Studies in Systems, Decision and Control 437

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

# Society 5.0

## Cyber-Solutions for Human-Centric Technologies

Springer

# Studies in Systems, Decision and Control

Volume 437

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.

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

# Society 5.0

## Cyber-Solutions for Human-Centric Technologies

Springer

*Editors*
Alla G. Kravets
CAD&RD
Volgograd State Technical University
Volgograd, Russia

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

Maxim V. Shcherbakov
CAD&RD
Volgograd State Technical University
Volgograd, Russia

ISSN 2198-4182 ISSN 2198-4190 (electronic)
Studies in Systems, Decision and Control
ISBN 978-3-031-35874-6 ISBN 978-3-031-35875-3 (eBook)
https://doi.org/10.1007/978-3-031-35875-3



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

# Preface

Society 5.0 is a paradigm of a society which combines a human-centred approach and cyber-physical systems with robotics and artificial intelligence components. This book focuses on open issues of Society 5.0 in the framework of new challenges in the world towards improving the quality of human life in society.

There are five parts included in the book.

The first part *Society 5.0: Biomedicine and Healthcare* presents how cyber-physical systems help in healthcare, e.g. to prevent negative outcomes of pregnant women with hypertension, breast cancer diagnostics, diet monitoring, product package and so on.

In the second part, *Society 5.0: Human-Centric Cyber-Solutions* chapters present new AI-based cyber-physical solutions for various domains, such as planning, controlling production, using VR for training and others.

In the third part, *Society 5.0: Socio-economic Systems Modelling*, the chapters focus on open questions related to the modeling of socio-economic systems.

*Society 5.0: Industrial Cyber-Solutions* includes chapters concerning intelligent support for the oil market, mining companies and railway companies, maintenance and repair of a machine-building cyber-physical system and float glass production management.

The last part *Society 5.0: Cyber-Solutions Security* provides new results on cybersecurity for the corporate mobile networks, critical infrastructures in telecommunication, video conferencing system and industrial control systems of gas production enterprises.

This book is directed to researchers, practitioners, engineers, software developers, professors and students in engineering, computer science and sociology. We do hope the book will be useful for them to create new and efficient solutions for people.

Volgograd, Russia — Alla G. Kravets
St. Petersburg, Russia — Alexander A. Bolshakov
Volgograd, Russia — Maxim V. Shcherbakov
April 2023

# Contents

**Society 5.0: Biomedicine and Healthcare**

**Stochastic Analysis of Hankel and Toeplitz Matrices for Medical and Technical Applications** . . . . . 3
Vladimir Kulikov, Alexander Kulikov, and Valery Khranilov

**Analysis of Clinical Data in Pregnant Women with Hypertension Based on a New Stochastic Approach** . . . . . 17
Vladimir Kulikov, Svetlana Kolobova, Liudmila Borovkova, and Alexander Ignatyev

**Hierarchical Fuzzy Inference of Adequacy of Highly Informative Diagnostic Signs of Breast Cancer** . . . . . 31
Ilya Germashev, Victoria Dubovskaya, and Alexander Losev

**Development of a Personalized Diet Using Structural Optimization** . . . . 43
Marina A. Nikitina

**Development of an Algorithm for Preparing Semi-finished Products for Packaging** . . . . . 53
Mikhail V. Tarachkov, Oleg V. Tolstel, and Alexandr L. Kalabin

**Society 5.0: Human-Centric Cyber-Solutions**

**Computer Training System for Planning Multi-Assortment Discrete-Continuous Productions** . . . . . 65
Tamara Chistyakova, Christian Kohlert, Olga Shashikhina, Ivan Kornienko, and Alexander Plekhanov

**Intelligent Training System for Controlling Polymer Extrusion Processes in Emergency Situations** . . . . . 75
Andrey Polosin, Tamara Chistyakova, Christian Kohlert, and Frank Kleinert

**Method for Constructing Virtual Reality Simulators for Turning and Milling for an Engineering Education System for Building Cyber-Physical Systems** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 91
Roman Aslanov and Alexander Bolshakov

**Development of Software for the Organization of Training of TPP Workers on the Basis of Digital Twins of Equipment** . . . . . . . . . . . . . . . . . . 107
Vladimir Agibalov, Michael Belov, Aleksey Dolgushev, and Ivan Shcherbatov

**Heterogeneous Information System for the Integration of Departmental Databases on the State and Development of Human Capital** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 117
Alexandr Ivanov and Alexey Bogomolov

**Society 5.0: Socio-economic Systems Modelling**

**Application of Quantum-Like Mathematical Models Based on Status Functions for the Analysis of Socio-Economic Systems. Part 1. Introduction** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 131
Irina Veshneva and Alexander Bolshakov

**Application of Quantum-Like Mathematical Models Based on Status Functions for the Analysis of Socio-Economic Systems. Part 2: Modelling Based on Status Functions** . . . . . . . . . . . . . . . . . . . . . . . . . 145
Irina Veshneva and Alexander Bolshakov

**Development of a Decision Support System for Assessing and Monitoring Regional Competitiveness Risks** . . . . . . . . . . . . . . . . . . . . . 161
Galina Chernyshova, Ekaterina Piunova, Irina Veshneva, and Gleb Rokakh

**Mental Analysis of Stackelberg Leadership in Nonlinear Oligopoly Model for Cyber-Physical Systems of Oil Market** . . . . . . . . . . . . . . . . . . . . . 175
Mikhail Geraskin

**Society 5.0: Industrial Cyber-Solutions**

**Planning the Optimal Railway Transport Load of a Mining Company** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 189
Evgeny Eletin, Galina Borovkova, and Alexander Galkin

**Intelligent System for Railway Joint Diagnostics** . . . . . . . . . . . . . . . . . . . . . . 199
N. Efimushkin, N. Efimushkina, and S. Orlov

**Information Support for Maintenance and Repair of a Cyber-Physical System of a Machine-Building Profile** . . . 213
Yulia V. Nemtinova, Andrey B. Borisenko, Vladimir A. Nemtinov, Alexey S. Protasov, and Kirill V. Nemtinov

**A System Dynamics Model for Float Glass Production Management** . . . 227
Dmitry Petrov

**Society 5.0: Cyber-Solutions Security**

**Interval Analysis of Security for Information and Telecommunication Resources of Critical Infrastructures** . . . 241
Igor Kotenko and Igor Parashchuk

**Assigning Access Rights to Applications in the Corporate Mobile Network: Software Development** . . . 251
Alla G. Kravets, N. A. Salnikova, and E. L. Shestopalova

**Scaling Networks and Capturing Keys Using Combined Systems of Sets** . . . 267
Alexander Frolov and Natalya Kochetova

**Video Conferencing Subsystem of the Secure Serverless Internet (TheOoL.Net)** . . . 279
A. V. Nenashev and A. Yu. Tolstenko

**Formation of an Optimal Set of Protective Measures to Handle the Information Security Risks of Industrial Control Systems of Gas Production Enterprises** . . . 299
Alexander Bolshakov, Darya Fomina, and Andrey Rimsha

**Application of PID Control Principles in the Tasks of Modeling the Movement of Wheeled Vehicles Equipped with an Anti-Lock Braking System** . . . 321
Grigory Boyko, Alexey Fedin, and Jozef Redl

# Society 5.0: Biomedicine and Healthcare

# Stochastic Analysis of Hankel and Toeplitz Matrices for Medical and Technical Applications

**Vladimir Kulikov, Alexander Kulikov, and Valery Khranilov**

**Abstract** The chapter states the results of modern methods for stochastic analysis of singular spectra of special matrices within the singular representation of matrix operators (SVD-algorithm). Using the method for the approximate solution of the first kind Fredholm equation the reconstruction of distributions of the singular spectra for the following matrices was carried out: (1) a Hankel limiting ill-conditioned matrix; (2) a Toeplitz matrix; (3) a special two-diagonal matrix. Toeplitz and Hankel matrices are quite widely presented in applied and theoretical branches of mathematics (functional analysis, theory of random processes), technical applications, biomedicine (mathematical simulation of the functional states of human systems and organs by differential equations, the study of time series in diagnosis using the method of the singular spectral analysis). A new approach aimed at finding the laws for distribution of singular values in the perturbation method is required to control and predict the stability of solutions for systems of algebraic equations to which differential equations are reduced as well as to study the possibilities of stochastic vector regularization. It is shown that for the correct solution of inverse problems in medical and technical applications, including cyber-physical systems (CPS), the EC-regularization method can be used. The results of the studies can be useful for creating mathematical models of human body structures, organs and systems which display their stochastic and chaotic properties, including for modern medical expert systems based on the diagnosis of random variables and processes.

**Keywords** Hankel and Toeplitz matrices · Stochastic analysis · Identification · Singular values · Matrix pseudo-inversion · Diagnostics · Expert systems

V. Kulikov (✉) · A. Kulikov · V. Khranilov
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

V. Khranilov
e-mail: hranilov@nntu.ru

 
A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_1

## 1 Introduction

The chapter states the results of modern methods for stochastic analysis of singular spectra of special matrices as part of the method for singular representation of matrix operators (SVD algorithm). The method for the approximate solution of the first kind Fredholm equation is used to identify the distributions of singular spectra of matrices: a Hankel limiting ill-conditioned matrix; a Toeplitz matrix; a special two-diagonal matrix.

Toeplitz and Hankel matrices are quite widely represented in applied and theoretical branches of mathematics (functional analysis, theory of random processes), technical applications leading to boundary value problems for partial differential equations (aero- and hydrodynamics, thermal physics) [1–6].

Mathematical models are also used in the form of boundary value problems, systems of differential equations, stochastic optimization and inverse problems are solved in works on biomedicine of S.I. Kabanikhin, the Corresponding Member of the Russian Academy of Sciences (ICM and MG SB RAS), G.A. Bocharov, Professor (ICM RAS n.a. G.I. Marchuk) [7–11].

Hankel and Toeplitz matrices can be used in biomedicine for mathematical simulation of functional states of human systems and organs, for solving inverse problems with differential equations in the study of the infectious disease dissemination, when analyzing time series in diagnosis within expert systems, including CPS, using the method of singular spectral analysis [12–15]. The new authors' approach aimed at finding the laws of distribution of singular values in the perturbation method is required to control and predict the stability of solutions for systems of algebraic equations approximating differential equations in numerical methods as well as to study possibilities of stochastic vector regularization [16, 17].

The identified distribution densities of small singular values serve to control and predict the stability of normal pseudo-solutions of algebraic systems of equations under conditions of perturbation of matrices and right-hand sides of systems. In the chapter the procedure is implemented at two essentially different levels of perturbation of the diagonal elements in the matrices (the normal law of element distortion). The analysis of the inverse problem solutions (ill-conditioned SLAE with a high-order square matrix) is carried out for resistance to errors of varying intensity.

It is expected that the results of the performed studies can be used to create mathematical models of the human body structures, organs and systems, which exhibit stochastic and chaotic properties, also for modern expert medical systems based on the diagnosis of random processes [18].

## 2 Method of Solution and Outcomes of Computational Experiments

In [16] there was for the first time ever studied the effect of the emergence of poly-modal (including non-Gaussian) laws of distribution of singular values for an ill-conditioned matrix in the SVD algorithm operation under conditions of varying the artificial or experimental data of the problem.

This chapter states the results of the new approach development on the example of matrices of the said three types. The eight-order Hilbert matrix is herein considered as an ill-conditioned Hankel matrix. The low order of this test matrix enables us to compactly present the identification results without loss of generality. The identification of the distribution densities of singular values for this matrix under perturbations of the main diagonal elements is performed at two different levels of Gaussian noise (the computation environment is MATLAB package). In fact, the well-known Monte Carlo algorithm is used for stochastic tests [19]. Figure 1 shows the Hankel (Hilbert) ($8 \times 8$) matrix for illustration purposes. This is a symmetric matrix with diminishing diagonal elements.

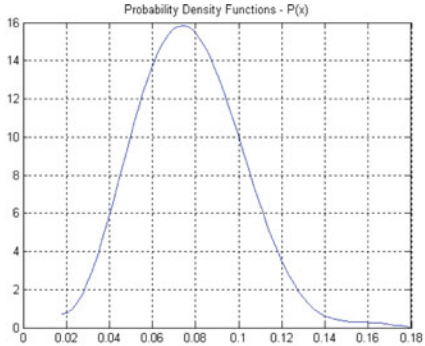
1000 perturbation cycles are chosen at each level. Eight distributions are identified as per the number of singular values. Figure 2 shows the distributions of the singular values for the matrix H under Gaussian perturbation of the diagonal by the formula (1) *hii + 0.1*randn.*

The conditioning number of the initial Hilbert matrix is $cond(H) = 1.53e + 10$; the rank is 8. The singular spectrum after 1000 perturbations is determined by the vector: {1.6556; 0.26975; 0.1226; 0.093268; 0.082738; 0.07313; 0.042492; 0.018888}. The conditioning number of the perturbed matrix is $cond(HV) = 87{,}6$. (Fig. 3).

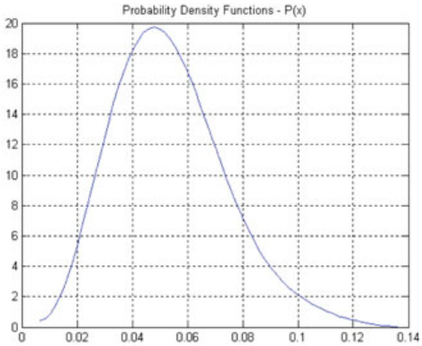
The analysis of the identified densities shows that the transformation of the normal distribution law typical for the first singular values becomes apparent starting from the sixth singular value. The sixth and seventh values have Pearson-type distributions and the smallest (eighth) one is characterized by a slightly expressed polymodal density. In general, the density resembles a distorted gamma distribution.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.5 | 0.33333 | 0.25 | 0.2 | 0.16667 | 0.14286 | 0.125 |
| 2 | 0.5 | 0.33333 | 0.25 | 0.2 | 0.16667 | 0.14286 | 0.125 | 0.11111 |
| 3 | 0.33333 | 0.25 | 0.2 | 0.16667 | 0.14286 | 0.125 | 0.11111 | 0.1 |
| 4 | 0.25 | 0.2 | 0.16667 | 0.14286 | 0.125 | 0.11111 | 0.1 | 0.090909 |
| 5 | 0.2 | 0.16667 | 0.14286 | 0.125 | 0.11111 | 0.1 | 0.090909 | 0.083333 |
| 6 | 0.16667 | 0.14286 | 0.125 | 0.11111 | 0.1 | 0.090909 | 0.083333 | 0.076923 |
| 7 | 0.14286 | 0.125 | 0.11111 | 0.1 | 0.090909 | 0.083333 | 0.076923 | 0.071429 |
| 8 | 0.125 | 0.11111 | 0.1 | 0.090909 | 0.083333 | 0.076923 | 0.071429 | 0.066667 |

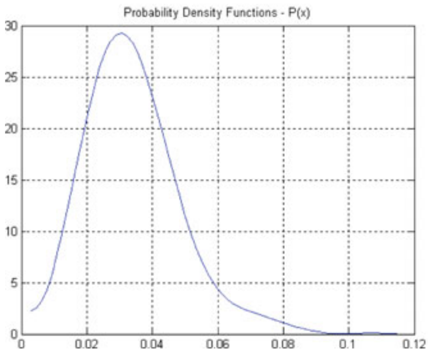
**Fig. 1** Matrix **H** is the initial Hankel matrix



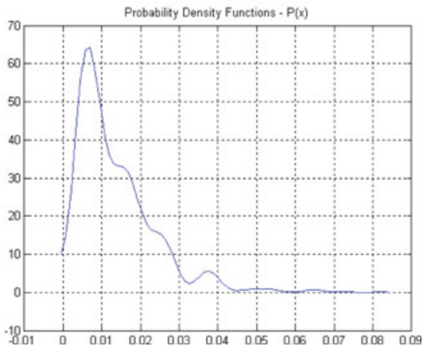

S5 NOPT = 4 $t_e = 0.97$ s
k = 2.032




S6 NOPT =5 $t_e = 1.00$ s
k = 2.002




S7 NOPT = 6 $t_e = 1.03$ s
k = 1.961




S8 NOPT =18 $t_e = 1.11$ s
k = 1.621

**Fig. 2** Identified distribution densities of singular values numbered 5–8 in the cycle of 1000 perturbations by normal noise. **NOPT** is the optimum of the expansion terms for P(x), **k** is the entropy coefficient, $\mathbf{t_e}$ is the time to perform identification P(x). High level of diagonal elements perturbation in the Hankel matrix

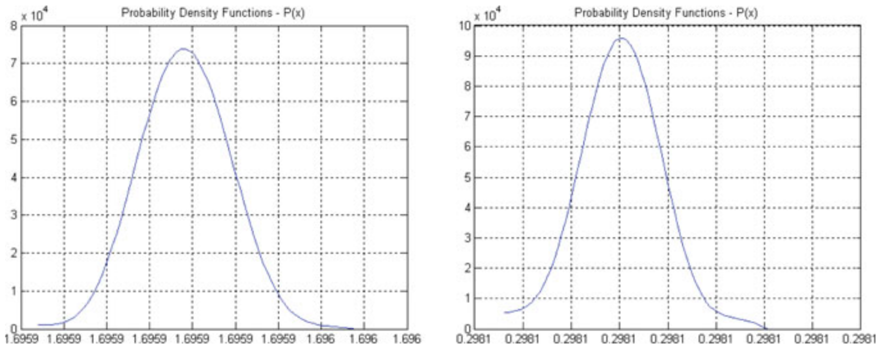
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.88735 | 0.5 | 0.33333 | 0.25 | 0.2 | 0.16667 | 0.14286 | 0.125 |
| 2 | 0.5 | 0.4239 | 0.25 | 0.2 | 0.16667 | 0.14286 | 0.125 | 0.11111 |
| 3 | 0.33333 | 0.25 | 0.16766 | 0.16667 | 0.14286 | 0.125 | 0.11111 | 0.1 |
| 4 | 0.25 | 0.2 | 0.16667 | 0.22597 | 0.125 | 0.11111 | 0.1 | 0.090909 |
| 5 | 0.2 | 0.16667 | 0.14286 | 0.125 | 0.0099511 | 0.1 | 0.090909 | 0.083333 |
| 6 | 0.16667 | 0.14286 | 0.125 | 0.11111 | 0.1 | 0.018934 | 0.083333 | 0.076923 |
| 7 | 0.14286 | 0.125 | 0.11111 | 0.1 | 0.090909 | 0.083333 | 0.0024089 | 0.071429 |
| 8 | 0.125 | 0.11111 | 0.1 | 0.090909 | 0.083333 | 0.076923 | 0.071429 | 0.2045 |

**Fig. 3** Matrix $\mathbf{H_V}$ is the perturbed Hankel matrix (after the 1000th iteration)

The preliminary conclusion is as follows: firstly, the effect of the emergence of polymodal, non-Gaussian distribution laws of small singular values is confirmed for a Hankel ill-conditioned matrix; secondly, the diagonal perturbation with the specified intensity can be used for stochastic regularization when solving the SLAE.

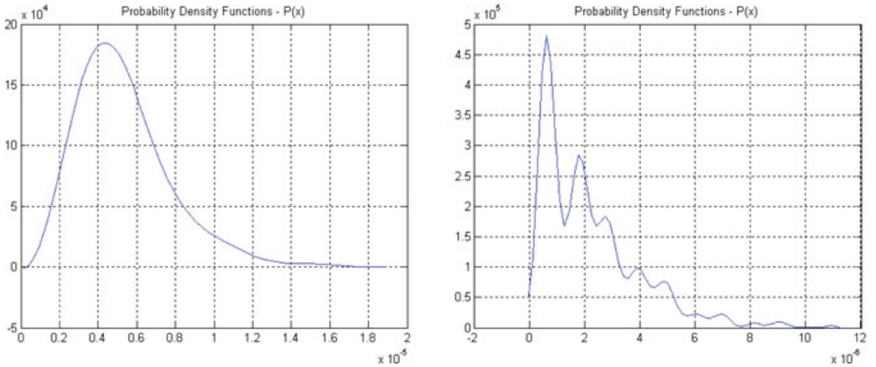
With identification at a low level of perturbation: (2) $h_{ii} + 0.00001*randn$ a more expressed polymodal effect is found in the distribution of the minimum singular value (Fig. 4).

The singular spectrum after the 1000th iteration (perturbation) is determined by the vector: {1.6959; 0.29812; 0.026217; 0.0014732; 6.0361e-005; 1.2659e-005; 7.1063e-006; 2.3334e-006}. The conditioning number of the perturbed matrix is $cond(HV) = 7.3e + 05.$




S1 NOPT = 5 elapsed_time = 0.89 s
k = 0.993

S2 NOPT = 5 elapsed_time = 0.89 s
k = 1.226




S7 NOPT = 7 elapsed_time = 0.98 s
k = 1.893

S8 NOPT = 22 elapsed_time = 1.41 s
k = 1.545

**Fig. 4** Identified distribution densities of singular values numbered 1, 2 and 7, 8 in the cycle of 1000 perturbations by normal noise. Low level of diagonal elements perturbation in the Hankel matrix

Thus, the low-intensive perturbations of the Hankel ill-conditioned matrix cannot ensure the regularization of SLAEs within the range of Lagrange parameters typical of A.N. Tikhonov classical regularization problems; that is determined by the singular spectrum structure of the Hilbert matrix [20].

