

On the other hand, models that are trained on a pre-prepared set of data and determine average regularities don’t always have sufficient accuracy. Therefore, those models are of interest, that can take into account the specifics of a particular signal and adjust their parameters to it directly during their use. This requires either using models based on rules, which are checked by the signal adaptively, either continuing training the models on incoming data, that is difficult due to the lack of markup or its sparseness.

Finally, the interpretability of the result obtained by the model is important, since it allows us to determine the quality of predictions more reliably and provides more information for specialists. Approaches based on rules or their generalizations such as decision trees have this property since, for each forecast of the model, a decision rule can be explicitly written out.

The Cetlin linguistic method [18] allows automatically detect arrhythmias using a system of rules by ECG signal, analyzing the sequences of RR-intervals. Its simplicity and real-time operation are of practical value. Reference [18] describes a device, that

implements the Cetlin method in hardware, and was used to track the condition of patients in intensive care, which indicates its effectiveness. But at that time, there were no open databases that could objectively evaluate the quality of the method, compare it with others, and determine the impact of each of the parameters and choose the best ones. The main research goal is to assess the quality of the Cetlin method on signals from MIT-BIH.

To train models for automatic arrhythmia detection, according to the AAMI EC57:2012 standard, open databases, such as MIT-BIH, are used, where for each R-peak its class is given, which belongs to one of the clusters: N, SVEB, VEB, F, and Q. That makes it necessary to train such models and use such methods, that are able to predict the class of each R-peak, which does not allow them to be used directly for other tasks, such as assessment of a general human state or determining the presence of abnormalities on a daily or minute ECG recording.

The Cetlin method doesn't solve the problem of classification of R-peaks, so it's impossible to directly assess the quality of its work on signals from MIT-BIH. The inability to evaluate the quality of the method doesn't allow us to find the impact of parameters on its operation and identify its limitations and possibilities of modification. Therefore, it's necessary to develop new approaches of assessing the quality of the Cetlin method as a representative of the family of interval methods by labels on R-peaks.

Finally, since the actual ECG signals are noisy and come to the input of the method as a discrete sequence of amplitudes, it's necessary to solve the problems of filtering and segmentation in order to extract RR-intervals or other features. Since these procedures have a certain margin of error, there is a separate subproblem of the imposition of segmentation results with true labels. To solve this problem one needs to develop special algorithms for calculating quality metrics.

2 Approaches to Assess the Quality of Interval Methods

To assess the quality and make it possible to compare the performance of the interval method when changing its parameters, one can use the following approaches:

- estimation of interval quality metrics values
- estimating of peak quality metrics values
- regression analysis of the direct mapping between the initial distribution on peaks and the final distribution on intervals
- regression analysis of the reverse mapping between the final distribution on intervals and initial distribution on peaks
- qualitative analysis of results and interpretation of received errors.

The last approach will be shown in Sect. 3.3, all others described in detail below.

2.1 Interval Quality Metrics

When calculating this and subsequent metrics, the interval is considered to be a sequence of RR-intervals. The Cetlin method examines sequences of 3 RR-intervals. As part of the approach, the following metrics can be calculated:

- number of TP-intervals—abnormal intervals containing at least one abnormal peak
- number of TN-intervals—normal intervals, that don't contain abnormal peaks
- number of FP-intervals—abnormal intervals containing only normal peaks
- number of FN-intervals—normal intervals containing at least one abnormal peak.

By analogy with the metrics proposed by the AAMI EC57:2012 standard, one can introduce aggregated quality metrics: accuracy (Acc), sensitivity (Se), positive predictivity (P^+), and false-positive rate (FPR), calculated using the following formulas:

$$Acc = (TP + TN)/(TP + TN + FP + FN);$$

$$Se = TP/(TP + FN); P^+ = TP/(TP + FP); FPR = FP/(TN + FP)$$

2.2 Peak Quality Metrics

Interval quality metrics only indirectly take into account peaks classes, so in this approach, the following metrics are used:

- number of normal TPN-peaks—normal peaks inside normal intervals
- number of TPN-peaks for each of abnormal class clusters—abnormal peaks of each cluster inside abnormal intervals
- number of normal FPN-peaks—normal peaks inside abnormal intervals
- number of FPN-peaks for each of abnormal class clusters—abnormal peaks of each cluster inside normal intervals.

It should be noted that the third metric is less informative, since there may be normal peaks in abnormal intervals. By analogy with intervals metrics, one can introduce an aggregated quality metric—the percentage of correctly recognized peaks of each cluster of classes, calculated using the formula:

$$Acc_k = TPN_k/(TPN_k + FPN_k), k \in \{N, SVEB, VEB, F\}$$

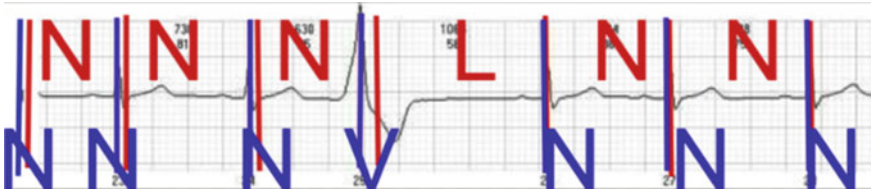


Fig. 1 Segmentation, that can be counted as FN and FP one after the other

2.3 Algorithmic Details of Metric Calculation

The correctness of metrics calculations depends on the segmentation quality. Therefore, it's necessary to correctly combine segmentation results with the true positions of the peaks from the markup. Figure 1 shows an example, where the formal calculation of the interval metrics increases values of FP and FN despite the correct operation of the method, since according to the markup, the abnormal V-peak is in the first normal Cetlin interval, but makes the following RR-interval Longer, that is, it affects the second Cetlin interval.

To count correctly in such situations, one should follow these rules:

- for all peaks except the last one, metrics are counted without changes
- if the current interval is normal and the last peak in it is normal, it is counted for the current interval, even if the peak is to the right of its end
- if the current interval is abnormal and the last peak is normal, it's also counted there
- if the current interval is normal and the last peak in it is abnormal, then to avoid the situation from Fig. 1, it is transferred to the next interval and is counted already there
- if the current interval and last peak are abnormal, the peak is counted for the current interval, but it is also transferred to the next one since it could affect both intervals.

In the last two rules, when analyzing the next interval, one needs to process the transferred peak as follows:

- if it wasn't calculated (rule 4), then it must be counted for calculating all metrics
- otherwise (rule 5)—if the interval is normal, that peak is ignored to avoid FN, and if not—it is counted only in the interval metrics.

2.4 Regression Analysis of Mappings Between Distributions

Previous approaches don't take into account the interval classes proposed by Cetlin and how they correspond to the peaks classes from MIT-BIH. Regression analysis

allows drawing conclusions, which of Cetlin's classes peaks of each cluster of classes from MIT-BIH mostly often fall in and vice versa.

There are the following Cetlin's classes of intervals: normal mode (NORM); extrasystoles without compensatory pause (without extending the next interval) (EXTRA-WITHOUT); heart block (BLOCK); paired extrasystoles (PAIR-EXTRA); extrasystoles with compensatory pause (EXTRA-WITH); paired extrasystoles with compensatory pause (PAIR-EXTRA-WITH); extrasystoles, after which there are blocks (EXTRA-BLOCK); paroxysmal tachycardia (TACHYC).

Peaks according to MIT-BIH markup can be one of the following classes: N—Normal (N/L/R); SVEB—Supraventricular ectopic beat (A/a/J/S/e/j); VEB—Ventricular ectopic beat (V/E); F—Fusion beat (F); Q—Unknown beat (Q).

The essence of the approach is to count the number of peaks and intervals of different classes, represent them as distributions, and analyze the mappings between the resulting distributions. Analysis of the direct mapping allows us to determine the sensitivity of the method to each of the peaks classes, and the reverse—to assess the quality of the method as a classifier and detector of the anomalies present in the signal section. Such analysis is more resistant to the alignment of signal splitting into intervals since it analyzes general distributions that make statistical sense on 30-min ECG signal recording.

The disadvantage of this approach is that it doesn't take into account the correctness of mappings from one class to another. This means that the analysis is not sensitive to random permutation of peak annotations of peaks annotations or intervals classes while preserving their total amount. At the same time, it allows us to look at the big picture, identify the limitations and capabilities of the method as a classifier and is useful when used together with other approaches.

3 Quality Research of the Cetlin Method

The proposed approaches were applied for the first time in the task of analyzing the quality of the Cetlin method on ECG records from MIT-BIH. The main subtasks, where they were used, are shown below.

3.1 Selection of Algorithms for Pre-processing of Signals

The signals in MIT-BIH are represented as a discrete sequence of samples with a frequency of 360 Hz. Before analyzing RR-intervals, one needs to perform filtering and segmentation in order to extract the positions of R-peaks, which are then used for calculating RR-intervals. To solve the first problem, median filters were used in accordance with the recommendations from [11].

Segmentation algorithms were chosen based on two assumptions: high segmentation accuracy and speed of operation since the Cetlin method is focused on real-time

Table 1 Comparison of algorithms for pre-processing of signals from MIT-BIH

Algorithm	Quality metric (%)							
	Acc	Se	P ⁺	FPR	Acc _N	Acc _{SVEB}	Acc _{VEB}	Acc _F
TMA, without filtering	96.25	90.09	77.90	3.05	91.95	82.35	93.99	0
TMA, with filtering	95.84	89.08	75.98	3.38	91.94	70.59	93.49	0
SWT, without filtering	95.08	88.45	71.86	4.15	91.49	64.71	93.17	0
SWT, with filtering	96.41	90.07	78.92	2.86	92.12	82.35	93.99	0

operating. To ensure the possibility of working in real-time mode, one implemented a scheme with storing signals samples within a time window, the size of which was determined in preliminary testing and qualitative analysis, and equals to 3 RR-intervals. This window was sufficient for determining the position of the first R-peak more precisely by most algorithms.

As a result of preliminary testing on the signal 101, two segmentation algorithms were selected: Two Moving Average (TMA) [19] and Stationary Wavelet Transform (SWT) [20]—and in 2 configurations: without filtering and with median filter according to recommendations from [11]. To check their correctness on the other signals, additional research was made, the results of which are shown in Table 1.

The table shows that both methods cope with the task with close accuracy for all quality metrics, but in the process of qualitative analysis, it was revealed that SWT works faster by 2 times, which is essential for real-time analysis, and therefore was chosen for further experiments. Filtering was also used since the results were better with it, but even without it, according to the table, the accuracy is high.

3.2 Analysis of the Results of the Method on Signals from MIT-BIH

As a part of the research, the Cetlin method was applied to all signals from MIT-BIH in accordance with the inter-patient partitioning scheme [17]. Table 2 shows the results of the method. It can be seen that the Cetlin method works quite well, which is confirmed by the high values of quality metrics. In particular, it recognizes the main types of rhythmic anomalies SVEB and VEB well.

However, the method makes a lot of false positives relate to inertia (using the average rhythm) when changing the rhythm and segmentation errors, as well as omits morphological anomalies (Acc_F) due to the lack of influence on the RR-intervals duration. Low values of Se/P⁺ were most often found in signals where there are few abnormal peaks, so skipping even one led to a large error.

Table 2 The final values of the quality metrics for ECG signals from MIT-BIH

Set of signals according to the inter-patient partitioning scheme	Quality metric (%)							
	Acc	Se	P ⁺	FPR	Acc N	Acc SVEB	Acc VEB	Acc F
Training set	84.1	78.8	63.2	14.3	77.3	84.8	84.0	24.7
Test set	84.7	75.4	62.3	12.8	80.4	75.6	82.4	42.8
All signals	83.5	81.8	63.9	16.0	74.2	89.2	86.0	5.4

Regression analysis was done with two models: random forest on 1000 decision trees to determine features importances and linear regression to determine correlations between features to find the existence of dependencies between peaks an intervals classes.

3.2.1 Direct Mapping Analysis

Table 3 shows the features' weights according to the random forest model obtained for direct mapping.

Based on the table, one can draw the following conclusions:

- VEB, N, and SVEB are important classes. The VEB class has the greatest weight since it changes the morphology very much and often leads to the appearance of long intervals, which leads to the appearance of different Cetlin's intervals
- N has large weight due to an unbalanced set towards the normal rhythm, morphological arrhythmias, and segmentation errors
- SVEBs are usually registered correctly since they often directly affect the rhythm and are determined by the method as extrasystoles
- F and Q are of low importance because they are rare and the method is not sensitive to them. The high value for class F is not due to its importance, but rather to its occurrence in different Cetlin's intervals, depending on adjacent peaks.

For the linear regression model, the following values of R^2 metric were obtained according to the inter-patient split [11]: for the training set—0.527; for the test set—0.682, and for all signals—0.58. A high value of this metric indicates that there are correlations between classes. To identify them. Linear regression coefficients were obtained (see Table 4). The meaningful correlations are given below:

Table 3 Importance of peaks classes according to the random forest model (1000 trees) for detecting anomalous interval by the Cetlin method

N	SVEB	VEB	F	Q
0.482	0.113	0.263	0.125	0.016

Table 4 Linear regression coefficients of a direct mapping between the initial distribution of peak classes and the final distribution of Cetlin’s intervals

R-peak class					Class of Cetlin’s interval
N	SVEB	VEB	F	Q	
0.869	-0.183	-1.35	3.888	-4.485	NORM
0.038	0.105	0.463	-0.582	-9.971	EXTRA-WITHOUT
0.05	0.178	0.690	0.015	-24.996	BLOCK
0.017	0.102	0.137	-0.34	13.801	PAIR-EXTRA
0.009	0.257	0.795	-1.349	4.567	EXTRA-WITH
0.001	0.138	0.055	-0.126	12.59	PAIR-EXTRA-WITH
0.001	0.071	0.175	-0.351	11.671	EXTRA-BLOCK
0.016	0.332	0.034	-0.156	-2.206	TACHYC

- peaks of class F often belong to the NORM class, since the method doesn’t feel them, which leads to a large number of FN
- the high weight of Q is caused by a small number of such peaks. The correlation with the NORM class is negative, meaning the method is sensitive to segmentation errors
- the correlation between SVEB and VEB and the NORM class is negative, that is the method is more likely to correctly classify them as anomalous
- N-peaks directly correlate with NORM class with a coefficient close to 1
- BLOCK and TACHYC are caused by rhythm changes, and not by their actual presence (inertia of the method). The division into blocks and extrasystoles is due to the position of the peak in the interval—extrasystoles can create a pair of Short and Long
- VEB negatively affect the NORM class and are defined as EXTRA-WITH, BLOCK, EXTRA-BLOCK, and EXTRA-WITHOUT, as they lead to Long intervals
- EXTRA-BLOCK—most often doesn’t precede the BLOCK and is just not accurately recognized EXTRA-WITH due to changes in the signal morphology due to the peak
- PAIR-EXTRA, EXTRA-BLOCK, and PAIR-EXTRA-WITH are often formed due to segmentation errors.