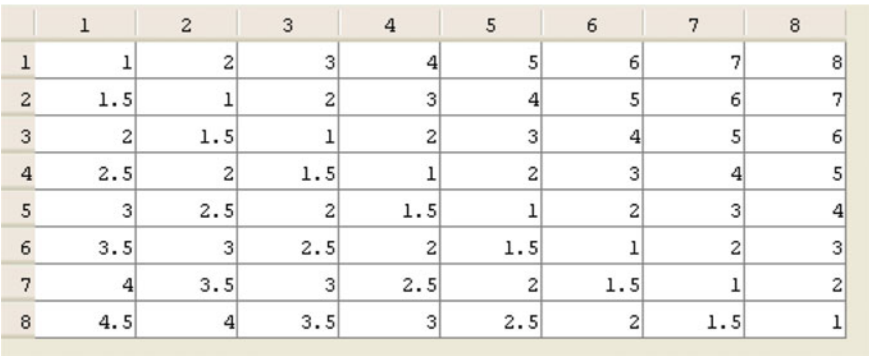
## 3 Identification of the Distribution Densities of Singular Values for a Toeplitz Matrix

A square eighth-order matrix is considered as a Toeplitz matrix. The Toeplitz matrix ($8 \times 8$) is shown in Fig. 5. This is an asymmetric matrix with constant diagonal elements (on the main and secondary diagonals).

As before, the identification of the distribution densities of singular values for this matrix is carried out within the method of the main diagonal element perturbation at two different levels of Gaussian noise. The number of perturbations at each level is also 1000. Command codes for setting the matrix (MATLAB) are as follows: c = (1:0.5:4.5); r = (1:8); T = toeplitz(c, r). The conditioning number of the initial Toeplitz matrix is cond (T) = 65.4; the rank is 8. The singular spectrum is the vector: {25.508; 8.5779; 3.0117; 1.2053; 0.76973; 0.54161; 0.44128; 0.38979}, which shows that the matrix is fairly well conditioned.

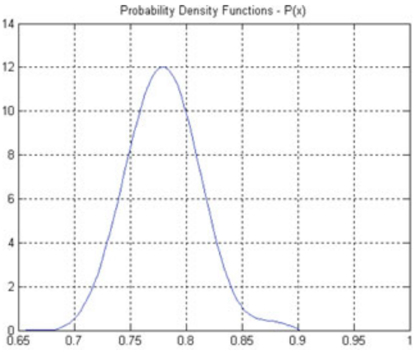
Initially, the identification of the singular value distribution for matrix T was performed under Gaussian perturbation of the diagonal by the formula tii + 0.1*randn. The analysis of the identified densities suggests that the singular spectrum of the Toeplitz matrix has practically no transformation of the normal distribution law (Fig. 6).

If the perturbation level is increased by an order of magnitude: $t_{ii}$ + randn (randn is the Gaussian random number generator program), the polymodal effect is again found in the distribution of the minimum singular value (Fig. 7).

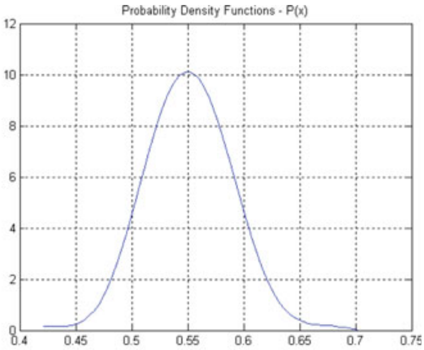


| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 2 | 1.5 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 3 | 2 | 1.5 | 1 | 2 | 3 | 4 | 5 | 6 |
| 4 | 2.5 | 2 | 1.5 | 1 | 2 | 3 | 4 | 5 |
| 5 | 3 | 2.5 | 2 | 1.5 | 1 | 2 | 3 | 4 |
| 6 | 3.5 | 3 | 2.5 | 2 | 1.5 | 1 | 2 | 3 |
| 7 | 4 | 3.5 | 3 | 2.5 | 2 | 1.5 | 1 | 2 |
| 8 | 4.5 | 4 | 3.5 | 3 | 2.5 | 2 | 1.5 | 1 |

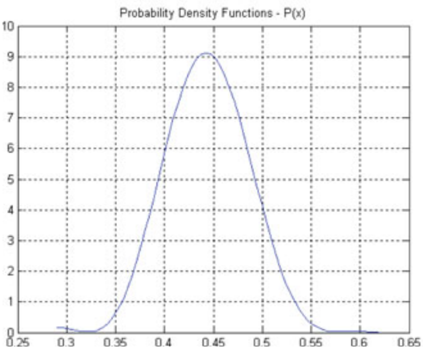
**Fig. 5** Initial Toeplitz matrix **T**



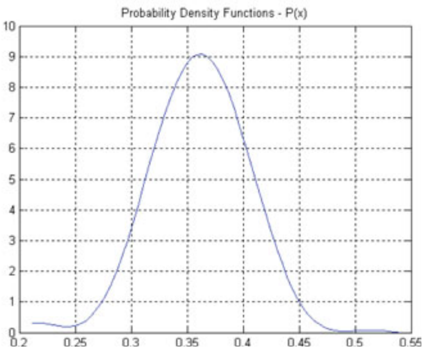

S5 NOPT = 6 elapsed_time = 0.89 s
k = 2.051




S6 NOPT = 5 elapsed_time = 0.92 s
k = 2.057




S7 NOPT = 5 elapsed_time = 0.92 s
k = 2.052



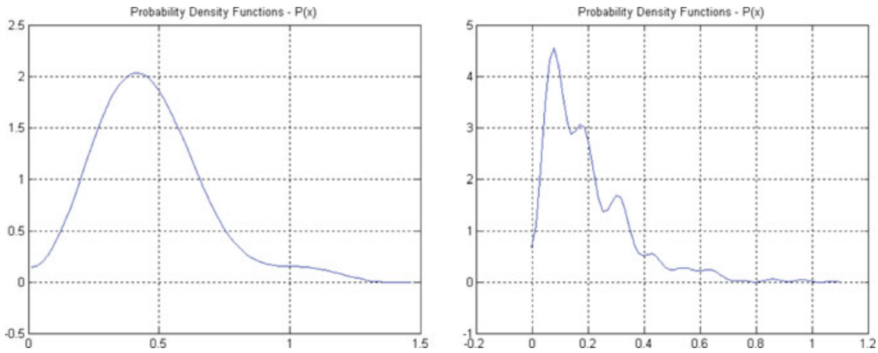

S8 NOPT = 5 elapsed_time = 0.87 s
k = 2.044

**Fig. 6** Identified distribution densities of singular values numbered 5–8 in a cycle of 1000 perturbations by normal noise. Relatively high level of perturbation of the diagonal elements in the Toeplitz matrix: $t_{ii} + 0.1$*randn

Thus, the tendency in the distribution of small singular values for the both matrices in the perturbation method is the same and has a polymodal character. The number of manifested modes depends on a noise level.

## 4 Analysis of Ill-Conditioned SLAE Solutions for Resistance to Errors of Varying Intensity

Let us turn first to an ill-conditioned general matrix. We consider as an example the matrix G(30 × 30) being the sum of two matrices: the two-diagonal extremely ill-conditioned matrix D2 and the stochastic matrix B with elements distributed under the normal law.




S7 NOPT = 5 elapsed_time = 1.03 s
k = 1.939

S8 NOPT =20 elapsed_time = 1.08 s
k = 1.585

**Fig. 7** Identified distribution densities of singular values numbered 7, 8 in a cycle of 1000 perturbations by normal noise. High level of the diagonal element perturbation in the Toeplitz matrix: $t_{ii}$ + randn

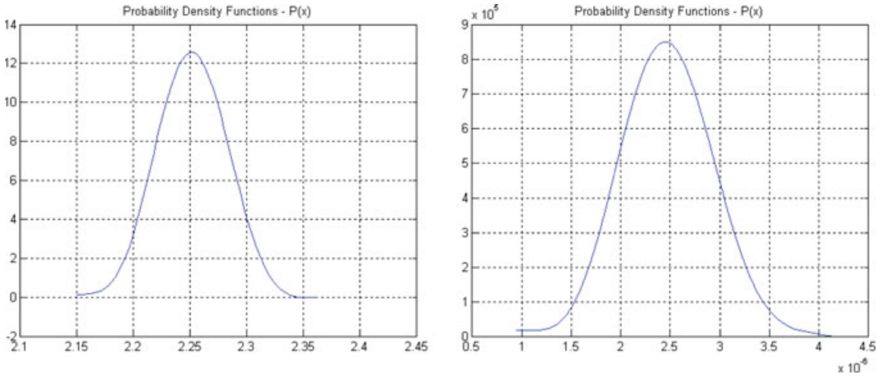
The matrix D2 (Fig. 8) is taken from the test example of the monograph [21]. The conditioning number of this matrix is 5.66e + 12.

The second matrix is assigned by the operator $B = 0.00001*randn(N)$; N = 30. The conditioning number of the matrix G (completely filled with elements) is 2.03e + 6; rank = 30. The singular spectrum contains thirty components: {5.0449; 5.0177; 4.9799; 4.9315; …, 2.3176; 2.2796; 2.4839e-006}. The matrix G has a smoothly decreasing singular spectrum (all components except for the last one are greater than unity), but it is still relatively ill-conditioned because of the smallness of the 30th singular value.

Figure 9 shows the distributions of small singular values of the matrix G under Gaussian perturbation of the main diagonal by the formula: $g_{ii} + 0.1*randn$.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.4 | 3.6667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1.4 | 3.6667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 1.4 | 3.6667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 1.4 | 3.6667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 1.4 | 3.6667 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 1.4 | 3.6667 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 1.4 | 3.6667 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.4 | 3.6667 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.4 | 3.6667 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.4 | 3.6667 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.4 | 3.6667 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.4 |

**Fig. 8** Fragment of the two-diagonal ill-conditioned matrix $\mathbf{D_2}$




S29 NOPT = 5 elapsed_time = 0.91 s
k = 2.058

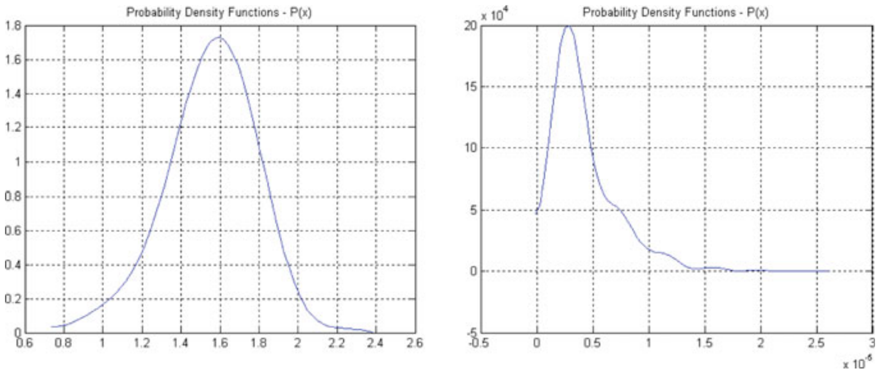
S30 NOPT = 5 elapsed_time = 0.91 s
k = 2.061

**Fig. 9** Identified distribution densities of singular values numbered 29, 30 in a cycle of 1000 perturbations by normal noise. Perturbation level of the matrix **G** is: $g_{ii} + 0.1$*randn

As in the case of the Toeplitz matrix, the total singular spectrum of the matrix G is characterized by the lack of transformation of the normal distribution law. Even the least singular value has a normal distribution density.

If the perturbation level is increased by an order of magnitude, the effect of the polymodality emergence will be revealed (Fig. 10), which appears at sufficiently high perturbation intensities.

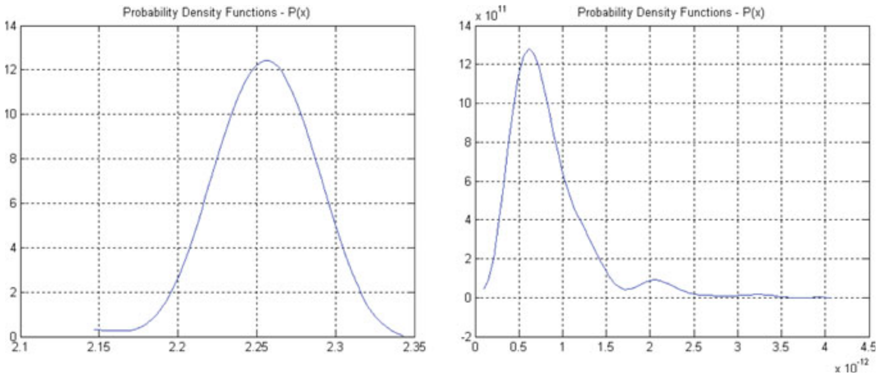
The both levels of the diagonal perturbation for the matrix G give no grounds for the stochastic diagonal regularization when solving the SLAE, because the conditioning number of the perturbed matrix does not decrease (on the 1000th iteration,




S29 NOPT = 6 elapsed_time = 0.89 s
k = 2.039

S30 NOPT = 13 elapsed_time = 1.25 s
k = 1.691

**Fig. 10** Identified distribution densities of singular values numbered 29 and 30 in a cycle of 1000 perturbations by normal noise. Perturbation level of the matrix **G** is: $g_{ii}$ + randn




S29 NOPT = 4 elapsed_time = 0.84 s
k = 2.051

S30 NOPT = 12 elapsed_time = 1.12 s
k = 1.594

**Fig. 11** Identified distribution densities of singular values numbered 29 and 30 in a cycle of 1000 perturbations by normal noise. Perturbation level of the matrix $\mathbf{D_2}$ is: dii + 0.1*randn

cond(G) = 8.6e + 05). We shall thereat note the following feature: the matrix D2 obtains in the perturbation cycle the weak bimodality of the 30th singular value (Fig. 11).

The introduction of a random matrix B (the summing of B and the extremely ill-conditioned matrix D2) helps to reduce the conditioning number (in the classical case of regularization by making consistent the "perturbation dispersion" with the operator's errors and the right-hand side of the SLAE). Hence, the stability (or instability) of the solutions and the possibility of stochastic regularization shall be determined not only by diagonal elements but also by the periphery of the ill-conditioned matrix. Note that the perturbation of the first non-zero auxiliary diagonal D2 does not give rise to the effect of polymodality of small singular numbers.

We analyze further the solutions of the inverse problem with a special matrix of a relatively high order for resistance to errors of varying intensity.

The initial two-diagonal matrix DM = D2 is given above and has the order of (30 × 30). The components of the right-hand side vector are determined by the formulas of the test case [21]:

$$f_k = \frac{152k + 118}{15(2k+1)(2k+3)} \quad (1 \leq k \leq M-1),\ f_M = \frac{7}{5(2M+1)} \quad M = 30.$$

The exact system solution for any M is a vector with components

$$x_k = \frac{1}{2k+1} \quad (1 \leq k \leq M).$$

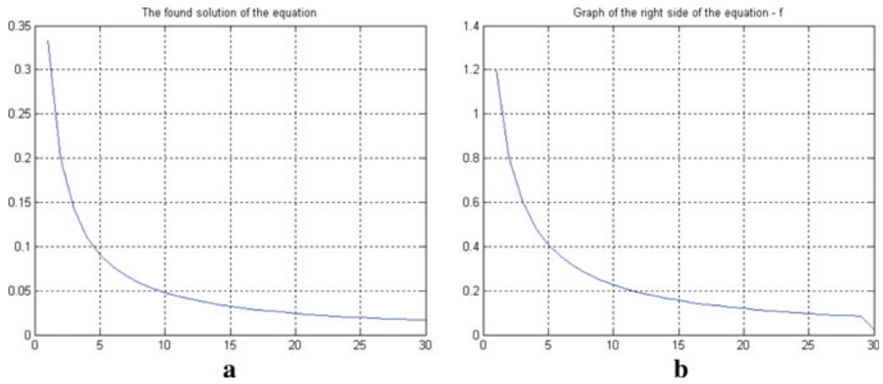



**Fig. 12** A SLAE solution. The exact solution coincides with graph **a**; **b** the right-hand side of the system of equations

The numerical solution of the system of linear equations using the Moore–Penrose pseudo-inversion method (SVD-algorithm of the MATLAB package) and the right-hand side of the system are shown in Fig. 12.

The analysis of SLAE solutions shows that with the minimal perturbation of the vector fk components the strongest distortion of the vector x occurs in subsequent solutions (Fig. 13a). Spikes thereat reach the level of 1.e+06 ÷ 1.e+08 units. The singular spectrum of the matrix DM is shown in Fig. 13b. The spectrum is smoothly descending; all singular values except for one (the minimum value, equal to 8.94e-013) are greater than unity. The conditioning number of the system is 5.66e + 012.

Thus, there is an ill-posedness of the pseudo-inversion problem in systems with ill-conditioned matrices. To find error-resistant initial data of solutions, known algorithms [22–25] can be used. In this chapter we demonstrate the efficiency of the EC-regularization method [26] studied and modified in [17]. Figure 14a shows the

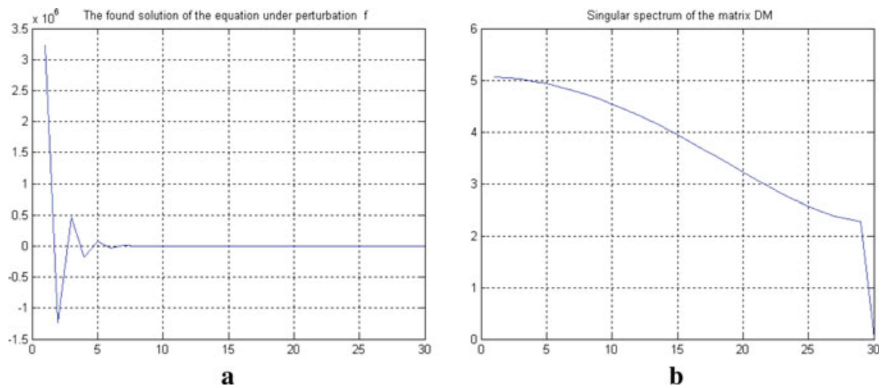



**Fig. 13** **a** SLAE solution at perturbation fp(j) = fk(j) + 0.000003*randn; **b** singular spectrum of the matrix $D_M$

comparison of SLAE solutions with DM by the both methods without perturbing the right-hand side of the system. Solutions with EC-regularization are stable (there is only one spike at the starting point) and practically do not change from the level of 3.e-10*randn of the perturbation of fk to the level equal to 0.03*randn (the graph is blue line). Some error at the initial point of the solution in the EC-regularization method is a sort of “payment” for the non-use of a priori information about the solution and may be otherwise reduced.

In the EC-regularization method with the further rising perturbation to values of 0.3*randn (Fig. 15a) some consistent changes of *fk* and solutions x occur (Fig. 15b).

Thus, the suggested method responds correctly to changes in *fk*. At the said noise levels, the Moore–Penrose pseudo-inversion method is absolutely unstable and requires the regularization procedure [27].

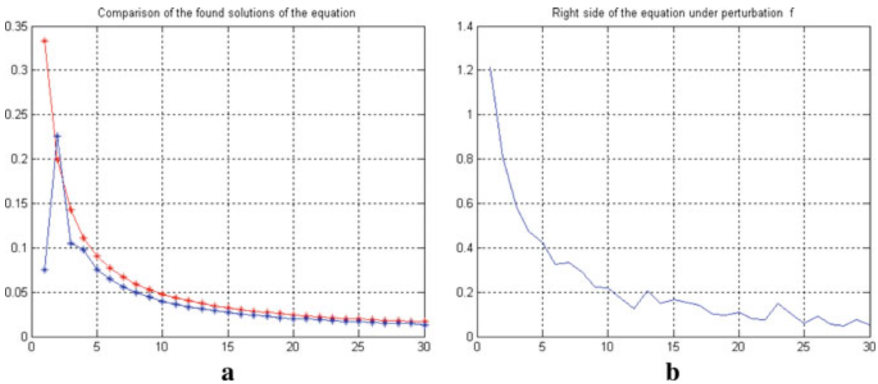



**Fig. 14** **a** Graphs of SLAE solutions without perturbation (red line is SVD algorithm; blue line is EC-regularization); **b** the right-hand side of the system for fp(j) = fk(j) + 0.03*randn

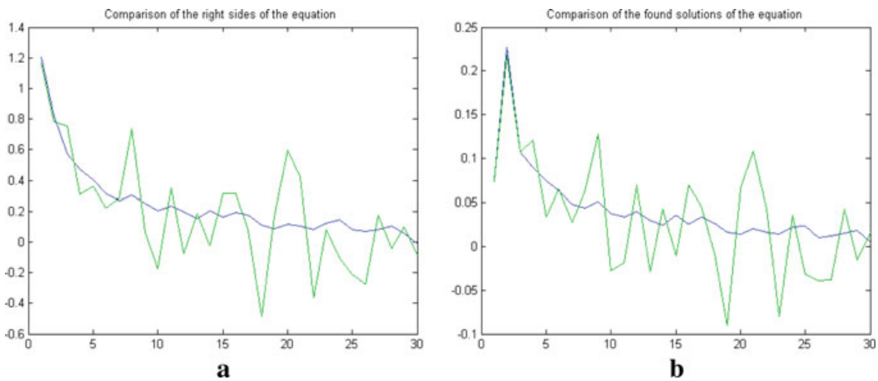



**Fig. 15** The right-hand sides of SLAE (30 × 30) under perturbation (blue line is level 0.03*randn; green line is 0.3*randn)—**a**; corresponding solutions of the system—**b**. Solution algorithm: EC-regularization

## 5 Results and Conclusions

The method for approximate solution of the first kind Fredholm equation is applied to identify distributions of singular spectra of the following matrices: a Hankel limiting ill-conditioned matrix, a Toeplitz matrix and a special two-diagonal matrix.