3.2.2 Analysis of Reverse Mapping

According to the random forest model (see Table 5) EXTRA-WITHOUT, PAIR-EXTRA and EXTRA-BLOCK have fewer weights, as they are often caused by errors. As in the direct mapping, PAIR-EXTRA has a low weight. EXTRA-WITH and PAIR-EXTRA-WITH have the greatest separating power since they are caused by both VEB and SVEB anomalies.

Table 5 The importance of the Cetlin's intervals for the method operating as a detector and classifier of peaks on examined interval

NORM	EXTRA-WITHOUT	BLOCK	PAIR-EXTRA	EXTRA-WITH	PAIR-EXTRA-WITH	EXTRA-BLOCK	TACHYC
0.166	0.019	0.108	0.020	0.236	0.227	0.055	0.169

Table 6 Coefficients of the linear regression of the reverse mapping between the final distributions of Cetlin's intervals and the initial distribution of peaks classes

NORM	Class of Cetlin's interval										R-peak class
	EXTRA-WITHOUT	BLOCK	PAIR-EXTRA	EXTRA-WITH	PAIR-EXTRA-WITH	EXTRA-BLOCK	TACHYC				
0.965	1.138	0.027	5.115	0.912	-3.8	-2.109	0.433	N			
0.011	0.426	-0.287	-2.079	-0.086	3.502	0.551	1.698	SVEB			
0.014	-0.371	0.966	-1.469	0.246	0.720	2.284	-0.929	VEB			
0.01	-0.191	0.292	-0.569	-0.072	0.559	0.271	-0.193	F			
0.0002	-0.0018	0.0005	0.0033	0.0002	0.0173	9.993	-0.0077	Q			

The linear regression model (see Table 6) reaches the following values of the R^2 metric: 0.865—for the training set, 0.95—for the test, and 0.865—for all signals, meaning that there is a significant linear relationship between classes:

- the model explains SVEB by interval classes of extrasystoles
- VEB is explained by intervals with pauses
- N correlates only with the NORM class
- the model F-peaks by the same features as VEB, i.e. F-peaks often fell in abnormal intervals if they were near the VEB peaks
- the method can be used as a basic arrhythmia presence detector and classifier, using only the first three classes
- the PAIR-EXTRA and EXTRA-BLOCK classes are not informative.

3.3 Selecting Values of Method Parameters to Improve Metrics

Worst of all, the method processed on signal 201. From the qualitative analysis of the examples in Fig. 2, it follows that there are changes in the “normal” rhythm, since the interval durations are similar for all of them, but the interval classes differ. Research of the method structure lets identify its parameters: X—interval duration for averaging, Y—the threshold of deviation from the “norm” for marking the RR-interval as Short, and Z—to Long. To improve the quality of the method’s operation

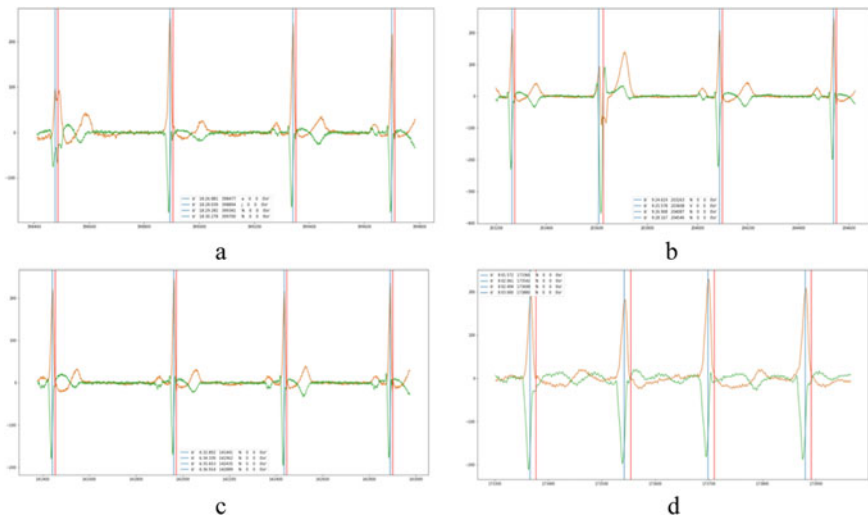


Fig. 2 Examples of Cetlin method operation on the signal 201: **a, b** normal mode [NNN]; **c** heart block [LLL]; **d** paroxysmal tachycardia [SSS]

Table 7 Comparison of the original and found configuration for signals

Signal number/configuration	Quality metric, % («→ means the absence of peaks in a signal)							
	Acc	Se	P ⁺	FPR	Acc _N	Acc _{SVEB}	Acc _{VEB}	Acc _F
201/25-25-2	52.23	63.7	47.57	57.1	40.52	99.27	47.47	50
201/30-10-2	59.6	93.69	53.61	69.71	21.25	100	90.91	50
221/25-25-2	81.12	94.58	75.66	33.86	39.49	–	94.19	–
221/30-10-2	71.93	92.45	67.56	52.57	28.38	–	91.92	–
222/25-25-2	67.39	97.93	47.12	45.32	47.03	98.34	–	–
222/30-10-2	56.04	99.59	39.83	62.01	32.57	100	–	–
231/25-25-2	65.26	100	1.09	34.87	65.02	100	100	–
231/30-10-2	61.61	66.67	0.99	38.54	61.37	100	100	–

on signal 201, a search is performed for such a configuration of X-Y-Z parameter values that achieves the best metric values.

The configuration for the signal 201 turned out to be 30-10-2, but Table 7 shows that it did not reach the best values for all quality metrics. This leads to the conclusion about the limitation of the Cetlin method and the need to use additional features, in particular, those, that take into account the morphology and solve the problem of inertia, as well as more complex models to achieve better results. At the same time, for the signals with similar properties found configuration has lowered metrics values. That indicates that there is no the only configuration of the method that gives the best result for all signals, and the need to implement an analogous procedure for searching parameters for each signal individually due to its uniqueness for each human to improve quality.

4 Discussion of Results and Directions for Further Research

The advantages of the Cetlin method are its simplicity and the ability to operate in real-time mode. However, the method is not intended for the classification of peaks but allows detecting the disturbances caused by them by analyzing the intervals. This caused the inability to evaluate it in accordance with AAMI EC57:2012 standard and the need to create new metrics and quality assessment approaches. Both of these tasks were completed and described in Sect. 2.

Section 3 shows that the method detects violations associated with peaks of the VEB and SVEB classes, but doesn't take into account the signal morphology and skips some of the abnormal peaks that have this type of violations. Also, the method has inertia—a change in heart rate variability leads to a large number of FPs, since the method does not have time to rebuild its knowledge of the “normal” interval

length. Finally, the Cetlin method doesn't allow to achieve the best quality even after parameter selection.