The developed approaches are required to control and predict the stability of solutions for systems of algebraic and differential equations as well as to study the possibility of stochastic vector regularization.

It is shown that the method of matrix pseudo-inversion based on the singular value decomposition without regularization results in large distortions of the normal pseudo-solution not only in case of the rapid decay of singular values to zero but also in case of a smoothly decreasing spectrum with at least one small singular value available. The EC-regularization method can be used for the correct and stable solution of the said inverse problems.

## References


1. Hartman, P.: On completely continuous Hankel matrices. Proc. Amer. Math. Soc. **9**, 862–866 (1958)
2. Rosenblum, M.: The absolute continuity of Toeplitz's matrices. Pacific J. Math. **10**, 987–996 (1960)
3. Heinig, G., Rost, K.: Algebraic Methods for Toeplitz-like Matrices and Operators. Birkhauser, Basel (1984)
4. Peller, V.V.: Hankel Operators and their Applications. Springer (2003)
5. Gerard, P., Pushnitski, A.: Inverse spectral theory for a class of non-compact Hankel operators. Mathematika **65**(1), 132–156 (2019)
6. Gerard, P., Pushnitski, A.: Weighted model spaces and Schmidt subspaces of Hankel operators. J. Lond. Math. Soc. **101**(1), 271–298 (2020)
7. Krivorotko, O.I., Kabanikhin, S.I., Sosnovskaya, M.I., Andornaya, D.V.: Sensitivity and identifiability analysis of COVID-19 pandemic models. Vavilov J. Genet. Breed. **25**(1), 82–91 (2021)
8. Krivorotko, O.I., Kabanikhin, S.I., et al.: Mathematical modeling and forecasting of COVID-19 in Moscow and Novosibirsk region. Numer. Anal. Appl. **13**(4), 332–348 (2020)
9. Krivorotko, O.I., Andornaya, D.V., Kabanikhin, S.I.: Sensitivity analysis and practical identifiability of some mathematical models in biology. J. Appl. Ind. Math. **14**(1), 115–130 (2020)
10. Luzyanina, T., Mrusek, S., Edwards, J.T., Roose, D., Ehl, S., Bocharov, G.: Computational analysis of CFSE proliferation assay. J. Math. Biol. **54**, 57–89 (2007)
11. Kabanikhin, S.I.: Definitions and examples of inverse and ill-posed problems. J. Inverse Ill-Posed Probl. **16**, 317–357 (2008)
12. Raue, A., Karlsson, J., Saccomani, M.P., Jirstrand, M., Timmer, J.: Comparison of approaches for parameter identifiability analysis of biological systems. Bioinformatics **30**(10), 1440–1448 (2014)
13. Rodriguez-Fernandez, M., Egea, J.A., Banga, J.R.: Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. BMC Bioinf **7**, 483–500 (2006)
14. Banks, H.T., Hu, S., Thompson, W.C.: Modeling and Inverse Problems in the Presence of Uncertainty. Chapman and Hall/CRC, New York (2014)

15. Golyandina, N., Nekrutkin, V., Zhigljavsky, A.: Analysis of Time Series Structure: SSA and Related Techniques. Chapman and Hall/CRC, London (2001)
16. Kulikov, V.B.: Stochastic properties of SVD decomposition. Identification of the distribution densities of singular values of matrix operators in solving SLAE. Cloud Sci. **7**(1), 49–60 (2020) (in Russian)
17. 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, pp. 179–191. Springer, Cham (2020)
18. Kulikov, V., Kulikov, A., Ignatyev, A.: The stochastic and singular analysis of fractal signals in cyber-physical systems of biomedicine. In: Kravets, A., Bolshakov, A., Shcherbakov, M. (eds.) Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, vol. 333, pp. 239–252. Springer, Cham (2021)
19. Metropolis, N., Ulam, S.: The Monte Carlo method. J. Amer. Stat. Assoc. **44**, 335–341 (1949)
20. Andrushevsky, N.M.: Stability Analysis of Solutions of Systems of Linear Algebraic Equations. MAX Press, Moscow (2008). (in Russian)
21. Biberdorf, E.A., Popova, N.I.: Guaranteed Accuracy of Modern Linear Algebra Algorithms. Publishing House of SB RAS, Novosibirsk (2006). (in Russian)
22. Bertero, M., Dovi, V.: Regularized and positive-constrained inverse methods in the problem of object restoration. Opt. Act. **28**(12), 1635–1649 (1981)
23. Engl, H.W., et al.: Regularization Methods for Invers Problems. Kluwer academic publishers (1996)
24. Gebali, F.: Solving systems of linear equations. In: Algorithms and Parallel Computing, pp. 305–321. John (2011)
25. Hoang, N.S., Ramm, A.G.: Solving ill-conditioned linear algebraic systems by the dynamical systems method (DSM). Inverse Prob. Sci. Eng. **16**(5), 617–630 (2008)
26. Hoerl, A.E., Kennard, R.W.: Ridge regression: biased estimation for non-orthogonal problems. Technometrics **12**, 55–67 (1970)
27. Kravets, A.G., Egunov, V.: The software cache optimization-based method for decreasing energy consumption of computational clusters. Energies **15**(20), 7509 (2022). https://doi.org/10.3390/en15207509

# Analysis of Clinical Data in Pregnant Women with Hypertension Based on a New Stochastic Approach

**Vladimir Kulikov, Svetlana Kolobova, Liudmila Borovkova, and Alexander Ignatyev**

**Abstract** Many world researchers have focused over recent years at studying risks associated with the development of pathologic pregnancy in women with arterial hypertension. The pathogenesis of gestational complications in case of hypertensive disorders is complex and associated mainly with endothelial dysfunction and reduced elasticity of a vascular wall. Nevertheless, the high prognostic value of the vascular elasticity assessment has been currently proven for the development of cardiovascular accidents in patients with hypertension in the general population. In order to early predict any risk of gestational complications, it is promising to study the state of the vascular wall elasticity based on the stochastic analysis, thus, optimizing the pregnancy management and improving perinatal outcomes in pregnant women with high blood pressure. The basis for the approach to the probabilistic (stochastic) analysis of clinical findings in pregnant women shall be methods for the correct identification of random values distribution laws in biomedicine. They are based on the solutions of inverse ill-posed problems elaborated by the school of Academician A.N. Tikhonov. To study data of the 24-h blood pressure (BP) monitoring in the course of the pregnant women supervision, distribution laws (probability distribution densities) have been identified, which have, among others, some polymodal features reflecting the current state of patients. Based on the stochastic analysis two marker indicators have been identified (the reflected wave transit time RWTT and the estimated pulse wave velocity in the aorta PWVao).

**Keywords** Pregnancy · Arterial hypertension · Endothelial dysfunction · Vascular wall elasticity · Risk prediction · Stochastic analysis · Ill-posed problems · Polymodal distribution densities

---

V. Kulikov (✉)
Nizhny Novgorod State Technical University N. a. R. E. Alekseev, 24 Minin St., Nizhny Novgorod 603950, Russia
e-mail: vb.kulikov@yandex.ru

S. Kolobova · L. Borovkova
Privolzhsky Research Medical University, 10/1, Minin and Pozharsky Sq., Nizhny Novgorod 603950, Russia

A. Ignatyev
Medical Diagnostics, Ltd, 12B Novikov-Priboya St., Nizhny Novgorod 603064, Russia



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_2

## 1 Introduction

Arterial hypertension (AH) remains under current conditions as one of the most urgent problems for clinical medicine because of high risk for development of cardiovascular complications (stroke, myocardial infarction, cardiac and renal failure) and the mortality. An important trend of cardiology is the search for early markers of damaged target organs such as blood vessels, heart, kidneys and brain in patients suffering from hypertension [1]. Pregnancy hypertension is an interdisciplinary problem of cardiology and obstetrics as it results in an unfavorable gestational process and a perinatal pathology. AH is a condition when the systolic blood pressure (SBP) of ≥140 mm Hg and/or diastolic blood pressure (DBP) of ≥90 mm Hg is recorded in pregnant women. Hypertension complicates 10–20% of pregnancies in the world, 5–30% in Russia. It should be noted that within the recent two decades the frequency of pregnancy hypertension has increased by 40–50%. Such increasing AH frequency mainly occurs because of chronic forms thereof (hypertensive disease) associated with the morbidity of female obesity, diabetes mellitus, older average age of primiparas as well as the widespread use of assisted reproductive technologies [2].

Hypertensive disease (HD) raises the risk of major obstetric syndromes (preeclampsia, placental insufficiency, fetal growth retardation, preterm labor, detachment of normally situated placenta), operative delivery by caesarean section, perinatal pathology, the need for intensive care for newborns, the maternal and infant mortality [3].

The pathogenesis of pregnancy complications in women suffering from hypertensive disorders is complex, still understudied and not yet fully understood. It is largely associated with generalized dysregulation of vascular tone, endothelial dysfunction and impaired placentation. Structural changes in vascular walls are attended with their reduced elasticity (stiffening, higher rigidity), narrowing of the artery lumen and impaired blood and oxygen flow to the placenta and fetus. The literature contains some evidence that endothelial dysfunction develops even before any clinical signs of pregnancy and fetal complications [2–4]. Hence, the attention of domestic and foreign scientists to the assessment of the vascular wall stiffness has significantly increased in order to early predict the risk of the pregnancy pathology development in case of hypertension.

Based on the world and Russian studies conducted over the past decade, it has been shown that one of the evidence of target organ damage is the hypertensive vascular remodeling resulted from elevated blood pressure. To date, it has been proven that higher arterial wall stiffness is associated with the increased risk for the cardiovascular mortality of HD patients in the general population [1].

The arterial wall stiffness (rigidity) affects the capability of large vessels to smooth the pulsation caused by cardiac output and to convert shock pressure in the ascending aorta into the stable blood flow in peripheral vessels. The increased arterial wall stiffness aggravates in case of HD the damping function of the arteries, acceleration of reflected waves, the rise in systolic and pulse pressure, the reduction of diastolic pressure and the risk of myocardial ischemia. According to the HARVEST (Hypertension and Ambulatory Recording VEnetia STudy) the vascular wall stiffness may be higher even at an early stage of the disease; it enables to improve the risk stratification for cardiovascular accidents of HD patients in the general population [1, 5].

## 2 Current Technologies for Functional Diagnosing of the Cardiovascular System

The blood pressure (BP) is a dynamic physiological parameter. The 24-h fluctuations in blood pressure constitute the sum of responses to external pressor stimuli, spontaneous and regulatory fluctuations.

A highly informative method for the ambulatory blood pressure identification is the 24-h blood pressure monitoring (BPM) using the BPLab 24-h soft hardware monitoring system (Petr Telegin LLC, Nizhny Novgorod). The BPM is an applanation tonometry method. The BPLab system is designed for the automatic non-invasive measurement of blood pressure, pulse rate as well as for assessing any complementary parameters characterizing the physical activity and hemodynamics.

The principle of the BPM technique is as follows:

1. Blood pressure of a patient, who lives in natural living conditions and can freely move, is periodically measured using a portable wearable monitor. Such measurements are taken both day and night. The monitoring duration shall last at least 24 h.
2. Upon the monitoring completion a table of obtained blood pressure measurement data is analyzed in order to assess indicators of the 24-h BP profile.

Such wearable monitors ensure the BP measurement by indirect methods based on the analysis of the dependence of biological signal parameters on the value of overpressure in a cuff applied to the patient's shoulder. The BP measuring methods used herein are oscillometric and auscultatory ones. The said monitors provide the recording of primary signals for each measurement, the patient's body position and activity within time periods corresponding to BP measurements taken.

The BPLab software is designed for BP and pulse wave monitoring; it is provided in Standard and Advanced versions.

To diagnose HD, the Standard version of the BPLab software is used, which provides the analysis of the 24-h peripheral BP profile with the calculation of statistical indicators (mean values, variability, pressure load indices SBP, DBP, mean blood pressure, pulse rate per 24 h, day, night and within a special interval).

To stratify the risk of target organ damage in hypertension, an innovative advanced version of the BPLab software with Vasotens 24 technology is used, which enables to assess daily trends in hemodynamic parameters (arterial stiffness and central aortic pressure findings).

The functional capabilities of the BPLab Standard and Advanced software are presented in Table 1.

The "Gold Standard" for measuring aortic stiffness is the valuation of the aortic carotid-femoral pulse wave velocity (Pulse Wave Velocity, PWV) using the dual pulse wave transducer technique. This factor is directly correlated with the vascular wall stiffness and is inversely proportional to its elasticity. Some recent studies show that the aortic stiffening assessed by the degree of the increased PWV therein is an independent predictor of the risk for cardiovascular complications in HD patients [1, 5].

However, as shown in [6], the traditional PWV (two-point) measurement has some disadvantages. In particular, the study procedure is time-consuming and needs sufficient work experience. Furthermore, the artery elasticity and, hence, the measured PWV depend essentially on the BP value at the time of measurement. Thus, the instantaneous PWV measurement shall only be valid for the pressure, thereat the measurement has been taken, so, it limits the use thereof when comparing subjects and evaluating patient-specific change dynamics. This parameter is not identified in the BPLab software.

To simplify the measurement procedure, some authors have suggested alternative rigidity indices, which correlate quite well with the traditionally measured PWV but are easier to use. The said indices include [7, 8]:

– reflected wave transit time (RWTT);
– aortic pulse wave velocity determined by the propagation time for the reflected wave, PWVao.

Moreover, there are some indices, which have independent significance but also essentially depend on the rigidity of the main arteries. These indices include:

– augmentation index (AIx), an indicator, which characterizes, first of all, the intensity of the reflected wave and its furthering the raise of the pulse blood pressure;
– the maximum rate of increasing blood pressure (dP/dt)max, an indicator, which characterizes, first of all, the myocardial contractility.

To assess hemodynamic parameters, the central aortic pressure (CAP) is also measured, the level thereof is regulated by elastic properties of large vessels, structural and functional elements of medium-sized arteries and resistive vessels, and is the most informative indicator as compared to peripheral blood pressure, which identifies the state of the arterial wall [1].

European experts consider the vascular wall stiffness and CBP as independent markers of the cardiovascular risk for HD patients in the general population [1, 5].

High blood pressure in pregnant women is an indication for the BPM in order to diagnose hypertension, its severity and differential diagnosis of hypertensive disorders [9]. But the experience of using the BPM to assess the arterial wall stiffness and the risk of gestational complications has shown that the results of the study of pregnant women suffering from hypertension are less informative and contradictory.

**Table 1** Capabilities of the BPLab software versions

| Daily trends under study | Description | Standard edition | Extended edition |
|---|---|---|---|
| Systolic BP (SBP) | The highest value of excessive blood pressure in the artery (reached at the moments of ventricular systole) | + | + |
| Diastolic BP (DBP) | The lowest value of excessive blood pressure in the artery (reached at the end of the diastolic period) | + | + |
| Mean hemodynamic blood pressure (MAP) | The resultant of all those pressure variables which occur during one involution of the heart | + | + |
| Pulse rate (PR) or heart rate (HR) | Number of heartbeats recorded per a minute | + | + |
| Pulse BP (SBP–DBP) | Difference between systolic and diastolic pressure | + | + |
| Double product index (SBP * PR)/100 | Myocardial oxygen consumption index | + | + |
| (dP/dt) max (maximum rate of pressure rise) | Maximum time derivative of arterial pressure (at the leading edge of the pulse wave), mm Hg. p. s | – | + |
| RWTT (reflected wave transit time) | Delay of a reflected wave relative to a direct wave | – | + |
| PWVao (estimated pulse wave velocity in the aorta) | To be determined by the propagation time of the reflected wave, m/s | – | + |
| AIx (Augmentation Index) | The ratio of amplitudes of the pulse wave direct component and component thereof reflected from the aorta bifurcation, % | – | + |
| SBP (central aortic systolic pressure) | Averaged form of pulsations in the ascending aorta | – | + |
| AIxao (aortic augmentation index) | The ratio of augmentation pressure (the difference between the amplitude determined by the direct wave and the amplitude at the time of maximum direct and reflected wave summation) to the pulse wave amplitude, % | – | + |
| PPA (pulse pressure amplification) | The ratio of the pulse pressure in the brachial artery to the central pulse pressure | – | + |
| ED (left ventricular ejection duration) | Time interval T from the onset of pulsation to the incisure (the time when the aortic valve closes) | – | + |

(continued)

**Table 1** (continued)

| Daily trends under study | Description | Standard edition | Extended edition |
|---|---|---|---|
| SEVR (subendocardial blood flow efficiency index) | The ratio of areas under the aortic pressure pulsation curve corresponding to vascular diastole (time period when the aortic valve is closed) and vascular systole (time period when the aortic valve is open), % | − | + |

Thus, most of the vascular elasticity studies were carried out in pregnant women during the preeclampsia development (trimester II–III). Some studies showed the increased PWV and AIx as compared to normotensive pregnant women [10], others noted the increased PWV with the unchanged augmentation index [11] and in some studies these parameters did not differ from physiological pregnancy [12]. There is also information in the literature that in women with clinical evidence of preeclampsia only an increased augmentation index has been identified, while other stiffness parameters have not been studied [13].

Furthermore, with preeclampsia developed at full-term pregnancy an increase in the PWV and AIx was demonstrated with higher levels of these parameters remaining also in the postpartum period, so, it may be evidence of persistent changes in the vascular wall. The same authors examined women with a history of preeclampsia. Findings of high augmentation index, pulse wave velocity, reflected wave propagation time and the arterial stiffness index in the second trimester have been obtained, which may evidence more rigid vessels as compared to pregnant women without a history of preeclampsia [10].

In the studies of pulse hemodynamics conducted in the 2nd–3rd trimester of pregnancy it was noted that in women with hypertension the central systolic blood pressure, central diastolic blood pressure, central mean hemodynamic pressure and central pulse blood pressure are higher than those in physiological pregnancy. Pregnant women with hypertension complicated by preeclampsia had thereat the highest levels of the central systolic BP and central diastolic BP [14]. There is also information on patients with AH whose central systolic BP and central pulse BP have been lower throughout the entire pregnancy than the values of peripheral systolic BP and peripheral pulse BP and it may indicate the reduced elasticity of the main vessels in case of cardiovascular diseases [14, 15].

The studies conducted in the second pregnancy trimester show that in women with HD the augmentation index, arterial stiffness index and the maximum rate of blood pressure rise are higher under clinical manifestations of preeclampsia, while the propagation time of the reflected wave and pulse pressure amplification do not differ from those of normotensive women [16, 17].

In single researches, the prognostic CBP value and the effect of difference between the central and peripheral arterial pressure on the gestation progression and outcome have been studied during the examination of women in the second half of pregnancy. Thus, when conducting the BPM in the second trimester the predictors of the fetal growth retardation development in pregnant women with hypertension are the decreased elasticity of the vascular wall and the reduced difference between the CBP and peripheral blood pressure. The difference between the central and peripheral systolic blood pressure equal to $8.3 \pm 0.9$ mm Hg.; the difference between central and peripheral pulse blood pressure equal to $9.8 \pm 1.0$ mm Hg; the level of central systolic blood pressure $\geq$118 mm Hg; peripheral diastolic blood pressure $\geq$82 mm Hg; peripheral mean hemodynamic pressure $\geq$99 mm Hg are of prognostic significance for the risk of the fetal growth retardation. The direct relationship of birth complications (premature birth, untimely discharge of amniotic fluid, a long anhydrous period, anomalies in labor) and increasing central systolic blood pressure in the II–III trimesters $\geq$118 mm Hg was identified in the studies of the same authors [18].

Thus, to date in the most studies the pulse wave analysis has been carried in the second half of pregnancy, the elastic properties of blood vessels have been examined during the clinical manifestations of preeclampsia and the prognostic significance of single vascular stiffness parameters have been assessed in the second half of pregnancy. It is relevant to study the prognostic significance of arterial wall stiffness parameters in the first trimester of pregnancy in order to stratify the risk of gestational complications in pregnant women with HD. The early prediction of major obstetric syndromes shall optimize the pregnancy management, improve perinatal outcomes and reduce the maternal and infant mortality in women with HD.

## 3 New Approach to the Processing of Clinical Findings in Pregnant Women

Purpose of the study is the development of marker indicators for the vascular wall stiffness in order to predict major obstetric syndromes in pregnant women with hypertension based on the stochastic analysis.

### *3.1 Materials and Methods*

The study of 19 pregnant women has been carried out who suffer from stage I hypertensive disease, arterial hypertension of stage 1, risk 1–2.

All the said women have been examined under Order N 1130n of the Ministry of Health of the Russian Federation dated October 20, 2020 On the Approval of the Procedure for the Delivery of Medical Care in Obstetrics and Gynecology.

All those pregnant women with gestational age of 10–13, 18–21 and 30–34 weeks have had the 24-h blood pressure monitoring and the arterial stiffness assessment using the BPLab 24-h BP Monitoring System with Vasotens24 technology (Petr Telegin LLC, Nizhny Novgorod).