To overcome the limitations of the method, it makes sense to continue research in the following directions:

- consider other possible parameters of the method: sequence length (in basic configuration length equals to 3), as well as a different alphabet (in the original work, only the characters “S”, “L” and “N” were used). The specificity of this problem is the need to define new interval classes based on peaks classes from MIT-BIH, which can be solved using clustering methods
- add morphological features, from simple ones proposed in the Uspensky interval linguistic method [21] to more complex ones used in modern ensemble models [12]
- to take into account the changes of the signal state, it makes sense to add heart rate variability indicators as new features
- in addition to classification models, it's also useful to build regression models to get forecasts of the position of the next peak, the duration of RR-intervals or the values of the signal itself. Such models would allow improving the quality of segmentation, as well as to make a forecast based on current observations in terms of building cyber-physical systems of continuous monitoring
- use a multi-level approach to modeling, in which the results of basic models, one of which may be the Cetlin method, are combined to get and forecast changes of the general state (for example, by quorum). At the same time, differences between predictions and incoming true signal samples can be used as feedback to adjust the basic models to changes in the system state.

Solving the last problem will allow creating adaptive monitoring systems for cyber-physical systems using existing methods since the basic models can work relatively independently of each other and process a specific subset of biosignals and indicators from the sensors of technical subsystems. When building such a system, an important requirement for all components is their capability to work in real-time mode, as well as the interpretability of the results. The Cetlin method meets both requirements and can be used as one of the basic components of such a system. As one of the simple feedback options for the Cetlin method, one can use the signal of general state change to reset the “norm” state of the method or adjust parameter values.

5 Conclusion

The Cetlin method is a simple but effective method for automatic arrhythmias detection, which analyzes not R-peaks, as most modern methods do, but the time intervals of the ECG signal, which refers it to the family of interval methods. The proposed metrics allow not only to evaluate the quality of interval methods for automatic

arrhythmia detection using existing signal bases but also to compare their effectiveness with other methods. This possibility encourages the creation of new interval methods for analyzing heart rate variability, tracking its disorders, and assessing the human state not by individual R-peaks, but by RR-intervals and other long time intervals, which simplifies the design process of such methods, as well as allows to obtain more interpretable models and results of their work for specialists.

Acknowledgements The research was carried out with the financial support of the Innovation Assistance Fund (grant UMNİK № 11303GU/2018).

References

1. Buldakova, T.I., Suyatinov, S.I.: Hierarchy of human operator models for digital twin. In: 2019 International Russian Automation Conference (RusAutoCon), pp. 1–5. IEEE, Sochi, Russia (2019). <https://doi.org/10.1109/RUSAUTOCON.2019.8867602>.
2. Fainzilberg, L.S.: Technology of constructing the telemedical system based on the generative model of creation of artificial ECG of realistic form. *Clin. Inform. Telemedicine* **8**(9), 89–98 (2012)
3. Cai, H., Venkatasubramanian, K. K.: Patient identity verification based on physiological signal fusion. In: 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), pp. 90–95. IEEE, Philadelphia, PA (2017). <https://doi.org/10.1109/CHASE.2017.65>.
4. Buldakova, T.I., Krivosheeva, D.A.: Data protection during remote monitoring of Person's State. In: Dolinina O., Brovko A., Pechenkin V., Lvov A., Zhmud V., Kreinovich V. (eds.) ICIT 2019: Recent Research in Control Engineering and Decision Making, pp. 3–14. Springer, Saratov, Russia (2019). https://doi.org/10.1007/978-3-030-12072-6_1
5. Permyakov, S.A., Kuznetsov, A.A., Sushkova, L.T., Chepenko, V.V.: Informational model of ECG amplitude-phase coupling based on statistical approach. *Infocommunication Technol.* **3**(15), 261–268 (2017)
6. Karavayev, A.S., Ishbulatov, Y.M., Kiselev, A.R.: Model of the human cardiovascular system with an autonomous loop of regulation of average blood pressure. *Human Physiol.* **1**(43), 70–80 (2017)
7. Barquero-Perez, O., Goya-Esteban, R., Caamano, A.J., Sarabia-Cachadina, E., Martinez-Garcia, C., Rojo-Alvarez, J.L.: Evolution of the heart rate variability complexity during Kangchenjung Climbing. *Comput. Cardiol.* **42**, 1033–1036 (2015)
8. Acharya, UR., Fujita, H., Sudarshan, Ghista, D.N., Eugene L.W.J., Koh, J.E.W.: Automated Prediction of sudden cardiac death risk using Kolmogorov Complexity and recurrence quantification analysis features extracted from HRV Signals. In: 2015 IEEE International Conference on Systems, Man, and Cybernetics, pp. 1110–1115. IEEE, Hong-Kong (2015). <https://doi.org/10.1109/SMC.2015.199>
9. Melkonian, D., Barin, E.: Non-linear dynamics and fractal composition of human electrocardiogram: a proposed universal formula for ECG waveforms. *Fractal Geometry Nonlinear Anal. Med. Biol.* **2**(2), 1–9 (2016)
10. Almeida, R., Dias C., Silva, M.E., Rocha, A.P.: ARFIMA-GARCH modelling of HRV: clinical application in acute brain injury. In: Complexity and Nonlinearity in Cardiovascular Signals. Springer, Berlin (2017). doi: https://doi.org/10.1007/978-3-319-58709-7_17.
11. Luz, E.J.S., Schwartz, W.R., Camara-Chavez, G., Menotti, D.: ECG-based heartbeat classification for arrhythmia detection: a survey. *Comput. Methods Programs Biomed.* **127**, 144–164 (2016)

12. Mondéjar-Guerra, V., Novo, J., Rouco, J., Penedo, M.G., Ortega, M.: Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers. *Biomed. Signal Process. Control* **47**, 41–48 (2019). ISSN 1746–8094. <https://doi.org/10.1016/j.bspc.2018.08.007>.
13. Warrick, P., Homsy, M.N.: Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks. In: *Computing in Cardiology (CinC)*, pp. 1–4. IEEE (2017). <https://doi.org/10.22489/CINC.2017.161-460>
14. Hou, B., Yang, J., Wang, P., Yan, R.: LSTM Based Auto-encoder model for ECG arrhythmias classification. *IEEE Trans. Instrum. Meas.* **69**(4), 1232–1240 (2019). <https://doi.org/10.1109/TIM.2019.2910342>
15. Bouchikhi, S., Boublenza, A., Chikh, M.A.: Discrete hidden Markov model classifier for premature ventricular contraction detection. *Int. J. Biomed. Eng. Technol.* **17**(4), 371–386 (2015). <https://doi.org/10.1504/IJBET.2015.069403>
16. Pimentel, M.A., Santos, M.D., Springer, D.B., Clifford, G.D.: Hidden semi-Markov model-based heartbeat detection using multimodal data and signal quality indices. In: *Computing in Cardiology*, pp. 553–556. IEEE, Cambridge, MA (2014)
17. Murugan, S., Selvaraj, J., Sahayadhas, A.: Detection and analysis: driver state with electrocardiogram (ECG). *Phys. Eng. Sci. Med.* **43**, 525–537 (2020). <https://doi.org/10.1007/s13246-020-00853-8>
18. Cetlin, M.L., Gorohov, Y.S., Matusova, A.P., Melnikova, V.A., Tarantovich, T.M., Shabashov, V.M.: Device for recording and diagnostics of rhythm of cardiac activity. *News Higher Educ. Institutions* **1**(4), 165–172 (1961)
19. Elgendi, M., Jonkman, J., De Boer F.: Frequency bands effects on QRS detection. In: *Proceedings of the 3rd International Conference on Bio-inspired Systems and Signal Processing BIOSIGNALS 2010*, vol. 1, pp. 428–431. Institute for Systems and Technologies of Information, Control and Communication (INSTICC), Portugal (2010) doi:<https://doi.org/10.5220/0002742704280431>
20. Kalidas, V., Tamil, L.S.: Real-time QRS detector using stationary wavelet transform for automated ECG analysis. In: *2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 457–461. IEEE, Washington (2017). <https://doi.org/10.1109/BIBE.2017.00-12>
21. Uspenskiy, V.: Diagnostic system based on the information analysis of electrocardiogram. In: *Proceedings of MECO 2012. Advances and Challenges in Embedded Computing*. Bar, pp. 74–76. Montenegrin Association for New Technologies, Montenegro (2012)

Technology for Predictive Monitoring of the Performance of Cyber-Physical System Operators Under Noise Conditions



Igor Ushakov, Alexey Bogomolov, Sergey Dragan, and Sergey Soldatov

Abstract The methodological bases of predictive monitoring of the operability of cyber-physical systems operators in the conditions of aircraft noise exposure are considered based on the use of personalized acoustic hazard indicators. New opportunities for optimizing the use of individual and collective protection against aircraft noise and preserving the health of operators of cyber-physical systems whose professional activities are associated with its impact are shown.

Keywords Monitoring of working conditions · Personalized acoustic monitoring · Cyber-physical system · Acoustic safety · Acoustic hazard indicator · Health riskometry

1 Introduction

According to modern data, more than half of cyber-physical systems operators work under conditions of increased exposure to acoustic vibrations (noise, infrasound, and ultrasound), and about 25% of industrial personnel jobs do not meet noise standards

I. Ushakov (✉) · S. Dragan
State Scientific Center of the Russian Federation, Burnazyan Federal Medical Biophysical Center,
46, Zhivopisnaya St., Moscow 123103, Russia
e-mail: iushakov@fmbcfmba.ru

S. Dragan
e-mail: s.p.dragan@mail.ru

A. Bogomolov
St. Petersburg Institute of Informatics and Automation of the Russian Academy of Sciences, 39,
14th line of Vasilyevsky Island, St. Petersburg 199178, Russia
e-mail: a.v.bogomolov@gmail.com

S. Soldatov
Central Research Institute of the Air Force of the Ministry of Defense of Russia, 12-A,
Petrovsky-Razumovskaya Alley, Moscow 127083, Russia
e-mail: soldatov2304@yandex.ru

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

A. G. Kravets et al. (eds.), *Society 5.0: Cyberspace for Advanced
Human-Centered Society*, Studies in Systems, Decision and Control 333,
https://doi.org/10.1007/978-3-030-63563-3_21

[1, 2]. This leads to the fact that professional hearing loss occupies the first place in the structure of occupational diseases of operators of cyber-physical systems, with a continuing trend of increasing incidence [3].

Note the negative growth in the number of cases of occupational hearing loss of operators of cyber-physical systems due to both techno-genetic processes of development of cyber-physical systems, associated with an increase in the capacity of industrial equipment, increased speed and density of traffic flow, and shortcomings of the legislation in the field of noise control, the use of ineffective remedies, a lack of reliable methods of prediction of noise levels around the regulated frequency range, insufficient elaboration of medical and technical measures to create safe working conditions for operators of various cyber-physical systems [4–7].

Production and transport noise accompanying operators of cyber-physical systems contain mainly infrasound and low frequencies in their spectrum. In most cases, it is not possible to ensure the acoustic safety of operators of cyber-physical systems by reducing the noise level in the source of education, since reducing the power of the equipment of cyber-physical systems reduces its performance. Therefore, individual and collective noise protection technologies are the most acceptable way to ensure the acoustic security of cyber-physical system operators [8, 9].

The problem of acoustic safety of professional activities of cyber-physical system operators is compounded by the lack of standard means of individual and collective protection against noise, which provide effective protection for operators of cyber-physical systems. Noise protection measures used by operators of cyber-physical systems tend to reduce the functional comfort of professional activities. This, in turn, leads to improper use or non-use of security tools, which significantly increases the risks to the health and performance of operators of cyber-physical systems [10–12]. In addition, the passport characteristics of the acoustic efficiency of noise protection devices obtained under special conditions are significantly overestimated in comparison with the real efficiency due to their improper use, as well as a decrease in the elastic–plastic characteristics of earpads during operation [8, 9]. It should also be noted that the protective properties of anti-noise in the low and infrasound frequencies are practically absent [9].

One of the promising directions for ensuring the acoustic safety of cyber-physical system operators is the implementation of technologies for personalized acoustic monitoring based on the use of acoustic hazard indicators that function in real-time [7, 14].

2 Personalized Acoustic Hazard Indicators

A personalized acoustic hazard indicator is a personal information and measuring device that provides real-time monitoring of the acoustic environment over the entire normalized frequency range with information about its danger. In the calculated hazard assessment, determined by the amount of risk to the health and performance of operators of cyber-physical systems, it is possible to take into account individual

health characteristics, characteristics of the applied means of protection, and features of the professional activity of operators of cyber-physical systems (duration of work, motor activity, etc.). Methods of such calculations for various socio-professional groups of operators of cyber-physical systems are widely presented in the literature [15–19].

Informing operators of cyber-physical systems of the danger is carried out using light, sound, or vibration signals (for example, green continuous signal—low risk; yellow continuous signal—expressed risk; red continuous signal—high risk; red pulsating signal—very high risk). In addition, using a miniature tableau, it is possible to display the estimation of the reserved time of maintaining operability by operators of cyber-physical systems in specific acoustic conditions [7].

Information obtained with the help of personalized acoustic hazard indicators can be accumulated in the medical dosimetric register and periodically (depending on the specifics of professional activity) processed in the interests of forming personalized recommendations for preserving the health of operators of cyber-physical systems.

This information can be transmitted in real-time for centralized monitoring of the working conditions of operators of cyber-physical systems, allowing you to take preventive measures aimed at preserving the health and maintaining the performance of operators of cyber-physical systems.

A typical personalized acoustic hazard indicator includes a meter, a signal and information Board, a computer, a power supply unit, an amplifier unit, a filter unit, a detection unit, a micro-computer, a control unit; an information input unit, a RAM, a permanent storage device, and a switch [7].

Cumulative effects of acoustic factors of the professional activity of operators of cyber-physical systems are taken into account using the “sliding window” procedure, the duration of which is determined based on the features of professional activity and the features accompanying its performance of acoustic factors (intensity, exposure time, the spectral composition of acoustic vibrations) [3, 20–22].

Personalized acoustic hazard indicators for operators of cyber-physical systems are placed in the area of personnel workplaces or fixed on clothing (uniforms, equipment), without interfering with the professional activities of operators of cyber-physical systems. The meter included with the indicator oriented to be able to detect noise and signal information, the scoreboard must be in sight of the operator of a cyber-physical system. Taking into account the level of development of Industry 4.0 technologies, advances in microelectronics and information and telecommunications technologies, a personalized acoustic hazard indicator for operators of cyber-physical systems can be implemented in a miniature version.

At the beginning of the professional activity, the personalized acoustic hazard indicator is turned on (activated). From this point on, it starts recording acoustic environment indicators: based on the current sound pressure level and the characteristics of the accumulated noise dose, the acoustic hazard of the surrounding environment is determined. If the signal and information Board is green, you can not use anti-noise. When the tableau lights up yellow or red, you should apply anti-noise (in accordance with the characteristics of your professional activity) until the tableau lights up green again.

Taking into account the presence of a permanent storage device as part of a personalized acoustic hazard indicator, it can act as a personalized acoustic dosimeter, preserving the user's accumulated zero noise dose and other indicators that characterize the cumulative effects of a dangerous acoustic environment [14].