The basis for the approach to the probabilistic (stochastic) analysis of clinical data of pregnant women are the methods for correct identification of random variable distribution laws in biomedicine first published by the European Journal in 2015 [19]. They are based on the solutions of inverse ill-posed problems elaborated by the Academician A.N. Tikhonov School of Mathematics [20].

The theoretical basis for the method of identifying the distribution densities of random variables is the approximate solution of the first kind Fredholm equation and the principle of structural risk minimization. The algorithmic basis for the method consists in solving equivalent systems of linear equations to find expansion coefficients of the desired density with respect to the chosen system of basis functions.

On the premise that the shape of the distribution density is smooth the distribution laws are restored on the set of trigonometric functions with restricting the number of expansion terms N depending on the volume L of observed data by minimizing the guaranteed risk.

The basis for the approach to the probabilistic (stochastic) analysis of clinical data of pregnant women are the methods for correct identification of random variable distribution laws in biomedicine first published by the European Journal in 2015 [19]. They are based on the solutions of inverse ill-posed problems elaborated by the Academician A.N. Tikhonov School of Mathematics [20].

The theoretical basis for the method of identifying the distribution densities of random variables is the approximate solution of the first kind Fredholm equation and the principle of structural risk minimization. The algorithmic basis for the method consists in solving equivalent systems of linear equations to find expansion coefficients of the desired density with respect to the chosen system of basis functions.

### *3.2 Stochastic Perinatal Diagnosis*

On the premise that the shape of the distribution density is smooth the distribution laws are restored on the set of trigonometric functions with restricting the number of expansion terms N depending on the volume L of observed data by minimizing the guaranteed risk.

The outcomes of studies based on the suggested identification method have previously been used in medical applications: in gastroenterology and clinical immunology. For example, a fractal model of an EGEG signal (electrogastroenterographic signal) is considered in publications [19, 21, 22] as a fractional Brownian motion in terms of polymodal distributions.

In the conducted studies, the problems of identifying the laws of distribution for local sections of the EGEG signal (selective realizations) are solved. The identified polymodal characteristics contain information on the dynamics in states of the gastrointestinal tract (GIT) of patients, enable to identify classification features and calculate the classical statistical parameters of distributions.

As compared to the known methods of diagnosis and prediction, the suggested algorithms can identify the evolution of the distribution densities of random processes nearly in real time, provide valuable information for verifying stochastic solutions and allow the effective use of entropy analysis methods. The high sensitivity of entropy algorithms to changes in biomedical data parameters makes the suggested approach be in demand for the effective selection of drugs and high-quality individual therapy.

The elaborated approaches have enabled to successfully solve problems in practical medicine associated with monitoring of the female patients' condition within the perinatal period (stochastic perinatal diagnosis). Early in 2021 the researchers from the Nizhny Novgorod State Technical University n.a. R.E. Alekseev (NSTU) and the Privolzhsky Research Medical University (PRMU) carried out the following joint works:

1. Complex methods of stochastic analysis were used for the correct study of data obtained during the 24-h blood pressure monitoring in the course of the supervision for pregnant women.
2. The processing of medical data obtained in 2020 through monitoring of the cardiovascular system in a group of pregnant women under supervision was carried out as per eight important indicators based on the identification methods discussed in [19].
3. Distribution laws (probability distribution densities) were identified, which, among others, had polymodal features reflecting the current state of the patients (over 150 characteristics).
4. Based on the stochastic analysis two marker indicators were identified (the reflected wave transit time—RWTT and the estimated pulse wave velocity in the aorta—PWVao), which characterized the vascular system rigidity and were important for predicting the pregnancy course and successful delivery. Differentiable differences for a group of patients with norm and pathology are 15–20% and are reliably identified by the suggested method.
5. The identified marker features are confirmed by the study within the fractal concept framework by the Hurst method and Fourier spectral analysis method.

Figure 1 shows for illustration some actually "averaged" identified marker indicators of successfully delivered women (distribution density), which quite well correlate with individual characteristics of pregnant women. The distributions of births with pathology have characteristic shifts in density extrema, including polymodal features of the PWVao parameter (Fig. 2).

Figure 3 illustrates the significant variability of other important indices.

The identified distribution densities of the Aixao index (aortic augmentation index) for two successfully delivered patients (samples of 72 and 56 of counts, respectively) have some pronounced polymodal features and different clustering centers of the most probable values. It indicates the effect of several significant pregnancy factors on this indicator. Thus, for correct interpreting the Aixao index, it is crucially

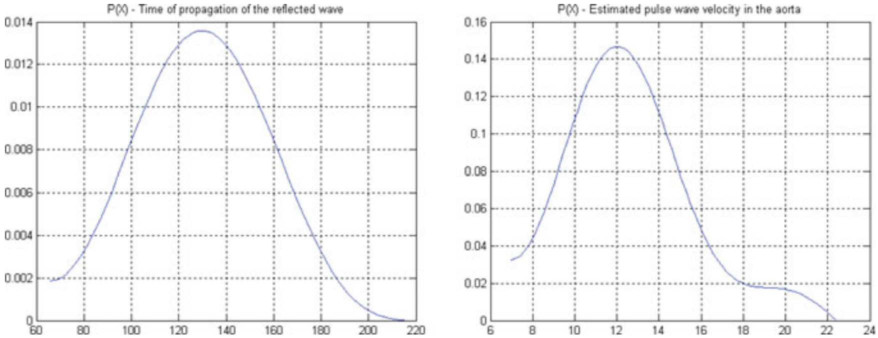



**Fig. 1** Identified distribution densities of marker indicators RWTT and PWVao in the group of successfully delivered women (combined daily samples 569 and 633 of count, respectively)

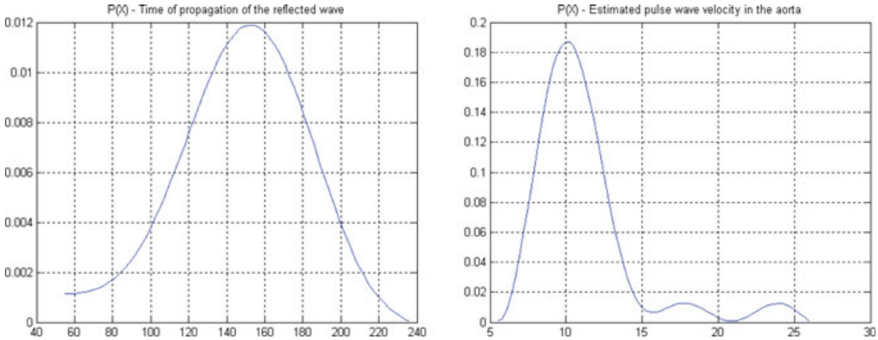



**Fig. 2** Identified distribution densities of marker indicators RWTT and PWVao in the group of births with pathology (combined samples of 168 counts)

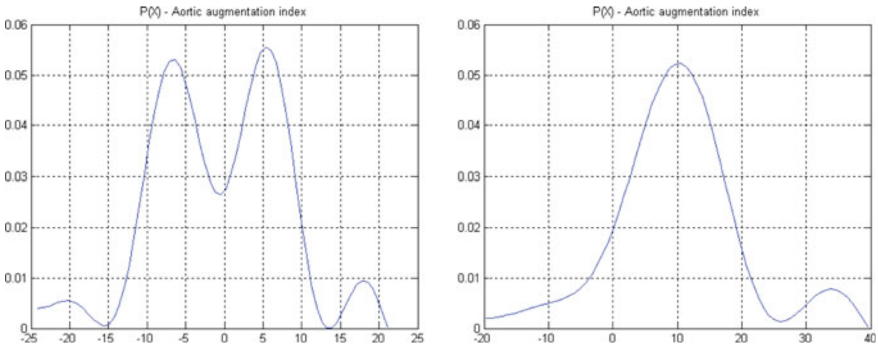



**Fig. 3** Identified distribution densities of the Aixao index for successfully delivered women (samples 72 and 56 of counts, respectively)

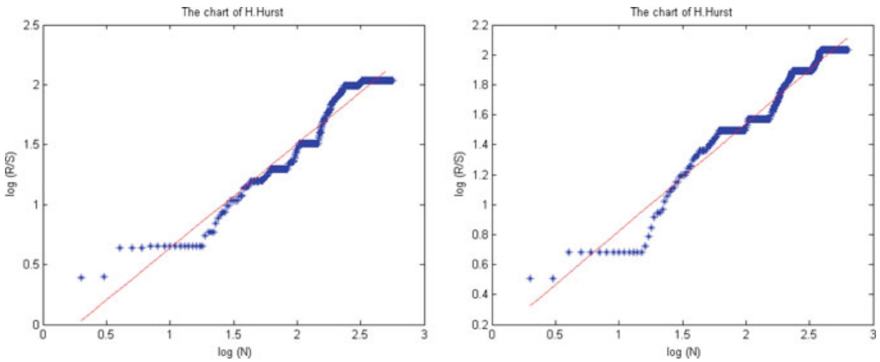



**Fig. 4** To the analysis of RWTT and PWVao indicators (the group of successfully delivered women) by the R/S method

required to take into account the polymodality and additional analysis. The distribution densities of the Aixao index in the group of those who gave birth successfully are generally bimodal.

The marker features of the reflected wave—RWTT and the estimated pulse wave velocity in the aorta—PWVao identified by the stochastic method have been successfully confirmed by the R/S analysis (Hurst method) [23, 24]. Figure 4 shows plots of dependencies lg(R/S) on lg(N) of successfully delivered patients (17 women). N is here the number of counts in the combined sample. These dependences enable us to calculate the Hurst exponent H: for RWTT it is 0.868, for PWVao H = 0.717.

Note that for the group of patients with birth pathology the values H = 0.828 and H = 0.986, respectively, have been obtained. Taking into account the high sensitivity of the Hurst index, the differences in the selected groups can be considered as significant. Moreover, according to the theory of fractals this indicator means at H > 0.5 the persistence of the analyzed indicators during the observation period.

This new proven diagnostic method is planned for the implementation at the Regional Clinical Hospital n.a. N.A. Semashko (Department of Obstetrics and Gynecology of the Privolzhsky Research Medical University).

## 4 Results and Conclusions

In order to early predict any risk of gestational complications, it is promising to study the state of the vascular wall elasticity based on the stochastic analysis, thus, optimizing the pregnancy management and improving perinatal outcomes in pregnant women with high blood pressure.

The basis for the approach to the probabilistic (stochastic) analysis of clinical findings in pregnant women shall be methods for the correct identification of random values distribution laws in biomedicine. To study data of the 24-hourblood pressure

(BP) monitoring in the course of the pregnant women supervision, distribution laws (probability distribution densities) have been identified, which have, among others, some polymodal features reflecting the current state of patients.

Based on the stochastic analysis two marker indicators have been identified (the reflected wave transit time RWTT and the estimated pulse wave velocity in the aorta PWVao), which characterize the vascular system rigidity and are important for predicting the pregnancy progression and successful delivery.

## References

1. Williams, B., Mancia, G., Spiering W. et al.: 2018 ESC/ESH guidelines for the management of arterial hypertension. The task force for the management of arterial hypertension of the European society of cardiology (ESC) and the European society of hypertension (ESH). J. Hypertens. **36**(12), 2284–2309 (2018)
2. Rouse, C.E., Eckert, L.O., Wylie, B.J., Lyell, D.J., Jeyabalan, A., Kochhar, S. et al.: The Brighton collaboration preeclampsia working group: hypertensive disorders of pregnancy: case definitions & guidelines for data collection, analysis, and presentation of immunization safety data. Vaccine. **34**, 6069–6076 (2016)
3. Anthony, J., Damasceno, A., Ojjii, D.: Hypertensive disorders of pregnancy: what the physician needs to know. Cardiovasc J. Afr. **27**(2), 104–110 (2018)
4. Rhodes, C., Beevers, D., Churchill, D.: A randomized trial of ambulatory blood pressure monitoring versus clinical blood pressure measurement in the management of hypertension in pregnancy. A feasibility study. Pregnancy Hypertens. **11**, 142–144 (2018)
5. Townsend, R.R., Wilkinson, I.B., Schiffrin, E.L., Chirinoset, J.A., Cockcroft, J.R., Heffernan, K.S., et al.: Recommendations for improving and standardizing vascular research on arterial stiffness: a scientific statement from the American heart association. Hypertension **63**(3), 698–722 (2015)
6. Omboni, S., Posokhov, I.N., Rogoza, A.N.: Evaluation of 24 hour arterial stiffness indices and central hemodynamics in healthy normotensive subjects versus treated or untreated hypertensive patients: a feasibility study. Int. J. Hypertens. **2015**(1), 1–10 (2015)
7. Mancia, G., Laurent, S., Agabiti-Rosei, E. et al.: Reappraisal of European guidelines on hypertension management: a European society of hypertension task force document. J. Hypertens. **27** (2009)
8. Mitchell, G.F., Parise, H., Benjamin, E.J., et al.: Changes in arterial stiffness and wave reflection with advancing age in healthy men and women: the Framingham heart study. Hypertension **43**, 1239–1245 (2004)
9. Phipps, E., Prasanna, D., Brima, W., Jim, B.: Preeclampsia: updates in pathogenesis, definitions and guidelines. Clin. J. Am. SocNephrol. **11**(6), 1102–1113 (2016)
10. Sufrin, S., Nessa, A., Islam, M.T., Khatun, A.A., Husain, M.F., Khatun, N., et al.: Blood pressure in third trimester of pregnancy. Mymensingh Med J. **25**(1), 18–22 (2016)
11. Cusimano, M.C., Pudwell, J., Roddy, M., et al.: The maternal health clinic: an initiative for cardiovascular risk identification in women with pregnancy-related complications. Am. J. Obstet. Gynecol. **210**(438), 1–9 (2014)
12. Nagar, T., Sharma, D., Choudhary, M., Khoiwal, S., Nagar, R.P., Pandita, A.: The role of uterine and umbilical arterial Doppler in high-risk pregnancy: a prospective observational study from India. Clin. Med. Insights: Reprod. Health. **9**, 1–5 (2015)
13. Cífková, R.: Should 24-hour ambulatory blood pressure monitoring become part of regular checkup in women with previous preeclampsia? Hypertension **71**, 59–60 (2018)
14. Varnier, N., Brown, M.A., Reynolds, M., et al.: Indications for delivery in pre-eclampsia. Pregnancy Hypertens. **11**, 12–17 (2018)

15. Seely, E.W., Ecker, J.: Chronic hypertension in pregnancy. Circulation **129**(11), 1254–1261 (2014)
16. Parikh, N.I., Norberg, M., Ingelsson, E., et al.: Association of pregnancy complications and characteristics with future risk of elevated blood pressure: the Västerbotten intervention program. Hypertension **69**(3), 475–483 (2017)
17. Antza, C., Cifkova, R., Kotsis, V.: Hypertensive complications of pregnancy: a clinical overview. Metabolism **86**, 102–111 (2018)
18. Dorogova, I.V., Usanov, V.D., Bartosh, L.F., Panina, E.S., Bocharnikov, D.Y.: Features of parameters of peripheral, central blood pressure and indicators of vessels rigidity in pregnant women with hypertensive disorders. Regionarnoe krovoobrashchenie i mikrotsirkulyatsiya. **2**(58), 51–59 (2016) (in Russian)
19. Kulikov, V.: The identification of the distribution density in the realization of stochastic processes by the regularization method. Appl. Mathem. Sci. **9**(137), 6827–6834 (2015)
20. Tikhonov, A.N., Arsenin, V.Y.: Methods for solving incorrect problems. Nauka, Moscow (1986) (in Russian)
21. Kulikov, V., Kulikov, A., Ignatyev, A.: The Stochastic and singular analysis of fractal signals in cyber-physical systems of biomedicine. In: Kravets, A., Bolshakov, A., Shcherbakov, M. (eds.) Society 5.0: cyberspace for advanced human-centered society. Studies in systems, decision and control, vol. 333, pp. 239–252. Springer, Cham (2021)
22. Kulikov, V., Kulikov, A., Khranilov, V., Ignatyev, A.: The Stochastic properties and singular analysis of Hankel trajectory matrices for random time series in biomedicine. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Society 5.0: human-centered society challenges and solutions. Studies in systems, decision and control, vol. 416, pp. 41–55. Springer, Cham (2022)
23. Peters, E.E.: Chaos and Order in the Capital Markets: A New View of Cycles, Prices and Market Volatility. John (1996)
24. Feder, J.: Fractals. Plenum Press, New York, London (1989)

# Hierarchical Fuzzy Inference of Adequacy of Highly Informative Diagnostic Signs of Breast Cancer

**Ilya Germashev, Victoria Dubovskaya, and Alexander Losev**

**Abstract** This chapter presents an approach to determining the influence of highly informative diagnostic signs on the classification result in the cyber-physical systems of diagnosis of patients in medicine. The concept of adequacy is introduced as a generalizing criterion, combining such contradictory characteristics as the level of influence on the correctness and erroneous of the diagnosis. A computational experiment was carried out to establish the level of adequacy of features within the framework of a specific classification method. The available data set contained strongly correlating features and allowed factor analysis only after removing a number of features. This problem was solved on the basis of a factor model with a reduced number of features, to which a fuzzy inference algorithm was applied. This made it possible to approximate the factor loads for the remaining features that were not included in the factor model.

**Keywords** Cyber-physical systems · Microwave thermometry · Breast cancer · Factor analysis · Fuzzy inference

## 1 Introduction

At the moment, cyber-physical systems have found their wide application in healthcare (see, for example, [1, 2]). Similar systems are also used as part of the task of diagnosing breast cancer based on thermometric data. The method of microwave radio thermometry (RTM) consists in determining the surface (skin) and deep body temperatures by measuring the intensity of its electromagnetic radiation.

I. Germashev (✉) · V. Dubovskaya · A. Losev
Volgograd State University, 100 Prospect Universitetsky, Volgograd 400062, Russia
e-mail: germashev@volsu.ru

V. Dubovskaya
e-mail: dubovskaja@volsu.ru

A. Losev
e-mail: alexander.losev@volsu.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_3

Measurements are carried out using antenna applicators, which are installed on the surface of the patient's body. The obtained temperature data are used by an artificial intelligence system, the purpose of which is to identify thermal anomalies inside the breast.

In [3–5], diagnostic problems and their solution by radio thermometry methods are presented. To date, within the framework of the conducted research, specialists have built:

1. A conceptual model of the behavior of temperature fields, which describes aspects characteristic of various diagnostic classes;
2. A mathematical model that formalizes heuristic hypotheses for the behavior of temperatures;
3. Feature space, which is used when classifying patients using machine learning methods.

The results of stages 1–3 can be found, for example, in the works of A.G. Losev, V.V. Levshinsky, E.A. Mazepa (for example, [6, 7, 14]). It should also be noted a number of studies that show that the use of microwave radiothermometry in combination with artificial intelligence methods shows high efficiency in the early diagnosis of diseases (see, for example, [8–13]).

The above approaches make it possible to apply machine learning methods to build classifier models that solve the problems of diagnosing a disease risk group in patients. This raises the question of what effect each feature has on the classification result. This knowledge allows, firstly, to identify the causes of the result obtained, which can form the basis for the creation of a recommender system to support decision-making by a specialist. Secondly, to correct the classifier model to obtain more adequate results.

This problem can be solved through the joint use of factor analysis and fuzzy inference algorithms.

## 2 Materials and Methods of Research

The measuring complex RTS-01-RES is used for the diagnosis of mammary glands. During the measurements, 44 temperatures are recorded in the RTM-range (deep) and IR-range (skin). The scheme of measuring the temperature of the mammary glands is shown in Fig. 1.

A sample of temperature data can be represented as:

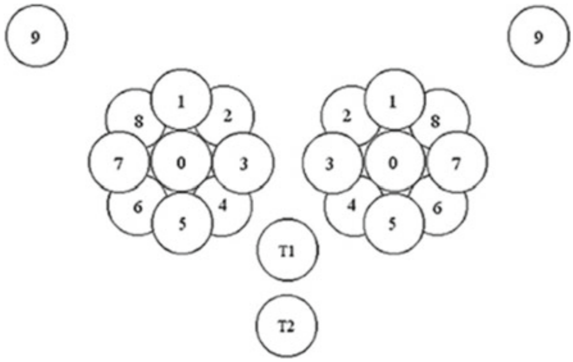



**Fig. 1** Measurement scheme

$$S = \begin{pmatrix} t_0^1 & \cdots & t_{43}^1 \\ \vdots & \ddots & \vdots \\ t_0^k & \cdots & t_{43}^k \\ t_0^{k+1} & \cdots & t_{43}^{k+1} \\ \vdots & \ddots & \vdots \\ t_0^n & \cdots & t_{43}^n \end{pmatrix},$$

where $t_i^j$, i = 0, …, 9—the i-th point temperature of the right breast of the j-th patient (RTM-range);

$t_i^j$, i = 10, …, 19—the [i–10]-th point temperature of the right breast of the j-th patient (IR-range);

$t_i^j$, i = 20, …, 29—the [i–20]-th point temperature of the left breast of the j-th patient (RTM-range);

$t_i^j$, i = 30, …, 39—the [i–30]-th point temperature of the left breast of the j-th patient (IR-range);

$t_{40}^j = T1$, $t_{41}^j = T2$—the reference points temperatures of the j-th patient (RTM-range);

$t_{42}^j = T1$, $t_{43}^j = T2$—the reference points temperatures of the j-th patient (IR-range).

As a result of the analysis carried out by specialists, qualitative signs of breast cancer were determined on the basis of temperature data. Among them, we can note:

- an increased amount of thermal symmetry between separate points in the affected mammary gland;
- increased temperature variation between individual points in the breast pore;
- the ratio of skin and depth temperatures;
- nipple temperature difference;

- increased nipple temperature in the affected mammary gland compared to the average breast temperature, taking into account the age-related temperature changes.

For each qualitative sign of cancer, it is possible to construct a set of modeling functions from temperatures, the value of which characterizes the sign. As an example, here are several functions that simulate:

Asymmetry of the temperature fields of the mammary glands:

$$\sum_{i=0}^{9} |t_i - t_{i+20}|;$$

1. The ratio of skin and depth temperatures:

$$\max_{i=\overline{0,9}} (t_i - t_{i+10}).$$

For more information about the process of constructing such functions, see [6]. As a result, the values of the features $p_{jr}$, where $r = 1, \ldots, s$, for each patient were calculated. And these signs are denoted by MG001, …, MG062, that is, s = 62.

In this work, a database of thermometric data containing information about patients of cancer centers was used. To conduct the study, the database was transformed in such a way that each element corresponds to a separate mammary gland of the patient.

The degree of breast disease is assessed on a 6-point scale [15]:

- th0—there are no temperature anomalies; a healthy mammary gland;
- th1—temperature anomalies were detected, but they are not typical for breast cancer (lower temperatures);
- th2—temperature anomalies were detected, but they are not typical for breast cancer (higher temperatures);
- th3—a high level of skin and deep temperatures;
- th4—multiple temperature anomalies were found;
- th5—thermogram is typical for patients with acute exacerbation or breast cancer.

The diagnosis is translated into a numerical type on this scale. Namely, we will assume that in cases of th0-th2 the mammary gland is healthy, in cases of th3-th5—the mammary gland is at risk group.

The entire data sample H contains 15,116 elements. The diagnostic class “Healthy” contains 13,570 elements, the diagnostic class “Risk group”—1546 elements.

The diagnostic state of the patient is described by 62 functions, which are a mathematical formalization of the qualitative signs of cancer.

To conduct factor analysis, it was necessary to classify all records from the database with one of the machine learning algorithms. Logistic regression was chosen as such an algorithm [16]. To classify each record, it was decided to use the cross-validation method. The method consists in sequential training and testing of the classification algorithm on different subsamples, mixed with each other in the training sample. To do this, the entire database was divided into 10 equal and non-overlapping parts (the maximum difference in their size is two records). Each part contains an equal number of records of each class. Next, 9 out of 10 parts were placed in the sample, on which the logistic regression was trained. The remaining part, not included in the training sample, was classified. The classification results were stored in the final database H. Thus, in addition to the real class, it also stores the class assigned by the logistic regression. After all the records from the considered part were classified, the algorithm was retrained on a new training set. The new sample was drawn up in such a way that one of the parts of the previous sample was interchanged with the classified sample. The classification process was repeated 10 times until all records from the database were classified. Thus, logistic regression classified correctly 90% of healthy mammary glands and 88% of glands at risk group.

Based on the obtained classification result, base $H$ was divided into two samples:

- $H^0$, consisting of 2766 elements, representing breast data for which the algorithm made an erroneous classification.
- $H^1$, consisting of 12,350 elements, representing breast data for which the classification algorithm determined the correct diagnosis.

For each sample, factor analysis was performed separately using maximum likelihood estimation.

To select factors, it is sufficient that the correlation matrix is positive definite, and also that the determinant of this matrix is not equal to zero.

However, these requirements are not faced on this set of features, which is confirmed by the analysis of the set using the Statistics program. This is due to the fact that some features have a strong correlation with each other (Table 1). Therefore, it was decided to remove such features with the subsequent restoration of their factor loadings using fuzzy inference.

The purpose of factor analysis is to obtain factor loadings of features. Factor loadings are the correlation of each sign with each identified factor. The closer the relationship of this feature with the factor under consideration, the higher the value of the factor loadings.

Thus, after constructing factor models, two characteristics are defined for each feature:

- the level of influence on the erroneous of the diagnosis made by the classifier (corresponding to the value of the factor load on the sample $H^0$);
- the level of influence on the correctness of the diagnosis made by the classifier (corresponding to the value of the factor load in the sample $H^1$).

It is proposed to combine these two, generally speaking, contradictory characteristics of adequacy, also using fuzzy inference.

**Table 1** Fragment of the matrix of feature correlations

| | MG001 | MG002 | MG003 | MG004 | MG005 | MG006 |
|---|---|---|---|---|---|---|
| MG001 | 1 | 0.977 | 0.797 | 0.669 | 0.036 | −0.092 |
| MG002 | 0.977 | 1 | 0.894 | 0.753 | 0.036 | −0.095 |
| MG003 | 0.797 | 0.894 | 1 | 0.806 | 0.034 | −0.094 |
| MG004 | 0.669 | 0.753 | 0.806 | 1 | 0.025 | −0.061 |
| MG005 | 0.036 | 0.036 | 0.034 | 0.025 | 1 | −0.576 |
| MG006 | −0.092 | −0.095 | −0.094 | −0.061 | −0.576 | 1 |
| MG007 | 0.350 | 0.394 | 0.430 | 0.477 | −0.308 | 0.433 |
| MG008 | −0.077 | −0.082 | −0.090 | −0.057 | −0.423 | 0.738 |

Let us present the definition of adequacy in this context. Adequacy is considered as the expectation of the researcher that the feature will have a positive impact on the correctness of the diagnosis, and vice versa, a negative impact on the misclassification.

For a correct description of the multidimensional dependencies "inputs-outputs" of the task of determining the level of adequacy of a feature within a specific classification method, it is proposed to use hierarchical fuzzy inference. In the course of this work, the following method was used: performing a fuzzy inference for intermediate variables (restoring factor loadings) with subsequent transfer of clear values of these variables to fuzzy systems of the next level of the hierarchy (determining the level of adequacy of a feature) [17–19].

## 3 Computational Experiment and its Discussion

For the selected 22 features (on the $H^0$) sample) and 25 features (on the $H^1$ sample), factor analysis was carried out using the Statistic software package (to conduct a factor analysis using the maximum likelihood estimation, only those features were left, the value of the correlation coefficient of which is less than 0.7).

The process of constructing factorial models is described in detail in [16]. As a result, a matrix of components for the selected features was obtained (Table 2).

High factor loadings obtained in the sample $H^1$ may mean that these variables contributed to the fact that the test sample instances fell into the class with the correct results of the classifier. For a sample $H^0$, on the contrary, may mean these features influenced the fact that the classifier made an incorrect conclusion on the test examples of the sample.

It is possible to calculate the level of influence for features missing in the factor model, using information about the level of correlation, with the help of fuzzy inference.

**Table 2** Factor matrix for the sample $H^0$ (factor extraction method: maximum likelihood estimation, factors extracted—5, iterations required - 8)

| Sign | Factor | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| MG001 | 0.353 | 0.259 | −0.396 | 0.494 | 0.166 |
| MG004 | 0.48 | 0.293 | −0.29 | 0.495 | 0.082 |
| MG005 | −0.302 | 0.746 | −0.146 | −0.165 | −0.175 |
| MG007 | 0.999 | −0.008 | 0.001 | −0.001 | −0.001 |
| MG008 | 0.313 | −0.536 | 0.154 | 0.012 | −0.134 |
| MG009 | 0.218 | −0.042 | −0.039 | −0.18 | 0.195 |
| MG012 | 0.267 | 0.323 | −0.01 | 0.65 | 0.146 |
| MG019 | 0.022 | 0.206 | 0.482 | 0.325 | 0.114 |
| MG023 | 0 | 0.12 | 0.175 | −0.004 | −0.097 |
| MG024 | −0.042 | 0.071 | 0.17 | 0.109 | 0.018 |

In particular, for variables that were not used in the process of constructing factor models, the levels of influence on correctness/erroneous were restored using Mamdani's fuzzy inference.

The following linguistic variables have been introduced:

(1) $\widetilde{C_i^1}$—"level of influence", which is a factor loading of the feature $i$, where $i \in I$—the numbers of the features that were used in the construction of the factor model;
(2) $\widetilde{C_i^2}$—"level of influence", which is a factor loading of the feature $i$, где $i \in J$—the numbers of features that were not used in the construction of the factor model;
(3) $\widetilde{R_{ij}}$—"correlation"—the value of the correlation coefficient between $p_{ir}$ and $p_{jr}$, где $r = 1, \ldots, s$.

Linguistic variables are characterized by a set of their meanings (terms): "medium strength", "strong", "significant strength".

A formal description of the linguistic variable "correlation" can be represented as follows:

$$\widetilde{R^1} = \big(\beta = \text{“}correlation\text{”}, T = \big\{t_1 = \text{“}medium\,strength\text{”}, t_2 = \text{“}strong\text{”}, \\ t_3 = \text{“}significant\,strength\text{”}\big\}, X = [0.7, 1], G = \emptyset, M = \emptyset\big),$$

where $t_1$—a term described by a fuzzy variable.

$\hat{t}_1 = (\alpha = \text{“}medium\ strength\text{”}, X = [0.7, 1], \mu_1(x) = \frac{-25x}{3} + \frac{41}{6}, 0.7 \le x \le 0.82)$;

$t_2$—*a term described by a fuzzy variable* $\hat{t}_2 = (\alpha = \text{“strong”}, X = [0.7, 1], \mu_2(x) = \begin{cases} \frac{100x}{13} - \frac{72}{13}, & 0.72 \le x \le 0.85, \\ \frac{-25x}{3} + \frac{97}{12}, & 0.85 < x \le 0.97. \end{cases})$;

*$t_3$—a term described by a fuzzy variable* $\hat{t}_3 = (\alpha =$ "*significant strength*", $X =$ [0.7, 1],

$$\mu_3(x) = \frac{100x}{13} - \frac{87}{13}, \quad 0.87 \le x \le 1).$$

A formal description of the linguistic variable "level of influence" can be represented as follows:

$$\widetilde{C^1} = \big(\beta = \text{``} \, level\, of\, influence \text{''}, T = \big\{t_1 = \text{``} \, medium\, strength \text{''}, t_2 = \text{``} \, strong \text{''}, t_3 = \text{``} \, significant\, strength \text{''}\big\}, X = [0.5, 1], G = \emptyset, M = \emptyset\big),$$

where $t_1$—a term described by a fuzzy variable.

$\hat{t}_1 = (\alpha =$ "*medium strength*", $X = [0.5, 1]$ , $\mu_1(x) = -5|x| + 3.5, 0.5 \le x \le 0.7)$;

*$t_2$—a term described by a fuzzy variable* $\hat{t}_2 = (\alpha =$ "*strong*", $X = [0.5, 1]$, $\mu_2(x) = \begin{cases} \frac{100|x|}{21} - \frac{18}{7}, & 0.54 \le x \le 0.75, \\ -5|x| + 4.75, & 0.75 < x \le 0.95 \end{cases}$);

*$t_3$—a term described by a fuzzy variable* $\hat{t}_3 = (\alpha =$ "*significant strength*", $X =$ [0.5, 1],

$$\mu_3(x) = \frac{100|x|}{21} - \frac{79}{21}, 0.79 \le x \le 1).$$

A fragment of the rule base used, consisting of 9 rules:

Rule 1: If ($\widetilde{C_i^1}$ "significant strength") and ($\widetilde{R_{ij}}$ "significant strength"), then ($\widetilde{C_i^2}$ "significant strength").

Rule 2: If ($\widetilde{C_i^1}$ "strong") and ($\widetilde{R_{ij}}$ "significant strength"), then ($\widetilde{C_i^2}$ "strong").

Rule 3: If ($\widetilde{C_i^1}$ "medium strength") and ($\widetilde{R_{ij}}$ "significant strength"), then ($\widetilde{C_i^2}$ "medium strength").

To implement the Mamdani fuzzy inference algorithm, the Matlab program was used. After defuzzification of the output variables, the values of the level of influence of features were obtained, which did not participate in the construction of the factor model (Table 3).

Thus, thanks to the factor analysis, it was possible to establish the levels of influence of features on the correctness/erroneous of the classification result. Further use

**Table 3** Fragment of the table of restored factor loadings

| Sign | Level of influence $\widetilde{C_i^2}$ |
|---|---|
| MG002 | 0.67 |
| MG010 | 0.68 |
| MG016 | 0.79 |
| MG039 | 0.80 |

of the Mamdani fuzzy inference algorithm made it possible to level out the limitations of the maximum likelihood estimation associated with linear dependence and high feature correlation.

The obtained levels of influence serve as input variables for the fuzzy conclusion of the adequacy of signs [20]: the value of factor loadings on a sample with correct diagnoses is the linguistic variable “correctness”, the value of factor loads on a sample with erroneous diagnoses is “erroneous”.

These linguistic variables are characterized by many of their meanings (terms): “practically absent”, “small”, “medium”, “high”.

A formal description of the linguistic variable "correctness" can be represented as follows:

$$\widetilde{Cor^1} = \big(\beta = \text{“} correctness\text{”}, T = \big\{t_1 = \text{“} practically\, absent\text{”}, \\ t_2 = \text{“} small\text{”}, t_3 = \text{“} medium\text{”}, t_4 = \text{“} high\text{”}\big\}, X = [0, 1], G = \emptyset, M = \emptyset\big),$$

where $t_1$—*a term described by a fuzzy variable* $\hat{t}_1 = (\alpha =$ “*practically absent*”, $X = [0, 1]$, $\mu_1(x) = 2x, 0 \le x \le 0.5)$;

$t_2$—*a term described by a fuzzy variable* $\hat{t}_2 = (\alpha =$ “*small*”, $X = [0, 1]$,

$$\mu_2(x) = 1 - \frac{x - 0.5}{0.2}, 0.5 \le x \le 0.7);$$

$t_3$—*a term described by a fuzzy variable* $\hat{t}_3 = (\alpha =$ “*medium*”, $X = [0, 1]$,

$$\mu_3(x) = \begin{cases} 1 - \frac{0.75 - x}{0.2}, \; 0.55 \le x \le 0.75, \\ 1 - \frac{x - 0.75}{0.2}, \; 0.75 \le x \le 0.95 \end{cases});$$

$t_4$—*a term described by a fuzzy variable* $\hat{t}_4 = (\alpha =$ “high”, $X = [0, 1]$, $\mu_4(x) = 1 - \frac{1-x}{0.2}, 0.8 \le x \le 1)$.

Similarly, to the previous example, a base of fuzzy rules was built. As a result of fuzzy inference, in particular the stages of fuzzification of input variables, aggregation of subconditions, activation of subconclusions, accumulation of conclusions and defuzzification of output variables [20], clear values of the adequacy of features were obtained (Table 4).

**Table 4** Fragment of the table of the level of adequacy of signs

| Sign | «Correctness» | «Erroneous» | «Adequacy» |
|---|---|---|---|
| MG002 | 0.67 | 0.84 | 0.56 |
| MG010 | 0.68 | 0.78 | 0.57 |
| MG016 | 0.79 | 0.57 | 0.74 |
| MG039 | 0.80 | 0.73 | 0.60 |

## 4 Conclusion

The resulting level of adequacy of features to the classification result will allow:

- first, interpret the results of the classifier;
- secondly, to improve the classifier by changing its mathematical model in the direction that will provide a more significant role for those features, the level of adequacy of which is marked as high.

## References

1. Kazakov, I.D., Shcherbakova, N.L., Brebels, A., Shcherbakov, M.V.: Accelerometer data based cyber-physical system for training intensity estimation. In: Kravets, et al. A.G. (eds.) Cyber-Physical Systems: Advances in Design & Modelling. Studies in Systems, Decision and Control, 2020, vol. 259, pp. 325–335. Springer Nature Switzerland AG (2020)
2. Demin, S., Panischev, O., Yunusov, V., Timashev, S.: The Application of Statistical Methods for the Analysis of Multi-Parameter Data of Complex Composite Objects in the Field of Cyber-Physical Systems. Studies in Systems, Decision and Control. 2020, vol. 350. pp. 27–38 (2020)
3. Vesnin, S.G., Sedankin, M.K., Gudkov, A.G., Leushin, V.Y., Sidorov, I.A., Porokhov, I.O.: A printed an-tenna with an infrared temperature sensor for a medical multichannel microwave radiometer. Biomed. Eng. **54**(4), 235–239 (2020)
4. Vesnin, S., Turnbull, A.K., Dixon, J.M., Goryanin, I.: Modern microwave thermometry for breast cancer. J. Mol. Imag. Dymanics **7**(2) (2017)
5. Goryanin, I., Karbainov, S., Shevelev, O., Tarakanov, A., Redpath, K., Vesnin, S., Ivanov, Y.: Passive microwave radiometry in biomedical studies (Review). Drug Discov. Today **4**(25), 757–763 (2020)
6. Levshinskii, V.V.: Mathematical models for analyzing and interpreting microwave radiometry data in medical diagnosis. J. Comput. Eng. Math. **8**(1), 3–14 (2021)
7. Polyakov, M.V., Popov, I.E., Losev, A.G., Khoperskov, A.V.: Application of computer simulation results and machine learning in the analysis of microwave radiothermometry data. Math. Phys. Comput. Model. **24**(2), 27–37 (2021)
8. Vavourakis, V., Eiben, B., Hipwell, J.H., Williams, N.R., Keshtgar, M., Hawkes, D.J.: Multi-scale mechano-biological finite element modelling of oncoplastic breast surgery—numerical study towards surgical planning and cosmetic outcome prediction. PLoS ONE **11**(7), art. no. e0159766 (2016)
9. Polyakov, M., Levshinskii, V., Khoperskov, A.: Modeling of brightness temperature in biological tissue. J. Phys: Conf. Ser. **1368**(4), 042057 (2019)
10. Zenovich, A.V., Baturin, N.A., Medvedev, D.A., Petrenko, A.Y.: Algorithms for the formation of two-dimensional characteristic and informative signs of diagnosis of diseases of the mammary glands by the methods of combined radiothermometry. Math. Phys. Comput. Simul. **4**(21), 44–56 (2018)
11. Bardati, F., Iudicello, S.: Modelingthe visibility of breast malignancy by a microwave radiometer. IEEE Trans Biomed Eng **55**(1), 214–221 (2008)
12. Araujo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., Polonia, A., Campilho, A.: Classification of breast cancer histology images using convolutional neural networks. PLoS ONE **12**(6), e0177544 (2016)
13. Yassin, N.I.R., Omran, S., El Houby, E.M.F., Allam, H.: Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: a systematic review (Review). Comput. Methods Programs Biomed. **156**(2018), 25–45 (2018)

14. Mazepa, E., Dubovskaya, V.: Neural Network Modeling in the Construction of Multiple Nonlinear Regression of RTM-Diagnostic Data. 2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), 2020, pp. 439–444 (2020)
15. Lashkari, A., Pak, F., Firouzmand, M.: Full intelligent cancer classification of thermal breast images to assist physician in clinical diagnostic applications. J. Med. Signals Sens. **1**(6), 12–24 (2016)
16. Germashev, I., Dubovskaya, V., Losev, A.: Substantiation of the Result of the Work of the Cyber-Physical System for Breast Cancer Diagnostics. Cyber-Physical Systems: Modelling and Intelligent Control. Studies in Systems, Decision and Control, 2021, vol. 416, pp. 3–14 (2021)
17. Terano, T., Asai, K., Sugeno, M. (eds.): Applied Fuzzy Systems. Omsya, Tokyo (1989)
18. Wei, X.-J., Zhang, D.-Q., Huang, S.-J.: A variable selection method for a hierarchical interval type-2 TSK fuzzy inference system. Fuzzy Sets Syst. **438**, 46–61 (2022)
19. Wyrwoł, B.: Implementation of the FATI hierarchical fuzzy inference system using the immutability decomposition method. Fuzzy Sets Syst. **381**, 105–123 (2020)
20. Germashev, I., Dubovskaya, V., Losev, A., Popov, I.: Fuzzy Inference of the effectiveness factors of the computational model for the diagnosis of the breast cancer. In: 2021 3rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), 2021, pp. 528–533 (2021)

# Development of a Personalized Diet Using Structural Optimization

**Marina A. Nikitina**

**Abstract** Development of a personalized human diet with due account for a variety of different factors is associated with system analysis and formalization of accumulated data and knowledge, as well as digital intelligent technologies for their processing and making optimal decisions. The methodology of optimization and formation of personalized diets based on structural-parametric modeling is presented. The proposed approach allows one to solve the following tasks: (1) to analyze the daily diet or individual meals (breakfast, lunch, afternoon snack, dinner, additional meals (snacks)) with a known quantitative set of finished products in terms of energy value and chemical composition in order to reveal dietary disorders; (2) to calculate for a specified list of finished products their optimal quantity, as close as possible in all respects to the reference diet recommended depending on mental and physical activity; (3) to optimize the diet depending on the task at hand by selecting a group of finished products from a complete or selected list of archival data, equally taking into account all the necessary parameters of each product quality; (4) to adjust the diet taking into account dietary deviations in certain parameters of chemical composition and energy value by additionally introduced products of increased biological value for special purposes, multivitamin and multivitamin-mineral supplements, as well as natural bioactive substances.