If you want to determine the danger of acoustic environment taking into account the protective properties of the PPE against noise meter should be placed on the inner surface of this tool, providing, if possible, his being at a distance from the body (clothing, protection). In this case, the connection of the meter with the other components of the personalized indicator should be provided via a wireless Protocol, ensuring the visibility of the protection tool when the employee performs professional tasks.

For centralized monitoring of the performance of a group of operators of cyber-physical systems in real-time, the switch of a personalized acoustic hazard indicator must be configured with the ability to implement the transfer of information over a wireless Protocol to the monitoring point.

If the professional activity of operators of cyber-physical systems is carried out in a large room (for example, an Assembly shop), the acoustic environment in which can be considered homogeneous (for example, it is determined by an external noise source and the noise of the ventilation system), use a single personalized indicator, the size of the signal and information Board is selected taking into account the visibility of all operators of cyber-physical systems. In this case, the indicators of adverse cumulative effects of noise exposure for a particular operator of the cyber-physical system should be determined taking into account its working schedule (being in the field at specific times).

The same approach applies to the use of collective protection of operators of cyber-physical systems from noise. To do this, personalized NY hazard indicator acoustic environment equipped with two scoreboards: one placed outside collective protection (for signal workers to go inside remedies) and the other inside collective protection (alarm about the possibility to leave the remedy for the normalization of the acoustic environment). In this case, it is acceptable to use several spatially separated meters as part of a personalized acoustic hazard indicator with the determination of indicators of the current acoustic situation based on the "most dangerous" readings of one of the meters.

3 The Methodological Approach to the Selection of Personal Protection Equipment for Operators of Cyber-Physical Systems from Noise

Personal protective equipment for operators of cyber-physical systems against noise can be divided into groups [8, 9, 16].

1. Personal protective equipment for the hearing organ to block the air path of sound propagation. These include anti-noise headphones, earplugs, which are

most often used in practice. Anti-noise of this group can provide a reduction in sound levels up to 30 dB mainly in the medium and high-frequency sound range. To increase efficiency, a combination of anti-noise headphones with earplugs is recommended, which allows you to increase their efficiency by 10 dB, as well as increase the sound absorption of low frequencies to 10–15 dB.

When choosing anti-noise from this group, follow the following guidelines:

- if the sound level in the workplace is up to 100 dBA, and the noise spectrum is dominated by medium and high frequencies, then you must use anti-noise headphones or earbuds;
 - when the sound level at the workplace is 100–110 dBA, and the noise spectrum is wide with a maximum in the mid-and high-frequency range, you should use a combination of anti-noise headphones with earbuds.
2. Personal protection of the head from noise to block the air and bone pathways of sound propagation. This includes an anti-noise helmet. Protection of the organ of hearing is provided comprising hat-mA, and the protection of the bony structures of the head—actually the helmet itself, multilayer structure which contributes to sound absorption. The anti-noise helmet itself can be soft, made of various fabrics and materials (foam, leather); semi-rigid, which is based on various types of polymers (plastics); hard—using metal and Kevlar.

It is known that at high levels, the acoustic signal enters the cochlea both through the outer ear and the bone-tissue structures of the skull. If the bone and tissue hearing thresholds exceed the air ones by 20–40 dB, then at the maximum permissible sound levels of 80 dBA, extra aural personal protective equipment for the head is recommended to be used at sound levels above 100–110 dBA. The acoustic efficiency of anti-noise helmets varies in a wide range (from 30 to 50 dB) and mainly depends on their design.

When choosing the antinoise of this group should be guided by the following: the sound level in the workplace is over 110 dBA and the noise spectrum is dominated by medium and high frequencies.

3. Means of individual protection of the trunk from noise. They are designed to protect the chest and abdominal organs from noise. The most commonly used means of noise protection of this class is an anti-noise vest. Vest protection is achieved by using a combination of materials with different sound-absorbing properties.

At high sound levels above 120 dBA, the operator of the cyber-physical system, located in the acoustic field, feels the vibration of the entire body. The mechanism of this phenomenon is associated with active irritation of the mechanic receptors and proprioceptors, which contributes to the activation of the Central and autonomic nervous systems. The effect of acoustic fields of such intensity contributes to the development of subjective discomfort in operators of cyber-physical systems in the form of vibrations of the thoracic and abdominal organs. The presence of low and infrasound frequencies in the noise spectrum significantly increases the sensation of

vibration up to the appearance of pain in the nasopharynx, speech disorders, feelings of fear, dizziness, nausea, drowsiness, etc.[2–4].

To protect operators of cyber-physical systems from high-intensity noise, it is proposed to use an anti-noise vest, which is designed to cover the surface of the chest and abdominal cavities and reduce the level of vibration caused by an acoustic wave. Laboratory studies have shown that the sound absorption of the anti-noise vest reaches 10–15 dB in the region of medium and high frequencies [9].

Choose the anti-noise of this group should be at the sound level at the workplace of operators of cyber-physical systems over 120 dBA and the presence of low frequencies in the spectrum with a sound pressure level of more than 100 dB. The anti-noise vest is recommended to be used together with the anti-noise helmet.

4 Methodology for Determining the Boundaries of Territories Potentially Hazardous in Terms of Noise Exposure

Currently, for the calculation of “noise contour” is used three Methods: a method based on the use of the point of closest approach to the IP-source of the noise, the method of segmentation, and the modeling method. The most common practice used segmentation method based on a database of noise characteristics and its sources containing analytical expressions dependencies noise-power-distance [1, 11, 19].

In the absence of such databases apply the “direct” approach. The standards of the sanitary norms normalized noise parameters in residential areas are the equivalent sound level (LA_{eqv}) and maximum sound level (LA), measured in dBA.

Actual measurements performed using the measuring path including a pressure transmitter (microphone), measuring (matching) amplifier, and recorder providing registration of sound pressure signals.

As the sound pressure transducer is usually used with microphones’ natural frequency over 30 kHz and a dynamic range is not worse than 80 dB (the nominal frequency response of the microphone was not already 1–20,000 Hz unevenness ± 1 dB) with omnidirectional diagram.

Measuring sensor signals arrive at the preamplifier and then to normalizing amplifier an analog–digital converter unit in advance and storing signals (recording device). For an array of input signals to the recorder using an analog-to-digital converter with a dynamic range of at least 70 dB and the number of bits is not less than 12, as well as software to the noise analyzer level recorder and the statistical analyzer.

Checkpoints (number, coordinates) is determined directly by measurement organization. All measurements were carried out in autonomous and automatic operation in the open air (outside the audible shadow) or at least 2 m away from the reflective structures of buildings. When measured control the absence of obstacles, distorting the sound field. Designated measure noise on a flat surface under conditions of excessive lack of excessive sound attenuation, the climatic conditions (atmospheric

pressure, air temperature, humidity, wind speed) must meet conditions acceptable functioning of the means of measuring and auxiliary devices. The measurements are not carried out during a storm and atmospheric precipitation, at a wind speed of 5 m/s. If the wind velocity exceeds 1 m/s, then the measurements must be used windscreen microphone [22].

Work center a microphone at a height of 1.2–1.5 m above the earth's surface or the surface structures, orienting the axis of the microphone in the direction of the anticipated maximum radiation noise.

The measurements are not carried out under unfavorable combinations of temperature and relative humidity of the ambient air (when the attenuation of sound in air to one-third octave band exceeds 10 dB for 100 m with a center frequency of 8 kHz), and in cases where the background noise is different from the source of the noise levels of less than 10 dB.