**Keywords** Personalized diet · Digital nutrition · Parametric description · Structural optimization

## 1 Introduction

Diet is a major factor in human health, and choosing the right set of food products can be a great improvement for health [1]. Adherence to a healthy lifestyle has become a very important aspect in people's lives. The latter requires maintaining a healthy nutrition, taking into account the nature and amount of food products consumed,

M. A. Nikitina (✉)
V.M. Gorbatov Federal Research Center for Food Systems of RAS, Moscow, Russia
e-mail: nikitinama@yandex.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_4

which allows one to regulate the intake of calories and nutrients. The studies [2, 3] show that poor nutrition increases the risk of non-communicable diseases such as obesity, diabetes, etc. The approaches to automate the meal planning process are primarily associated with mathematical methods: linear optimization, metaheuristics, case-based reasoning, fuzzy logic.

A large number of research studies on nutrition planning are associated with the use of linear optimization. The first approach to solve this problem was made back in 1945 by the American economist Stigler [4]. A mathematical model was proposed and a trial and error approach was applied. As part of this approach, non-structured food products has been listed that meet standard nutritional requirements while optimizing food products prices.

Over the past two decades, the "Stigler's model" has been adopted and extended using different variations of the linear programming paradigm [5], integer programming [6], mixed integer-linear programming [7], and target programming [8]. In solving the problems of meal planning, the following studies were carried out [9–11]. The authors in [11] use an exchange system for menu planning. A food substitution option is offered from the same food group (for example, replacing 1/2 cup of pasta with 1/3 cup of white rice, since both foods are in the starchy category). However, the work does not consider the compatibility of products from different categories, as well as the sizes of portions.

In [8], the authors pay attention to human food preferences when considering nutritional needs, but ignore the compatibility between food products. This approach does not attach foods into a meal plan (breakfast, lunch, dinner), but rather presents a set of recommended foods as an unstructured list. The problem in applying linear optimization is primarily due to the fact that because of the large number of restrictions, the resolvability of the problem decreases [7]. Therefore, it is not possible to consider many factors such as nutritional requirements, food preferences, diversity and compatibility.

To create personalized nutrition plans, the authors of [12–16] used a metaheuristic approach with genetic algorithms. A detailed review of these studies is presented in the article [17]. Bianchi et al. [18] and El-Ghazi [19] notes that the metaheuristic approach does not guarantee a global optimal solution, but rather provides a "good" solution to the problem.

The papers [20–23] used the study case reasoning techniques. Scientists and researchers try to recommend nutrition plans by identifying the best nutrition plan from a set of existing ones and then tailoring them to meet the needs of the target patient by revising the existing menu. A detailed review of these studies is presented in the article [17].

Fuzzy logic when compiling a menu with due account for the patient's profile is presented in [24–28]. In [26], the food intake consumed by the patient is used as the initial data. The result is a conclusion about whether the food is healthy and with what probability. In [24, 25], the authors present a solution aimed at patients with diabetes mellitus. As input data, one must specify the products consumed for breakfast and lunch. As a result, the patient receives the remaining calorie intake and the necessary portions of food for dinner.

Using fuzzy ontologies of food and personal profiles in [27], a solution is presented that offers several options for dishes from each category of products that meet the requirements for dinner portions. However, it is not necessary to cook a full dinner, as the solution does not combine products to form a complete dinner, but rather offers product options.

There are various mobile and web applications related to nutrition and health. Typically, these are calorie tracking tools such as MyFitnessPal [29], MyPlate [30], MyNetDiary [31] and others. They help patients monitor their daily calorie and nutrient intake. These applications take the food consumed by the patient as input data, and calculate the amount of calories and food nutrients contained in the food products consumed.

Meal planning tools such as EatThisMuch [32], Fitness Meal Planner [33], Make-MyPlate [34], Yum-Me [35] help one create daily meal plans based on the required calorific value provided by the patient. One such tool is MakeMyPlate [34], a mobile app that provides patients with pre-optimally balanced, nutritionally-recommended daily diets that meet user-specified calorie levels. Each meal plan allows you to find alternatives while keeping your plate balanced. The patient takes an alternative from the available database without checking whether the replacing product/meal is equivalent in calories to the original one (which may lead to an excess or decrease in the recommended caloric content and the amount of food nutrients). Besides, the MakeMyPlate mobile app also does not take into account the patient's food preferences.

Another solution, EatThisMuch [32], takes as input data the patient's basic health information (gender, age, height, weight, and activity level) in addition to body fat percentage. Along with this, textual information is entered: (1) "what do you want to do with your weight"—to maintain, reduce, gain weight and gain muscle mass); (2) the preferred type of diet—Mediterranean, vegetarian, etc.; (3) do you have any food preferences and how often do you want to "see your preference" in the weekly diet. As a result, the app gives the user daily food rations as the conclusion. While this is a powerful mobile app, it has a few limitations. First, it allows the user to create meal plans for the current day only. Meal planning for two days or more requires a premium subscription. Secondly, there is no consideration of the user's preferences in a particular dish in form of a gradient assessment, for example, from "undesirable" to "highly preferred", which would be more useful when composing adapted meal plans. Another online fitness app Fitness Meal Planner [33] repeats most of the app features and limitations EatThisMuch [32].

However, in all studies and mobile apps, as a rule, the criterion of energy value is used. The difficulty of making optimal decisions is in a personalized approach, the selection of products and dishes in the diet, taking into account genetics and biomedical requirements, considering genetic characteristics and parameters of a particular person (age, gender, physiological state, labor intensity, regional, ethnic and ecological characteristics, as well as the state of digestibility of individual food products), structural relationships and restrictions at the component, elemental and monostructural levels.

In this regard, it is of interest to form a methodology for optimizing personalized diets, which allows one to solve the following tasks: (1) to analyze the daily diet or individual meals (breakfast, lunch, afternoon snack, dinner, additional meals (snacks)) with a known quantitative set of finished products in terms of energy value and chemical composition in order to reveal dietary disorders; (2) to calculate for a specified list of finished products their optimal quantity, as close as possible in all respects to the reference diet recommended depending on mental and physical activity; (3) to optimize the diet depending on the task at hand by selecting a group of finished products from a complete or selected list of archival data, equally taking into account all the necessary parameters of each product quality; (4) to adjust the diet taking into account dietary deviations in certain parameters of chemical composition and energy value by additionally introduced products of increased biological value for special purposes, multivitamin and multivitamin-mineral supplements, as well as natural bioactive substances.

## 2 Structural-Parametric Model of Personalized Human Nutrition

The software implementation of the personalized diet optimization is based on the knowledge base in the form of a structural-parametric description [36] of many genotype factors and variables of the physiological state of a living system, the relationships between them, the composition and properties of food raw materials and food products, reflecting the results of scientific research and the experience of nutritionists, physicians, technologists and other experts, taking into account the biomedical requirements and physiological needs of a living body. The genetic and physiological characteristics of a person, his/her carbohydrate, lipid and protein metabolism can practically be determined in the process of genetic and medical examination, in particular, basing on a biochemical blood test.

A structural-parametric model (SPM) of personalized adequate human nutrition is formed basing on the knowledge contained in the parametric description of a person and his/her preferences in the diet. SPM is a cellular matrix [37, 38] (see Fig. 1) and reflects the whole variety of existing known and unknown relationships between the human genotype, variables of the human physiological state and parameters of the diet and dietary regime.

Table 1 presents a part of the generalized data and knowledge about common genotypes associated with the manifestation of disease states [39].

To compose an adequate balanced diet with due account for the “health passport” of a person, the risk of diseases and the gastrointestinal tract status, a dialog algorithm is used (see Fig. 2). According to descriptions of medical and biological state of a person, as well as biomarkers, a parametric model of his/her adequate nutrition is formed as specific parameters, norms and ratios of the required nutrients and components per day [40, 41].

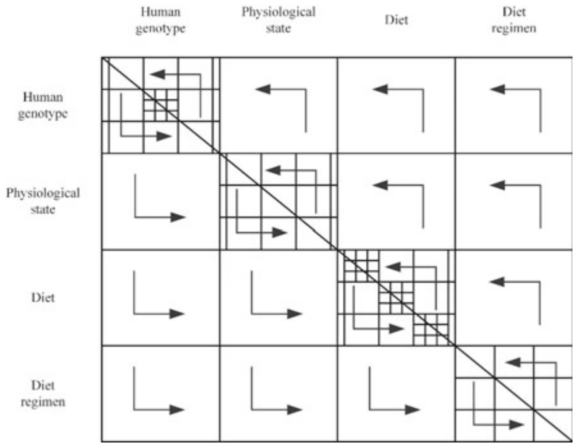



**Fig. 1** Structural matrix model of personalized human nutrition

**Table 1** Genes, genotypes, diseases

| Gene/SNP | Genotypes | Pathology | Diet regimen/ preferred type of diet |
|---|---|---|---|
| Features of fat and carbohydrate metabolism | | | |
| PPARG rs180128 | (Pro/Pro) | Obesity<br>Diabetes mellitus | Diet low in fat |
| ADRB2 rs1042714 | (Gln/Glu or Glu/Glu) | Obesity<br>Diabetes mellitus | Diet low in carbohydrates |
| ADRB2 rs1042713 | (Arg/Gly or Gly/Gly) | Obesity | Diet low in carbohydrates |
| SR (HTR2A) rs6313 | (Arg/Gly or Gly/Gly) | Increased tendency to impulsive eating, increased appetite, and a tendency to "comfort eating"<br>Increase in body mass index | |

In constructing a parametric model of a person, it is necessary to consider the following groups of properties:

1. Personal data (age, gender, physical activity, region of residence, type of activity, psycho-emotional state, etc.);
2. Critical properties (body mass index, blood pressure, glucose level, atherogenic index, etc.);
3. Secondary properties (carbohydrate metabolism, protein metabolism, lipid metabolism, etc.).

The parametric human model template is presented in Table 2.

Structural optimization of the diet is carried out using the quadratic criterion of the minimum deviation from the reference structure of nutrient indicators of nutritional and energy value for a specific determined group of people. As of today, reference

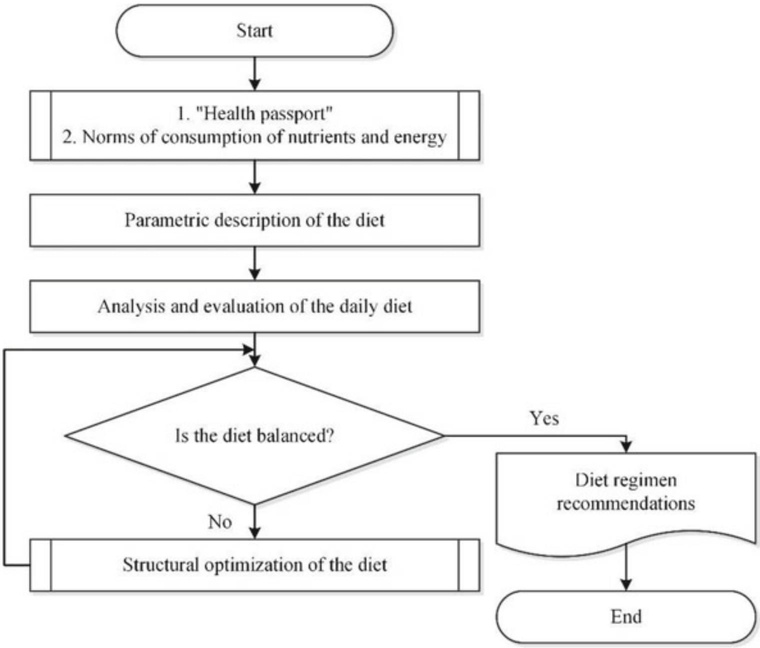



**Fig. 2** Block diagram of the dialogue algorithm for the analysis, evaluation and optimization of the diet

**Table 2** Parametric model of a person

| Group of properties | Parameter name | | |
|---|---|---|---|
| Personal data | Age | Region of residence (geographical, climatic, socio-cultural conditions) | |
| | Gender | Type of activity | Psycho-emotional state |
| Critical properties | Body weight (anthropometric data) | Man's height | Physical development |
| Secondary properties | Carbohydrate metabolism | Lipid metabolism | Mineral metabolism |
| | … | … | … |
| | Protein metabolism | Amino acid metabolism | Fatty acid metabolism |

indicators exist in an average form for various determined groups of people. No personalized recommendations are available.

The algorithm operation is to sequentially determine the imbalance in the $k$-th element with restrictions on the total volume of the daily ration, on the permissible limits for changing the mass fraction (volume) of the $j$-th product in the diet and its minimization (in case of deficiency) by increasing it to the upper limit of the mass share of one product $y_l$ with the maximum specific content $b_{kl}^{\max}$ of the deficient element and a corresponding decrease to the minimum mass fraction of another product $y_r$ with the minimum specific content $b_{kr}^{\min}$ of the $k$-th element. In the case of excess, the structural shift follows in the opposite direction. As a result of the mass fractions redistribution of the selected pair of products to $y_l^{\max}$ and $y_r^{\min}$, we obtain a new improved value of the optimization criterion. The value of the mass fractions redistribution of the components is selected based on the restrictions on the specified limits $y_l^{\max}$ and $y_r^{\min}$, as well as restrictions on the total volume of the diet, as

$$\delta = \min\{\Delta y_k, \Delta y_{kmax}, \Delta y_{kmin}\}$$

where $\Delta y_{kmax} = y_l^{max} - y_l$—an acceptable increase in the mass fraction of the $l$-th product;

$\Delta y_{kmin} = y_p - y_p^{min}$—an acceptable decrease in the mass fraction of the $p$-the product.

New values of mass fractions are determined as $y_l' = y_l + \delta$; $y_p' = y_p - \delta$ the procedure is repeated with finding the next pair of products to redistribute their mass fractions until finding a local minimum for the $k$–th element. The procedure continues until all the possibilities of redistributing the mass fractions of the components are exhausted, resulting in an alternative diet option with a minimum deviation of the $k$–th element from the given standard value in the structure of the adequate nutrition assortment.

In creating a diet for students with the gastrointestinal tract problems, it is necessary to consider scientifically based principles for prevention and specified by biomedical requirements. With this knowledge, clustering was carried out in "Food Products" database [42] and a list of products and dishes by categories (salads, first, second courses, side dishes, etc.) was obtained, which was used in the preparation of the diet.

As a result of structural optimization, there were obtained dietary options for the prevention of gastrointestinal tract diseases. Figure 3 shows one of them.

## 3 Conclusion

Composing a personalized human diet is directly related to digital technologies. Only they make it possible to consider all the necessary variety of factors: individual physiological characteristics (blood group, blood glucose level, body mass index, microbiota, acidity and gastrointestinal fluids, etc.), personal data (age, gender, height, weight, etc.), level and nature of food nutrients transformation under pepsin and trypsin action in the gastrointestinal tract and their digestibility.

| Breakfast | Weight, g. | Protein, g. | Fat, g. | Carbohydrates, g | Energy value., kCal | Vitamins | | | | | Mineral substances | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | C mg. | B2 mg. | B1 mg. | PP mg. | E mg. | Ca mg. | Fe mg. | P mg. | Mg mg. |
| 1. "Hercules" cereal porridge with butter | 125 | 3.00 | 5.00 | 18.50 | 131.25 | 0.00 | 0.03 | 0.09 | 0.25 | 1.38 | 23.75 | 1.00 | 87.50 | 36.25 |
| with butter | 5 | 0.03 | 4.13 | 0.47 | 37.40 | 0.00 | 0.01 | 0.00 | 0.01 | 0.01 | 1.10 | 0.01 | 0.95 | 0.15 |
| 2. Bun "Studentcheskaya" fortified | 1/30 | 2.20 | 1.50 | 16.10 | 86.70 | 0.00 | 0.06 | 0.09 | 1.62 | 0.00 | 24.42 | 0.99 | 0.00 | 0.00 |
| 3. "Rodnichok" processed cheese | 1/20 | 4.60 | 3.80 | 0.00 | 52.60 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 152.00 | 0.16 | 90.00 | 8.00 |
| 4. Cocoa with milk | 200 | 3.92 | 0.65 | 68.74 | 129.22 | 0.52 | 0.17 | 0.04 | 0.13 | 0.07 | 140.00 | 0.61 | 117.29 | 29.05 |
| **Total:** | | **13.75** | **15.08** | **103.81** | **437.17** | **0.52** | **0.27** | **0.22** | **2.01** | **1.46** | **341.27** | **2.77** | **295.74** | **73.45** |

| Dinner | Weight, g. | Protein, g. | Fat, g. | Carbohydrates, g | Energy value., kCal | Vitamins | | | | | Mineral substances | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | C mg. | B2 mg. | B1 mg. | PP mg. | E mg. | Ca mg. | Fe mg. | P mg. | Mg mg. |
| 1. Meat salad | 100 | 5.30 | 15.80 | 7.90 | 219.00 | 7.70 | 0.10 | 0.07 | 0.00 | 7.10 | 28.00 | 1.10 | 78.00 | 21.00 |
| 2. Rice soup on vegetable broth with meatballs | 300 | 6.60 | 6.60 | 17.10 | 153.00 | 16.40 | 0.10 | 0.10 | 1.97 | 1.38 | 24.06 | 1.35 | 100.75 | 31.55 |
| 3. Fish baked with potatoes, in Russian style | 250 | 17.25 | 10.25 | 12.00 | 230.00 | 12.42 | 0.19 | 0.17 | 2.63 | 1.91 | 127.14 | 1.32 | 264.71 | 43.99 |
| 4. "Tonus" bakery product | 60 | 7.20 | 0.54 | 25.20 | 129.60 | 0.00 | 0.30 | 0.50 | 4.30 | 0.00 | 26.40 | 7.02 | 342.00 | 87.60 |
| 5. Cherry compote | 200 | 1.20 | 0.60 | 47.40 | 202.00 | 4.00 | 0.04 | 0.04 | 0.40 | 0.40 | 20.00 | 0.80 | 34.00 | 16.00 |
| **Total:** | | **37.55** | **33.79** | **109.60** | **933.60** | **40.52** | **0.73** | **0.88** | **9.30** | **10.79** | **225.60** | **11.59** | **819.46** | **200.14** |
| | | | | | | | | | | | | | | |
| **IN ALL:** | | **51.30** | **48.87** | **213.41** | **1370.77** | **41.04** | **1.00** | **1.10** | **11.31** | **12.25** | **566.87** | **14.36** | **1115.20** | **273.59** |

| Protein | Fat | Carbohydrates |
|---|---|---|
| 1.05 | 1.00 | 4.37 |

**Fig. 3** Diet to prevent gastrointestinal tract diseases, recommended for a school or student canteen

**Acknowledgements** This chapter is prepared as part of scientific research theme No. 0585-2019-0008 under the state assignment of the federal state budgetary scientific institution 'V. M. Gorbatov's Federal Research Centre for Food Systems' of RAS.