Measurements were carried out to determine the sound pressure levels in octave frequency spectral bands with central frequencies of 2, 4, 8, 16, 31.5, 64, 125, 250, 500, 1000, 2000, 4000, and 8000 Hz. Measurements performed by the simultaneous transformation of sound pressure in residential areas to electrical signals, with subsequent recording and processing in two stages:

- the first stage (under natural conditions) provide the registration of sound pressure signals to control points of measurement;
- at the second stage, the processing of the registered information is carried out, determining the root mean square values sound pressure in the octave bands, maximum sound levels and the calculated values of the equivalent sound level in the day and night.

Dissimilar primary measurements of sound level results in the same distance from the measurement reference point to the noise source of 1000 m, in accordance with the expression:

$$L_{Aop} = L_A + 20 \times \lg(R/R_{op}),$$

wherein L_{Aop} —sound level dB (maximum or equivalent) for the distinction 1000 m, L_A —sound level dB (maximum or equivalent), $R_{op} = 1000$ m, R —distance to the noise source.

The obtained values are averaged normalized noise parameters from different types of sources in its different modes of operation are the basis for calculating the area boundaries, potentially hazardous levels of noise exposure. Is applied to map the lines of equal maximum and equivalent sound levels, at which the outer boundary and beyond audio noise does not exceed the permissible. Bounds checking circuits sound levels is performed using actual measurements.

5 Methods for Evaluating Acoustic Security of Cyber-Physical System Operators

Acoustic safety is understood as the state of human protection from the adverse effects of noise in the process of life [6, 22].

The calculation of the acoustic safety indicator for operators of cyber-physical systems—the acoustic safety coefficient—is based on a comparison of the measured acoustic environment indicators for noise and infrasound in places of human activity with the maximum permissible levels established by sanitary standards.

The acoustic safety factor of personnel (k , dB) is defined as

$$k = 20 \lg \frac{\sum_{i=1}^n 10^{\Delta_i/20}}{n},$$

where $n = 19$ —is the number of used indicators of the acoustic environment, Δ_i —is the difference between the maximum permissible level and the corresponding measured value of the i th indicator of the acoustic environment.

When all the indicators of the acoustic environment used for its calculation are equal to the maximum permissible, $k = 0$. The lower the measured values of the indicators of the acoustic situation compared to the maximum permissible levels, the greater the coefficient k and, accordingly, the better the acoustic security of cyber-physical system operators.

When calculating k , all indicators of the acoustic environment are considered equal, and the correctness of their convolution into an integral indicator is determined by summing the Δ_i values on a linear scale with the subsequent translation of the result into a logarithmic scale.

Many indicators of the acoustic environment used to calculate the value of the coefficient k , consists of subsets of fixed and variable indicators.

A subset of fixed indicators includes six indicators of the acoustic environment:

1. Equivalent sound level A per working shift ($L_{p,Aeq,8h}$, dBA), measured with frequency correction on the “A” scale and/or calculated for 8 h of a working shift. The standard equivalent maximum permissible sound level at the workplaces of specialists in most sectors (sub-sectors) of the economy is 80 dBA (for certain sectors of the economy an equivalent noise level at workplaces of up to 85 dBA is allowed provided that the acceptable risk to the health of workers is confirmed and a set of measures is taken to minimize risks health of operators of cyber-physical systems.).
2. The maximum sound level A , measured with temporary correction “slowly” $S = 1$ s ($L_{S,Amax}$, dBA), the maximum permissible level of which is 110 dBA.
3. The maximum sound level A , measured with temporary correction “pulse” $I = 40$ ms ($L_{I,Amax}$, dBA), the maximum permissible level of which is 125 dBA.
4. Peak sound level C —peak sound level corrected according to the “C” scale ($L_{p,Cpeak}$, dB), the maximum permissible level of which is 137 dB. It should

be noted that when registering pulsed or tonal noise, the maximum permissible levels are reduced by 5 dB [9, 18].

5. The equivalent overall level of infrasound per shift ($L_{p,ZI,eq,8h}$, dB) is the sound pressure level in the frequency range 1.4 ... 22 Hz. The maximum permissible level of this indicator at workplaces is established: in vehicles $L_{p,ZI,eq,8h} = 110$ dB, work of varying severity $L_{p,ZI,eq,8h} = 100$ dB and work of varying degrees of intellectual and emotional tension $L_{p,ZI,eq,8h} = 95$ dB.
6. The maximum general level of infrasound, measured with time correction S (slowly) in the frequency range 1.4 ... 22 Hz (L_{ZFmax} , dB), the maximum permissible level of which is $L_{ZFmax} = 120$ dB.

A subset of variable acoustic environment indicators includes up to thirteen indicators, the number of which is determined by the number of octave frequency bands in which to ensure or evaluate the acoustic safety of cyber-physical system operators:

1. equivalent sound pressure levels per shift in the octave frequency bands of 2, 4, 8, 16 Hz ($L_{p,1/1,eq,8h}$, dB). The maximum permissible levels of this indicator are differentiated for three types of work: in vehicles, work of varying degrees of severity, work of varying degrees of intellectual-emotional tension; at the same time, the maximum current total level of infrasound should not exceed 120 dB, and with a reduced working day (less than 40 h per week), the maximum permissible levels are applied without change;
2. the maximum allowable sound pressure levels in the octave frequency bands of 31.5, 63, 125, 250, 500, 1000, 2000, 4000, 8000 Hz. Extremely acceptable levels of this indicator are determined by the type of labor activity and the characteristics of the workplace [3, 7, 19]. It should be noted that at present these indicators of the acoustic situation are not standardized. However, to objectify the acoustic impact, they must be taken into account due to the fact that, as a rule, the spectrum of industrial, industrial, and transport noise along with high frequencies contains infrasound and low frequencies. The application of the standards leads to the fact that the use of the "A" scale for measuring noise involves filtering the noise, leading to an underestimation of its intensity at low frequencies. At a frequency of 500 Hz, the difference between the sound level and the sound pressure level will be 3 dB, and at a frequency of 22 Hz this difference will reach 50 dB. Therefore, in order to objectify the assessment of the acoustic safety of operators of cyber-physical systems, the developed method takes into account the sound pressure levels in octave bands with average geometric frequencies [19]. Therefore, to objectify the assessment of the acoustic safety of personnel, the developed method ensures that sound pressure levels in octave bands with geometric mean frequencies 31.5, 63, 125, 250, 500, 1000, 2000, 4000, 8000 Hz are taken into account.

By the value k , the acoustic security of cyber-physical system operators can be estimated as unsatisfactory if $k < 5$; satisfactory if $5 \leq k < 15$; good if $15 \leq k < 25$; excellent if $k \geq 25$.

This method allows us to quantify the acoustic safety of cyber-physical system operators, justify the priorities for improving acoustic safety, and evaluate the effectiveness of measures aimed at ensuring it.

6 Conclusion

The use of personalized acoustic situation indicators provides objective monitoring of the acoustic situation at the location of cyber-physical system operators, allowing to assess in real-time and identify changes in health and performance risks caused by the impact of acoustic factors, with the information of cyber-physical system operators and, if necessary, their managers. The introduction of a medical control system using a medical dosimetric register will allow an objective assessment of the entire range of organizational and medical-technical measures to preserve the health of operators of cyber-physical systems, whose professional activities are associated with the impact of potentially dangerous industrial and industrial noise.

The most effective way to reduce the risks to health and reliable operation of cyber-physical systems operators remains the use of noise protection. Provided by using personalized indicators acoustic environment and system of health control of the savings Deposit factual information on the influence of acoustic factors on operators of cyber-physical systems, adequate and reliable quantitative description of the patterns of change in health and performance open new possibilities for inspection, correction and justification of managerial decisions aimed at the preservation of health and prolongation of professional longevity, the operators of cyber-physical systems of many socio-occupational groups.

Acknowledgements This work was supported by a grant from the President of the Russian Federation for state support of leading scientific schools of the Russian Federation (NSH-2553.2020.8).

References

1. Vasiliev, A.V.: (2018) Approaches to assessing environmental risk when exposed to acoustic pollution. *Ecol. Ind. Russia* **3**, 25–27 (2018)
2. Zinkin, V.N., Bogomolov, A.V., Ahmetzjanov, I.M., Sheshegov, P.M.: (2011) Ecological aspects of life safety of the population exposed to aircraft noise. *Teoreticheskaja I Prikladnaja Jekologija* **3**, 97–101 (2011)
3. Zhdanko, I.M., Zinkin, V.N., Soldatov, S.K., Bogomolov, A.V., Sheshegov, P.M.: (2016) Fundamental and applied aspects of preventing the adverse effects of aviation noise. *Human Physiol.* **42**(7), 705–714 (2016)
4. Mehcheryakov, R.V., Galin, R.R.: Review on human–robot interaction during collaboration in a shared workspace. In: *Interactive Collaborative Robotics ICR-2019*. LNCS, vol. 11659, pp. 63–74. Springer, Cham (2019)

5. Dragan, S.P., Bogomolov, A.V., Kotlyar-Shapiro, A.D., Kondrat'eva, E.A.: Experimental study of displays in contralateral acoustic reflex auditory stimulation. *Doklady Biochem. Biophys.* **468**(1), 224–225 (2016)
6. Zinkin, V.N., Soldatov, S.K., Kukushkin, Yu.A., Afanasyev, R.V., Bogomolov, A.V., Akhmetzyanov, I.M., Svidovy, V.I., Pirozhkov, M.V.: (2008) Hygienic assessment of work-ing conditions of workers in the “noise” professions of aircraft repair plants. *Occup. Med. Ind. Ecol.* **4**, 40–42 (2008)
7. Kravets, A.G., Bolshakov, A.A., Shcherbakov, M.V.: *Cyber-Physical Systems: Industry 4.0 Challenges*, 334p. Springer, Berlin (2019). <https://doi.org/10.1007/978-3-030-32648-7>
8. Soldatov, S.K., Bogomolov, A.V., Zinkin, V.N., Averyanov, A.A., Rossels, A.V., Patskin, G.A., Sokolov, B.A.: (2011) Means and methods of protection against aircraft noise: state and development prospects. *Aerospace Environ. Med.* **45**(5), 3–11 (2011)
9. Dragan, S.P., Drozdov, S.V., Zinkin, V.N., Bogomolov, A.V., Soldatov, S.K.: (2013) Efficiency of acoustic noise protection. *Biomed. Eng.* **47**(3), 150–152 (2013)
10. Zinkin, V.N., Bogomolov, A.V., Akhmetzyanov, I.M., Sheshegov, P.M.: (2012) Aviation noise: specific features of biological action and protection. *Aerospace Environ. Med.* **46**(2), 9–16 (2012)
11. Bogomolov, A.V., Dragan, S.P., Zinkin, V.N., Alekhin, M.D.: Acoustic factor environmental safety monitoring information system. In: XXII International Conference on Soft Computing and Measurements (SCM) 2019, St. Petersburg, Russia, pp. 215–218 (2019)
12. Dragan, S.P., Bogomolov, A.V.: (2017) A method for noninvasive diagnostic examination of the tympanic membrane using probing polyharmonic sound signals. *Biomed. Eng.* **50**(6), 390–392 (2017)
13. Dragan, S.P., Bogomolov, A.V., Kotlyar-Shapiro, A.D., Kondrat'eva, E.A.: A method for investigation of the acoustic reflex on the basis of impedance measurements. *Biomed. Eng.* **51**(1), 72–76 (2017)
14. Ushakov, I.B., Bogomolov, A.V., Dragan, S.P., Soldatov, S.K.: (2017) Methodological foundations of personalized hygienic monitoring. *Aerospace Environ. Med.* **51**(6), 53–56 (2017)
15. Bogomolov, A.V., Dragan, S.P.: (2017) Method of acoustic qualimetry of collective noise protection equipment. *Hygiene Sanitation* **96**(8), 755–759 (2017)
16. Ponomarenko, V.A., Soldatov, S.K., Filatov, V.N., Bogomolov, A.V.: (2017) Providing personalized acoustic protection for aviation specialists (practical aspects. *Military Med. J.* **338**(4), 44–50 (2017)
17. Larkin, E.V., Bogomolov, A.V., Privalov, A.N.: (2017) A method for estimating the time intervals between transactions in speech-compression algorithms. *Autom. Document. Math. Linguistics* **51**(5), 214 (2017)
18. Dragan, S.P., Soldatov, S.K., Bogomolov, A.V., Drozdov, S.V., Polyakov, N.M.: (2013) Assessment of the acoustic effectiveness of personal protective equipment against extra-aural exposure to aircraft noise. *Aerospace Environ. Med.* **47**(5), 21–26 (2013)
19. Chueshev, A.V., Melekhova, O.N., Meshcheryakov, R.V.: Cloud robotic platform on basis of fog computing approach. *Lecture Notes in Computer Science*. 2018. 11097 LNAI, pp. 34–43 (2018). https://doi.org/10.1007/978-3-319-99582-3_4
20. Dragan, S.P., Bogomolov, A.V., Zinkin, V.N.: Methodical support of monitoring the acoustic safety of flight personnel. *AIP Conf. Proc.* **2019**, 020019 (2019)
21. Bogomolov, A.V., Dragan, S.P.: (2015) A new approach to the study of impedance characteristics of tympanic membrane. *Dokl. Biochem. Biophys.* **464**(1), 269–271 (2015)
22. Mehcheryakov, R.V., Galin, R.R.: (2020) Human-robot interaction efficiency and human-robot collaboration. *Stud. Syst. Decision Control* **272**, 55–63 (2020)
23. Bogomolov, A.V., Sviridyuk, G.A., Keller, A.V., Zinkin, V.N., Alekhin, M.D.: Information-logical modeling of information collection and processing at the evaluation of the functional reliability of the aviation ergate control system operator. In: *Proceedings of 2018 International Conference on Human Factors in Complex Technical Systems and Environments (ERGO)*. Piscataway, NJ, United States: Piscataway, NJ, United States, pp. 106–110 (2018)

24. Kravets, A.G., Bolshakov, A.A., Shcherbakov, M.V.: Cyber-Physical Systems: Advances in Design & Modelling, 340p. Springer, Berlin (2019). <https://doi.org/10.1007/978-3-030-32579-4>
25. Ushakov, V.: (2014) Electric Power Industry: Current State, Problems, Prospects, p. 234. Springer, Heidelberg (2014)
26. Larkin E., Bogomolov A., Gorbachev D., Privalov A.: About approach of the transactions flow to poisson one in robot control systems. Lecture Notes in Computer Science, vol. 10459, pp. 113–122. Springer, Berlin (2017). https://doi.org/10.1007/978-3-319-66471-2_13
27. Meshcheryakov, R.: Control of hyperlinked cyber-physical systems. Smart Innov. Syst. Technol. **2020**, 27–33 (2020)