## References

1. Lichtenstein, H., Appel, L.J., Vadiveloo, M., Hu, F.B., Kris-Etherton, P.M., Rebholz, C.M., et al.: Dietary guidance to improve cardiovascular health: a scientific statement from the American heart association. Circulation **144**, e472–e477 (2021)
2. Ayoub, J., Samra, A., Hlais, A., Bassil, M., Obeid, O.: Effect of phosphorus supplementation on weight gain and waist circumference of overweight/obese adults: a randomized clinical trial. Nutr. Diabetes **5**(12), e189 (2015)
3. Mattar, L., Zeeni, N., Bassil, M.: Effect of movie violence on mood, stress, appetite perception and food preferences in a random population. Eur. J. Clin. Nutr. **69**(8), 972–973 (2015)
4. Stigler, G.: The cost of subsistence. Am. J. Agr. Econ. **27**(2), 303–314 (1945)
5. Lancaster, L.: The history of the application of mathematical programming to menu planning. Eur. J. Oper. Res. **57**(3), 339–347 (1992)
6. Lenstra, H.: Integer programming with a fixed number of variables. Math. Oper. Res. **8**(4), 538–548 (1983)
7. Schniederjans, M.: Goal programming: methodology and applications, 1st edn. Springer, New York (1995)
8. McCann-Rugg, M., White, G., Endres, J.: Using goal programming to improve the calculation of diabetic diets. Comput. Oper. Res. **10**(4), 365–373 (1983)
9. Bassi, L.: The diet problem revisited. Am. Econ. **20**(2), 35–39 (1976)
10. Foytik, J.: Devising and using a computerized diet: an exploratory study. J. Consum. Aff. **15**(1), 158–169 (1981)

11. Valdez-Pena, H., Martinez-Alfaro, H.: Menu planning using the exchange diet system. In: International Conference on Systems, Man and Cybernetics, SMC'03, vol. 3, pp. 3044–3049. IEEE (2003)
12. Ainsworth, B., Haskell, W., Herrmann, S., Meckes, B.D.J.N., Tudor-Locke, C., Greer, V.J.J., Whitt-Glover, M., Leon, A.: 2011 compendium of physical activities: a second update of codes and MET values. Med. Sci. Sports Exerc. **43**(8), 1575–1581 (2011)
13. Fister, D., Fister, R.S.I.: Generating eating plans for athletes using the particle swarm optimization. In: 17th International Symposium on Computational Intelligence and Informatics (CINTI'2016), pp. 193–198. IEEE (2016)
14. Noor, S., Mohd, A., Ruhana, K., Mahamud, K., Arbin, N.: Self-adaptive hybrid genetic algorithm (SHGA). Far East J. Math. Sci. **103**, 171–190 (2018)
15. Seljak, B.: Computer-based dietary menu planning. J. Food Compos. Anal. **22**(5), 414–420 (2009)
16. Türkmenoglu, C., Etaner-Uyar, A.S., Kiraz, B.: Recommending healthy meal plans by optimising nature-inspired many- objective diet problem. Health Inform. J. **2127**(1), 146045822097671 (2021)
17. Salloum, G., Tekli, J.: Automated and personalized meal plan generation and relevance scoring using a multi-factor adaptation of the transportation problem. Soft. Comput. **26**, 2561–2585 (2022)
18. Bianchi, L., Dorigo, M., Gambardella, L., Gutjahr, W.: A survey on metaheuristics for stochastic combinatorial optimization. Nat. Comput. **8**(2), 239–287 (2009)
19. El-Ghazi, T.: Metaheuristics: From Design to Implementation. Wiley, New Jersey (2009)
20. Husain, W., Wei, L., Cheng, Z.N.S.: Application of data mining techniques in a personalized diet recommendation system for cancer patients. In: 2011 IEEE Colloquium on Humanities, Science and Engineering, pp. 239–244. IEEE (2011)
21. Khan, A.S., Hoffmann, A.: An advanced artificial intelligence tool for menu design. Nutr. Health **17**(1), 43–53 (2003)
22. Khan, A.S., Hoffmann, A.: Building a case-based diet recommendation system without a knowledge engineer. Artif. Intell. Med. **27**(2), 155–179 (2003)
23. Petot, G.J., Marling, C., Sterling, L.: An artificial intelligence system for computer-assisted menu planning. J. Am. Diet. Assoc. **98**(9), 1009–1014 (1998)
24. Lee, C., Wang, M., Hsu, C., Hagras, H.: A novel type-2 fuzzy ontology and its application to diet assessment. In: International Joint Conference on Web Intelligence and Intelligent Agent Technology, vol. 3, pp. 417–420. IEEE/WIC/ACM (2009)
25. Lee, C.S., Wang, M.H., Acampora, G., Hsu, C.Y., Hagras, H.: Diet assessment based on type-2 fuzzy ontology and fuzzy markup language. Int. J. Intell. Syst. **25**(12), 1187–1216 (2010)
26. Wang, M.H.: Intelligent ontological multi-agent for healthy diet planning. In: International Conference on Fuzzy Systems, pp. 751–756. IEEE (2009)
27. Lee, C., Wang, M., Habras, H., Chen, Z., Lan, S., Hsu, C., Kuo, S., Kuo, H., Cheng, H.: A novel genetic fuzzy markup language and its application to healthy diet assessment. Int. J. Uncertainty Fuzziness Knowl.-Based Syst. **20**(supp. 02), 247–278 (2012)
28. Lee, C., Lan, S.: Adaptive personalized diet linguisticrecommendation mechanism based on type-2 fuzzy sets and genetic fuzzy markup language. IEEE Trans. Fuzzy Syst. **23**(5), 1777–1802 (2015)
29. Evans, D.: MyFitnessPal. Br. J. Sports Med. **51**(14), 1101–1102 (2017)
30. Livestrong Foundation. MyPlateCalorie Counter. https://www.livestrong.com/myplate/. Accessed 1 May 2022
31. MyNetDiary. https://www.mynetdiary.com/. Accessed 1 May 2022
32. EatThisMuch Inc.: Eat this much. https://www.eatthismuch.com/. Accessed 1 May 2022
33. Fitness Meal Planner. http://www.fitnessmealplanner.com/. Accessed 1 May 2022
34. MakeMyPlate Inc.: Make my plate. http://www.makemyplate.co/. Accessed 1 May 2022
35. Yang, L., Hsieh, C.K., Yang, H., Dell, N., Belongie, S., Cole, C., Estrin, D.: Yum-Me: a personalized nutrient-based meal recommender system. ACM Trans. Inf. Syst. **36**(1), 1–31 (2018)

36. Ivashkin, Yu.A., Nikitina, M.A.: The concept of biological compatibility in optimization of a human diet. Sci. Intensive Technol. **19**(3), 33–44 (2018). (In Russian)
37. Nikitina, M.: Structural-parametric modeling in human healthy nutrition system. In: 6th International Conference Information Technology and Nanotechnology. Session Data Science, pp. 219–224. CEUR Workshop Proceedings (2020)
38. Ivashkin, Yu.A.: Structural-parametric modeling and identification of anomalous situations in complex technological systems. Manag. Prob. **3**, 39–43 (2003). (In Russian)
39. Development of ways to prevent diseases and improve health status by adjusting diets. In: Modern Technologies of Functional Food Products, pp. 422–425. DeLiplus, Moscow (2018). (In Russian)
40. Nikitina, M.A.: Digital technology in the development of healthy diet decision support system. In: Kravets, A., Bolshakov, A., Shcherbakov, M. (eds.) Society 5.0: Human-Centered Society Challenges and Solutions, vol. 416, pp. 65–74. Springer, Heidelberg (2022)
41. Nikitina, M.A.: Personalization in the structural optimization of the individual human nutrition diet. Math. Methods Technol. Technics **1**, 85–88 (2022). (In Russian)
42. Nikitina, M.A., Lisitsyn, A.B., Zakharov, A.N., Sus', E.B., Pilyugina, S.A., Dydykin, A.S., Ustinova, A.V.: Food products. Database Registration Certificate RU 2015620557 (2015). (In Russian)

# Development of an Algorithm for Preparing Semi-finished Products for Packaging

**Mikhail V. Tarachkov, Oleg V. Tolstel, and Alexandr L. Kalabin**

**Abstract** The algorithm of semi-finished products' preparation for packaging is presented and its efficiency is evaluated. The semi-finished products are chilled nuggets, which move along the high-speed conveyor line section in open cardboard boxes. The preparation consists of aligning the polyethylene packages with the nuggets so that they do not protrude beyond the walls of the box. This must be done to ensure that the box of nuggets seals correctly at the final packing stage. Leveling of semi-finished products is performed by a robot-manipulator DR-1, manufactured by Intelligent Robotics LLC, which is suspended above the conveyor belt on a special frame. Also included in the automation system is a detector of boxes with chilled semi-finished products, and a speed sensor for the conveyor belt. The estimation of errors during data transfer between sensors and actuators is given. Application of the proposed algorithm allows to perform of equalization of boxes with nuggets in automatic mode, thereby increasing productivity and reducing the number of rejects.

**Keywords** Food production · Robotic arm · Packaging of semi-finished products · Automated control system

## 1 Introduction

In the food processing company on the packaging line of finished products, there is a task of equalizing semi-finished products that are in the opened box. The necessity of the process is due to the fact that if the semi-finished products protrude outside the box, the packaging machine will not work correctly.

At the moment, the balancing is done by people. It is proposed to automate the alignment process. The publications [1, 2] describe the process of development of

M. V. Tarachkov (✉) · O. V. Tolstel
Immanuel Kant Baltic Federal University, Kaliningrad 236016, Russia
e-mail: mishklgpmi@mail.ru

A. L. Kalabin
Tver State Technical University, Tver 170026, Russia



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_5

robot-manipulator DR-1, which has the necessary characteristics for the implementation of equalization. It uses forward and inverse kinematics [3, 4] for positioning the end-effector. Delta robots with the same construction are used to sort waste [5].

The aim is to develop an algorithm that, based on the signal from the sensor of the passage of the box on the conveyor belt, gives control commands to the robot manipulator. The same task was solved in [6, 7].

The tasks defined are:

1. To assemble an installation for the equalization of semi-finished products in laboratory conditions.
2. Create a mathematical description of the equalization process.
3. Develop an algorithm for the equalization process.
4. Calculate the errors in the execution of the algorithm.
5. Conduct tests of the resulting automation system.

## 2 Hardware Description

The nugget packaging production line consists of a conveyor belt and a robot arm DR-1 located above it. The robot is mounted on a frame. A working body is attached to the robot manipulator, which is a vibrating platform. The conveyor line has a maximum movement speed of 0.55 m/s and adjustable width, which allows moving boxes with nuggets one after another (Fig. 1).

The nugget box moves in the direction of the blue arrow (Fig. 1) and has the characteristics shown in Table 1.

A laser pointer and a photo-sensor were used as a detector, fixed at the beginning of the conveyor line, with a distance of about 0.7 m to the center of the robot. This is the best way because using RGB-D cameras [8] and neural networks to determine the

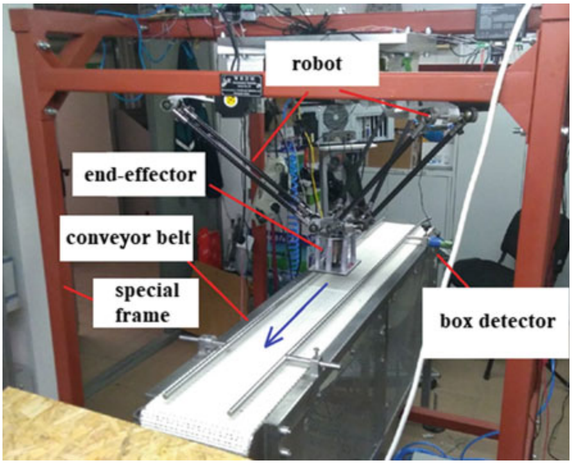



**Fig. 1** General view of the nugget mixing unit. A vibrating platform is used as the end-effector

**Table 1** Features of the nugget box

| Feature | Value |
|---|---|
| Length, m | 0.18 |
| Width, m | 0.13 |
| Height, m | 0.04 |
| Amount of nuggets, items | 14 |
| Weight, kg | 0.3 |

position of the boxes [9, 10] takes too much time and is redundant in this situation. Used a more classic way, like here [11].

The photosensor outputs a discrete signal of 0 V if the box is not present, and 3.3 V if the box interrupts the beam. The sensor was connected to the discrete input of the onboard microcontroller of the DR-1 robot.

## 3 Mathematical Description of the Equalization Process

For simplification, the installation shown in Fig. 1, considers schematically from the side (Fig. 2) and from above (Fig. 3).

The coordinate system, the center of which is located at point $M_0$, the $Z$ axis is directed downwards, the $X$ axis makes an angle $a$ with the central line of the conveyor, is presented. The $Y$ axis complements the right triplet. The point $M_0$ is the geometric center of the robot.

The angle $a$ appears because the robot is attached to the frame in such a way that it cannot touch it when it moves. Therefore, the central axis of the conveyor is not coaxial with any of the axes of the robot's coordinate system. The angle $a$ is constant.

During the experiments on debugging of motion planning system of robot manipulator DR-1 it was found that the highest speed can be obtained by setting the motion path with a minimum number of points. This is due to the fact that the robot manipulator must stop at a given point. Therefore, the trajectory in the form of a triangle is used when performing the nuggets flattening operation (Fig. 2).

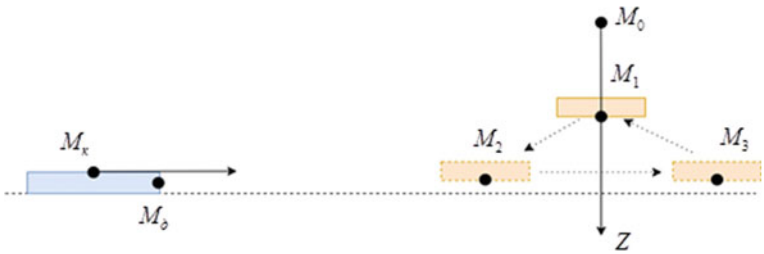



**Fig. 2** Diagram of the robot-manipulator working body movement when equalizing the box with nuggets, side view

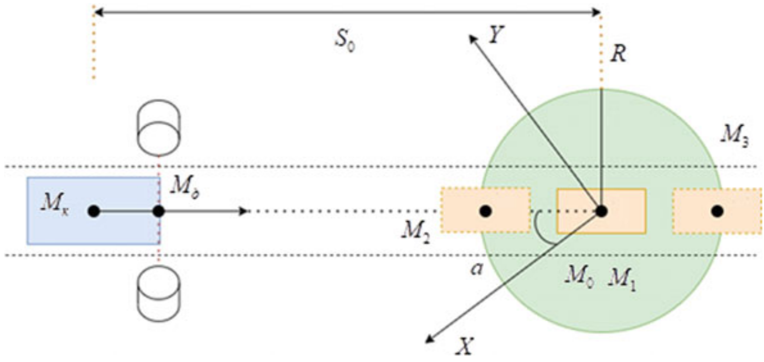



**Fig. 3** Schematic of the robotic arm movement when equalizing a box of nuggets, top view

In Fig. 2, the blue rectangle is the box with the nuggets, the point $M_\kappa$ is the center of its upper edge. The box appears in the line of sight of the robot manipulator at the moment of passing through the detector (point $M_\partial$). $M_0$ is the origin of the coordinate system of the robot manipulator. The point $M_1$ is the coordinate of the center of the bottom plate of the working body in the waiting position. The point $M_2$ is the coordinate when the bottom plate of the robot manipulator first contacts the box. At this point, the impact on the nuggets in the box by the robot manipulator itself and the pneumatic vibrator placed on the working organ is carried out. This impact continues up to the point $M_3$, the working body accompanies the box. At the point $M_3$, the tool stops impacting the box and starts moving to the waiting position—point $M_1$.

The point heights $M_1$, $M_2$, $M_3$ are chosen experimentally. The quality of the nuggets and the safety of the packs and half-finished products depend on them. During the tests, it was observed that the best way to level the nuggets is to set the height of points $M_2$ and $M_3$ at the height of the top edge of the box. This height minimizes the impact of the implementation on the carton. The height of the point $M_1$ is a few centimeters less ($Z$-axis pointing downwards) than $M_2$ and $M_3$, which is enough to prevent the box and the half-finished products in it from getting caught when the implement is moved.

At the moment the box passes through the detector (point $M_\partial$, time moment $t_0$), the coordinates of the center of the box $M_\kappa$ (1) and the projection of the distance $M_\kappa M_0$ on the $XY$ plane become known, denote $S_0$. It turns out that $S_0$ is the distance by which the box moves from the moment of its detection until it is directly below the middle of the robot.

$$\begin{cases} M_{\kappa_x} = M_{\partial_x} + \frac{l*\cos(a)}{2} \\ M_{\kappa_y} = M_{\partial_y} + \frac{l*\sin(a)}{2} \end{cases} \quad (1)$$

where $l$ is the length of the box, indexes $x$ and $y$ denote the projections on the corresponding axes.

The box moves to the projection of the point $M_0$ with constant velocity v, equal to the velocity of the conveyor.

Thus, we can write down the equation of motion of the center of the box in projections to the $X$ and $Y$ axes, assuming that $t$ is the current time:

$$\begin{cases} s_x(t) = s_0 * \cos(a) + \mathrm{v} * \cos(a) * (t - t_0) \\ s_y(t) = s_0 * \sin(a) + \mathrm{v} * \sin(a) * (t - t_0) \end{cases} \tag{2}$$

It should be explained that the velocity v will have a negative value in the coordinate system of the robot.

Let us define on the plane $XY$ a circle with a radius $R$ and center at $M_0$, which will be the effective working area of the robot manipulator. This means that the robot will not perform equilibrium until the box moves inside the circle:

$$\sqrt{s_x(t)^2 + s_y(t)^2 \leq R} \tag{3}$$

The radius $R$ was chosen experimentally. An experiment was conducted, in which $R$ was decreased from 0.2 to 0.03 m. The criterion was the qualitative alignment of the nuggets (the nugget pack does not protrude beyond the top edge of the box) and the fact that the robot has time to align all the boxes on the conveyor belt. A high $R$ value is excessive and doesn't equalize all the boxes. A small value of $R$, which degenerates into pushing from above on the box without accompanying it, is less effective from the point of view of leveling.

A delta value is introduced for tracking multiple boxes. This is the time it takes for one box to completely travel down the conveyor belt past the box detector. The delta is calculated by formula (4).

$$\mathrm{delta} = \frac{l}{\mathrm{v}} \tag{4}$$

where $l$ is the length of the box, v—conveyor velocity. The delta value allows excluding false triggers of the detector when one box moves and at the same time to signal the beginning of a new box, if they go one after another without separation.

Assign each box an ordinal number $i(i = 1\ldots)$, and for each $i$-th box, we will remember $t_{0_i}$—the time of passage of each box through the detector, from which we can calculate the position of the box at a time $t$.

## 4 Algorithm of Half-Finished Products Equalization

Based on the mathematical description of the semi-finished product equalization process, an algorithm has been developed, the description of which is given in natural language.

The algorithm is divided into 2 blocks, working in parallel.

The first block is responsible for processing signals from the box detector and recording time $t_{0_i}$ in a ring queue:

In case a signal from the box detector arrives, do the following. If the queue is empty or the difference between the current time t and the last element of the queue is greater than or equal to the delta, then add the current time t to the end of the queue.

An empty queue means that no box has passed through the detector yet. The time difference must be determined in order to separate 2 boxes that follow each other.

The second block is responsible for calculating the position of the box whose time $t_0$ is recorded at the beginning of the queue and sending control commands.

1. Set the signal "Operation allowed" to FALSE.
2. If the signal "Operation is permitted" is set to FALSE, go to step 3, otherwise go to step 5.
3. If the queue size is greater than 0, calculate the current position of the box according to formula (2), go to step 4, otherwise go to step 2.
4. If the current position of the box is within the working area of the robot (formula (3)), set the "Operation allowed" signal to TRUE and remove the first element from the queue, go to step 2.
5. Perform sequential movement from point $M_1$ to $M_2$, then from $M_2$ to $M_3$. Then from $M_3$ to $M_1$. Set the "Operation Allowed" signal to FALSE. Go to step 2.

The presented algorithm was implemented in the C + + 11 standard programming language and integrated into the existing software solution as a node in the Robot Operating System (ROS) software platform [12–14].

This node communicates with forward and inverse kinematics solver nodes via topics and [15] tells it works in real-time. It is suitable to use the Linux real-time kernel patch as done here [16] to increase performance. TCP/IP connection over Ethernet is used in robotics and is quite efficient [17]. The onboard microcontroller uses the operating system FreeRTOS, which is real-time [18].

## 5 Errors Estimation

It is especially important to estimate the errors when performing work on the equalization of boxes with nuggets because we consider the high-speed section of the production line, where errors can lead to damage to semi-finished products and packaging containers.

The permissible positioning error of the working body is $e = 0.02$ m. This is the difference between the length and width of the box and the similar parameters of the lower part of the working body.

At the speed of the conveyor in v = 0.55 m/c error in time can be no more than $t_{err} = 36$ ms. The calculation was carried out according to the formula (5).

**Table 2** Basic system delays

| Delay source | Value, ms |
|---|---|
| Interrogation of the detector by the onboard microcontroller | 0.333 |
| Data transfer between the onboard microcontroller and the robot control computer | Less 10 |
| Data transfer within the ROS software platform between the nodes of the system | Less 3 |
| Frequency of data checking by the developed algorithm | 10 |

$$t_{\text{err}} = \frac{e}{\text{v}} \tag{5}$$

Table 2 shows the main delays that can affect the positioning error.

Thus, in the worst case, the delay will be no more than 23 ms, which satisfies the conditions.

## 6 Tests

Three boxes of nuggets with nuggets protruding outside the package were used for the tests. The main condition for quality leveling was to level the nuggets in such a way that they do not protrude outside the package. Also, the integrity of the packaging and the semi-finished products was evaluated, here we should take into account that the equalization must be done exactly once in production conditions, and in the experimental conditions, it was done several times.

Boxes of nuggets were placed on the conveyor belt before the detector in various combinations:

1. One box.
2. Two boxes going across the gap.
3. Two boxes going side by side.
4. Three boxes going side by side.

The case with three boxes going in a row was considered most “thorough” as the most difficult to perform.

Twenty experiments were conducted. The boxes were consecutively fed on the conveyor so that there were no gaps between them (Fig. 4). After that, the boxes were influenced and the result was checked (Fig. 5). All boxes of nuggets were successfully aligned. Damage to the carton and some of the nuggets was observed. No damage to the polyethylene packaging was observed.

**Fig. 4** Boxes of nuggets before the test. The nuggets protrude over the edges of the box

**Fig. 5** The result of the algorithm. The nuggets do not protrude over the edges with the box

## 7 Conclusions

In the course of the study, all of the objectives have been achieved. The goal of the research has been achieved. The developed algorithm showed effectiveness in the equalization of boxes of nuggets, it can be used in industry after longer testing in the complex with the robot manipulator and the working body.

At the same time, the disadvantages of the robot manipulator design were identified:

1. Installation of a more rigid connection to the conveyor belt.
2. Reducing positioning error by transferring the algorithm from the control computer to the onboard microcontroller, which is a real-time system, for example [19].
3. Adding the ability to quickly calibrate the system in-house, for example [20].

Such improvements will allow more efficient and less costly use of the equalization algorithm and the entire automation system.

## References

1. Tarachkov, M.V., Kalenik, A.V.: Development of the DR-1 robot-manipulator, hybrid and synergetic intelligent systems. In: Kolesnikov, A.V. (ed.) Proceedings of the V All-Russian Pospelov Conference with International Participation, pp. 482–489 (2020)
2. Tarachkov, M.V., Lebedev, F.A.: Development of motion planning system for DR-1 robot-manipulator. Math. Methods Technol. Eng. **12**, 70–75 (2021)
3. Tyves, L.I.: Mechanisms of Robotics. The Concept of Interchanges in Kinematics, Dynamics and Motion Planning. Lenand, p. 28 (2018)
4. Zsombor-Murray, P.: Descriptive Geometric Kinematic Analysis of Clavel's «Delta» Robot, McGill University Department of Mechanical Engineering Centre for Intelligent Machines, p. 2 (2004)
5. Gemuev, ShSh., Vorotnikov, S.A.: Robotic complex for sorting solid household waste. Caspian J. Manag. High Technol. **2**(46), 194–207 (2019)
6. Koltygin, D.S., Sedelnikov, I.A.: Methodology for developing a robot control program for the Delta* robotic arm. Sci. Bull. NSTU **70**(1), 103–116 (2018)
7. Zhang, G., Li, Z., Ni, F., Liu, H.: A real-time robot control framework using ROS Control for 7-DoF light-weight robot. In: Proceedings of the 2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics Hong Kong, China, pp. 1754–1579 (2019)
8. Zhou, T., Fan, D.P., Cheng, M.M., et al.: RGB-D salient object detection: a survey. Comp. Vis. Media **7**, 37–69 (2021). https://doi.org/10.1007/s41095-020-0199-z
9. Arkin, E., Yadikar, N., Xu, X., et al.: A survey: object detection methods from CNN to transformer. Multimed. Tools Appl. **82**, 21353–21383 (2022). https://doi.org/10.1007/s11042-022-13801-3
10. Marullo, G., Tanzi, L., Piazzolla, P., et al.: 6D object position estimation from 2D images: a literature review. Multimed. Tools Appl. **82**, 24605–24643 (2022). https://doi.org/10.1007/s11042-022-14213-z
11. Jin, X., Li, M., Li, D., et al.: Development of automatic conveying system for vegetable seedlings. J. Wireless Commun. Netw. **2018**, 178 (2018). https://doi.org/10.1186/s13638-018-1200-8
12. Mahtani, A., Sanchez, L., Fernandez, E., Martinez, A.: Effective Robotics Programming with ROS, PacktPublishing Ltd., Birmingham B3 2PB, UK, p. 129 (2016)
13. Joseph, L. Mastering ROS for Robotics Programming, PacktPublishing Ltd., Birmingham B3 2PB, UK, p. 552 (2018)
14. Quigley, M., Conley, K., Gerkey, B.P., Faust, J., Foote, T., Leibs, J., Wheeler, R., Ng, A.Y.: Ros: an open-source robot operating system. In: ICRA Workshop on Open Source Software (2009)
15. Aarizou, M.L., Berrached, N.E.: ROS-based telerobotic application for transmitting high-bandwidth kinematic data over a limited network. Int. J. Control Autom. Syst. **17**, 445–453 (2019). https://doi.org/10.1007/s12555-018-0047-4
16. Adam, G.K., Petrellis, N., Doulos, L.T.: Performance assessment of Linux Kernels with PREEMPT_RT on ARM-based embedded devices. Electronics **10**, 1331 (2021). https://doi.org/10.3390/electronics10111331
17. Kim, D.W., Lee, H.D., de Silva, C.W., et al.: Service-provider intelligent humanoid robot using TCP/IP and CORBA. Int. J. Control Autom. Syst. **14**, 608–615 (2016). https://doi.org/10.1007/s12555-014-0441-5
18. Pinto, M.L., Wehrmeister, M.A., de Oliveira, A.S.: Real-time performance evaluation for robotics. J. Intell. Robot. Syst. **101**, 37 (2021). https://doi.org/10.1007/s10846-020-01301-1
19. Subbaraman, S., Patil, M.M., Nilkund, P.S.: Novel integrated development environment for implementing PLC on fPGA by converting ladder diagram to synthesizable VHDL code. In: 11th International Conference on Control Automation Robotics and Vision, pp. 1791–1795. IEEE (2010)

20. Chen, Z., Yang, X., Zhang, C., Jiang, S.: Extrinsic calibration of a laser range finder and a camera based on the automatic detection of line feature. In: 2016 9th International Congress on Image and Signal Processing, BioMedicalEngineering and Informatics (CISP-BMEI), pp. 448–453 (2016)

# Society 5.0: Human-Centric Cyber-Solutions

# Computer Training System for Planning Multi-Assortment Discrete-Continuous Productions

**Tamara Chistyakova, Christian Kohlert, Olga Shashikhina, Ivan Kornienko, and Alexander Plekhanov**

**Abstract** A specialized computer system for training production personnel of multi-assortment industrial enterprises in the skills of solving the problem of optimal production planning is presented. This task is characterized by the need to take into account a large range of products, and a variety of types and configurations of production equipment that can be reconfigured to different types of orders. To effectively solve the training problem, the functional structure of the simulator has been developed, which is a flexible computer system configured for a specific production. The structure includes ergonomic interfaces of the trainee and the instructor, information support: a library of mathematical models of target functions and a database of rules for their formation, a database of orders, equipment, and customers, as well as software implementations of various modern optimization methods. Based on the information and mathematical support, a scenario and a training protocol are developed in which various situations of production planning are modeled. The effectiveness of the computer simulator is confirmed by the results of testing at Russian and foreign industrial enterprises, as well as by its introduction into the educational process.

**Keywords** Computer simulators · Personnel training · Calendar planning · Production schedules · Discrete–continuous production

T. Chistyakova (✉) · O. Shashikhina · I. Kornienko · A. Plekhanov
Saint Petersburg State Institute of Technology (Technical University), Moskovsky Prospekt 26, Saint Petersburg 190013, Russia
e-mail: nov@technolog.edu.ru

O. Shashikhina
e-mail: shashikhina@technolog.edu.ru

C. Kohlert
Kohlert-Consulting, Rotsteinerstr. 14, 56414 Oberahr, Germany
e-mail: info@kohlert-consulting.com



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_6

# 1 Introduction

The efficiency of the functioning of modern multi-assortment enterprises directly depends on the effectiveness of the production planning process and the preparation of production schedules. The task of optimizing planning is relevant for various innovative civil and industrial enterprises producing multi-assortment products on high-tech equipment [1–3]. The task of effective training of production personnel in the planning process is especially urgent [4, 5].

Training in the process of finding the optimal production schedule is quite a difficult task due to the need to take into account many characteristics of the production process when solving: the type and configuration of production equipment, the parameters of multi-assortment products, as well as requirements for the quality and timing of orders. Due to the large dimension of the problem, in order to effectively find the optimal schedule, it is necessary to develop software implementations of modern optimization methods that can search for solutions in a large space of acceptable solutions in acceptable machine time [6].

Therefore, the digitalization of production, which is currently actively developing, requires the development of specialized simulators that allow staff to train scenarios and algorithms for drawing up optimal production schedules based on modern optimization methods, which take into account a variety of technological limitations [7, 8].

Therefore, it is advisable to develop a computer training system that is operatively configured for various characteristics of the production process of each enterprise, which has intelligent interfaces that provide variability in the use of the software package functionality for various planning situations and provides an effective solution to planning problems of different dimensions through the use of modern optimization methods.

The introduction of such a computer training system into the enterprise management system is required based on the high requirements for the qualification of production personnel and the importance of the planning result for the entire technological process [9, 10].

# 2 Information and Mathematical Description of the Production Planning Process

## *2.1 Formalized Description of the Object of Study*

The first stage of the implementation of the educational process is the development of a formalized description of the production planning of continuously discrete industrial productions as an object of study.

The input data vector is given $X = (O, E, Pd)$, consisting of the following parameters:

$\tilde{O} = \left\{ \underline{O_i} : \underline{O} = \left( N^O, K^O, P^O, Q^m, \tilde{R}, \tilde{G}, Cs \right), i = \underline{1, N} \right\}$—vector characterizing orders: $N^O$—order number; $K^O$—order code; $P^O$—type of products included in the order; $Q^m$—amount of material; $\tilde{R}$—vector characterizing the formulation of the material; $\tilde{G} = \{W^m, Th^m\}$—vector that characterizes the geometric characteristics of the material, such as $W^m$—width of the material, $Th^m$—material thickness; $Cs$—customer; $i$—order sequence number; $N$—number of orders.

$\tilde{E} = \left\{ \underline{E_j} : \underline{E} = \left( N^E, K^E, V^P, \underline{To^E}, \underline{C_E} \right), j = \underline{1, Me} \right\}$—vector characterizing production lines (aggregates, work centers): $N^E$—name of the equipment; $K^{Rc}$—equipment code; $V^P$—productivity; $\underline{To^E} = \left\{ \underline{To_r^E}, r = \underline{1, R} \right\}$—technological operations that can be performed on this equipment; $\underline{C_E}$—vector describing hardware settings; $j$—serial number of the production line (unit, work center); $Me$—number of production lines (units, work centers).

To formalize the parameters described above in an application to real industrial productions and for operational adjustment to a new type of enterprise, knowledge representation is used in the form of intelligence maps that reflect knowledge about production equipment and products involved in planning, their characteristics and hierarchical relationships between them. Intelligence cards are also used in the development of augmented databases as part of the information support of a computer system.

$\widetilde{Pd} = (\tau^o, Y^{cr})$—vector characterizing the planning parameters: $\tau^o = \left[\tau^b, \tau^e\right]$—planning period; $\tau^b$—planning start date; $\tau^e$—end date of planning; $Y^{cr}$—optimization criteria.

The vector of variable parameters is $Q = \left\{ Q : Q = (j, k, \tau_{oi}, \tau_i), i = \underline{1, N}, j = \underline{1, Me}, k = \underline{1, L_j}, L \in N \right\}$—vector describing the distribution of $i$ orders on $j$ lines (aggregates, work centers), $k = \underline{1, L_j}$—the sequence number of the execution of the *i-th* order on the *j–th* line (unit, work center) in the current schedule $Q$, $L_j$—the number of orders executed on the *j–th* line (unit, work center) in the schedule $Q$.

The result of planning (vector $Y$) will be such a vector of distribution of orders along production lines (aggregates, work centers) $Q^{opt} = \left\{ Q_i : Q = (j, k, \tau_{oi}, \tau_i)\ j = \underline{1, Me}, k = \underline{1, L_j}, L_j \subset N, i = \underline{1, N} \right\}$, for which the value of the objective function for each production line (unit, work center) and the entire production cycle will be optimal.

## 2.2 *Setting the Task of Optimizing Production Planning*

It is required to find such a value of the vector of variable parameters $Q^{opt} = \left\{ Q_i : Q = (j, k, \tau_{oi}, \tau_i)\ j = \underline{1, Me}, k = \underline{1, L_j}, L_j \subset N, i = \underline{1, N} \right\}$, that is, the placement and order of execution of $N$ orders on $Me$ production lines within the planning period $\left[\tau^b,\ \tau^e\right]$, which will provide the optimal value of the objective functions.

In the case of solving the problem for one production line (unit, work center), the total time for order fulfillment and equipment configuration is minimized:

$$F_j\left(Q^{opt}\right) = \min_{Q \in D \subset R^{|Q|}} \left( \sum_{k=1}^{L_j - 1} \omega^{Ch}_{O_{j,k},\, O_{j,k+1}} * \tau^{ch}\left(O_{j,k},\ O_{j,k+1}\right) + \sum_{k=1}^{L_j} \omega^O_k * \tau^O_{j,k} \right) \quad (1)$$

In order to ensure a minimum production cycle of manufacturing orders on several production lines (aggregates, work centers), the production time of orders on the production line (aggregate, work center) for which it is maximum is minimized:

$$F_C\left(Q^{opt}\right) = \min_{Q \in D \subset R^{|Q|}} \left(F_j(Q)\right) \quad (2)$$

where $\tau^{ch}\left(O_{j,k},\ O_{j,k+1}\right)$—the time of equipment setup from the execution of order $k$ to the next order under the number $k + 1$, $\tau^O_k$—execution time of the *k-th* order, $\omega_k$—the weight coefficient of the $k$-th order, $R^{|Q|}$—search space, $D$—the set of acceptable values of the vector $Q$.

The solution to the optimization problem is such a vector of distribution of orders along production lines (aggregates, work centers) within the planning interval $Q^{opt} = \left\{Q_i\ :\ Q = (j,\ k,\ \tau_{oi},\ \tau_i)\ j = \underline{1,\ Me},\ k = \underline{1,\ L_j},\ L_j \subset N,\ i = \underline{1,\ N}\right\}$, for which the value of the objective function for each production line (unit, work center) $F_j\left(Q^{opt}\right)$ and the total production cycle time $F_C\left(Q^{opt}\right)$ will be optimal, which will ensure uniform loading of production equipment.

Expert information received from the specialists of the enterprise is formalized in the process of filling in special conceptual tables and is the basis for forming a base of rules for reconfiguring equipment from manufacturing one type of product to another [11, 12].

The rules of equipment settings are presented in the form of a production model and can be used to correctly form the type of target functions. When calculating the value of the objective function, the total time of reconfigurations is based on implicative statements of the form "*IF* [parameter of the *i–1-th* order $P^O = \ldots$] *AND* [*i-th* order parameter $P^O = \ldots$] *AND* [line parameter of the *j-th* line $P^E = \ldots$], *THEN* reconfiguration time $\tau^{Ch}_m = \ldots$".

### 2.3 Methods of Solving the Problem of Optimal Production Planning

To effectively search for the optimal production plan in a computer system, various optimization methods are used for tasks of different dimensions, since with an increase in the number of orders, units of equipment and the number of its possible reconfigurations for the manufacture of multi-assortment products, the number of plan options (the range of acceptable values) increases exponentially.

To make a decision on choosing the optimal planning method, the total number of plan options for a given number of orders, equipment units and its configurations is calculated. Further, based on the number of options found, a decision is made on the applicability of the full search method, which makes it guaranteed to find the optimal solution, and if the permissible dimension of the problem is exceeded $\left(C_{\text{var}} > 10^6\right)$, the planning problem is solved by a more complex optimization method—a genetic algorithm.

The genetic evolutionary optimization algorithm in the application to the problem under consideration has a number of advantages that have determined its use and study. Due to the combination of the advantages of gradient and random methods (using selection, crossing and mutation operators), the search for the optimal value is directional, but occurs in a wide range of possible values [13–17].

The beginning of the search from different starting points (the formation of an initial population of solutions of varying magnitude), as well as the use of the mutation operator and the combination of suboptimal solutions from different populations make it possible to avoid falling into local extremes and ensure finding a solution close to the global extremum [18, 19].

The input parameters of the optimal planning problem are transformed into elements of a genetic algorithm. When comparing two schedules, the values of the fitness function are compared (depending on the optimization criterion) and the components of the objective function are calculated—the time of order production and the time of equipment reconfiguration. After the stop criterion occurs, the algorithm finishes its work and the user is provided with the received production plan with the best value of the fitness function in the last population.

## 3 Structure of the Computer System of Training in Production Planning

### 3.1 Functional Structure of the Computer System

Based on the formalized mathematical description of the planning process presented above, the functional structure of the simulator-training computer system was developed.

The structure includes four main interfaces: knowledge engineer, database administrator, instructor and trainee.

The database administrator interface allows you to keep up-to-date the information support of the software package.

Information support includes databases of equipment and products; a database of equipment reconfiguration rules; libraries of mathematical models of objective functions, criteria constraints and optimization methods, as well as a database of educational theoretical materials. Information support is configured for various objects of study and planning using a special knowledge engineer interface, which also allows it to be integrated into automated control systems for innovative continuous-discrete productions.

The instructor's interface allows you to form tasks, monitor the progress and results of their implementation, develop an educational scenario, conduct testing and form a protocol of learning outcomes.

The learner's interface includes modules that allow you to study expert knowledge about the planning object, study the procedure for importing data, form a planning situation (select orders and equipment for planning, as well as set deadlines for building a plan), pass testing and study reference materials.

The open architecture of the computer system allows the connection of new software modules, which makes it possible to expand the functionality of the simulator.

The development of the computer system was carried out using the object-oriented programming language C# in the Microsoft Visual Studio environment on the platform APS.NET 2.0. The database was developed using the SQLite relational database control system.

Figure 1 shows the functional structure of the simulator-training computer system.

### 3.2 *The Process of Learning Production Planning*

The process of training in production planning is carried out in accordance with the developed scenarios and methods.

The instructor's interface allows you to implement various training options: issuing recommendations, answering a student's question, conducting testing (according to the selected knowledge control model), exam mode.

The computer system provides the following formats of mastering the material for the student: studying theoretical material (viewing the information and reference system), passing testing, configuring the system for specific equipment, studying the process of building calendar schedules in the mode of independent (free) training. These procedures are the basis for the formation of professional competencies and skills of effective production scheduling.

Practical tasks to be solved in the computer system were formed on the basis of data provided by polymer materials manufacturing companies in Russia and Germany

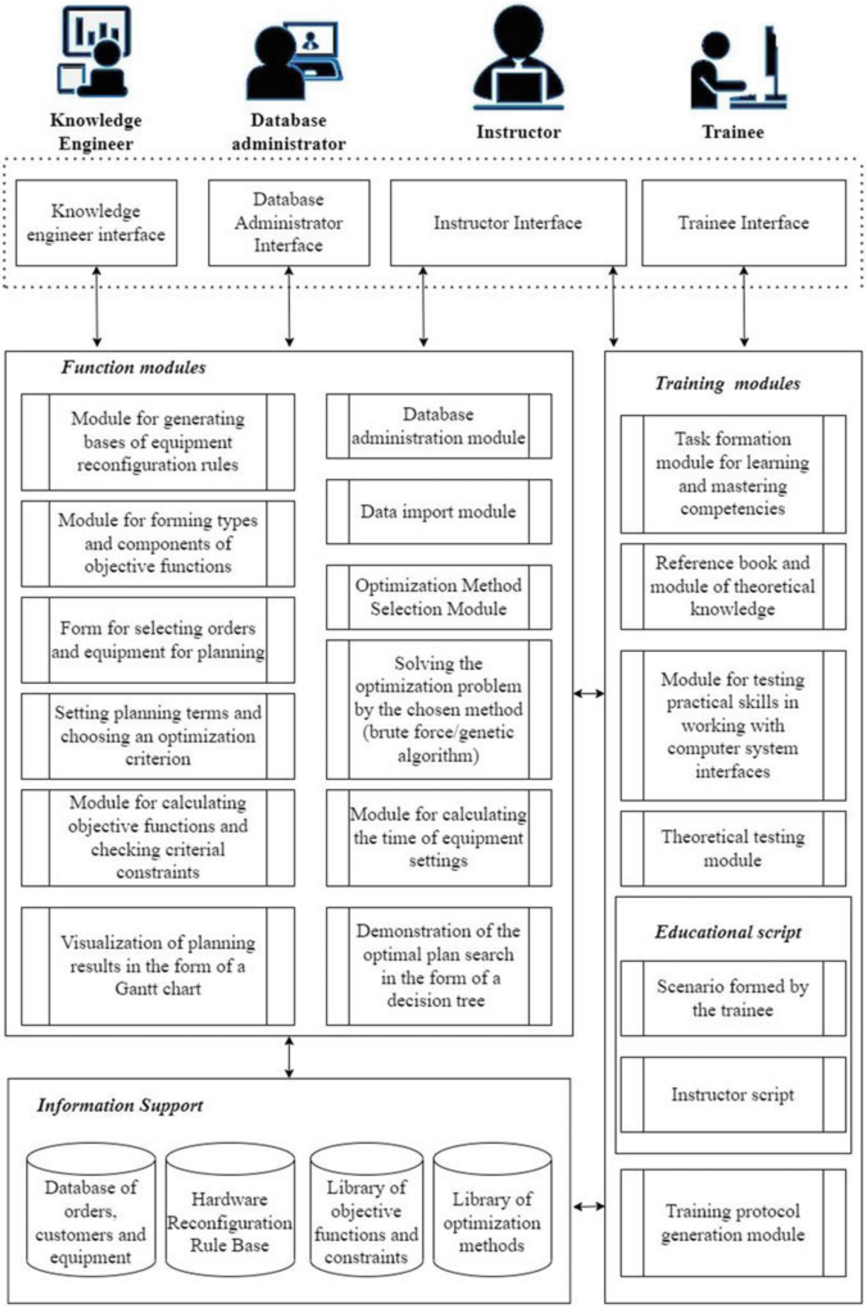



**Fig. 1** Functional structure of the training computer system

(Klöckner Pentaplast Rus and Maria Soell HTF), as well as by Russian metallurgical industries (Chelyabinsk Plant of Metal Structures) [1, 20].

Based on the results of the study, a training protocol is formed, which contains the data of the trainee, the date and time of training, the scenario and the name of the simulated situations. The protocol is at the disposal of the instructor and the teacher and is a document for assessing the qualifications of the student.

Consider the planning situation based on the data of the manufacturer of multilayer polymer films "Maria Soell HTF" (Germany): it is necessary to find the optimal order of manufacture of 58 different assortment orders, planning is carried out for 1.5 months, the number of different combinations of product parameters is more than 200. After solving the problem using the interfaces of the computer system, the trainee needs to compare the resulting plan with the plan formed in production.

In Fig. 2, you can see the visualization of the constructed optimal plan as a result of mastering the competence of production planning. On the diagram, you can see the order of production of orders, the start and end date of production of each order, the time of its manufacture and the time of reconfiguration from the previous order to the current one.

The effectiveness of the computer simulator is confirmed by the results of training students and production personnel according to Russian and foreign industrial enterprises, as well as the introduction into the educational process.

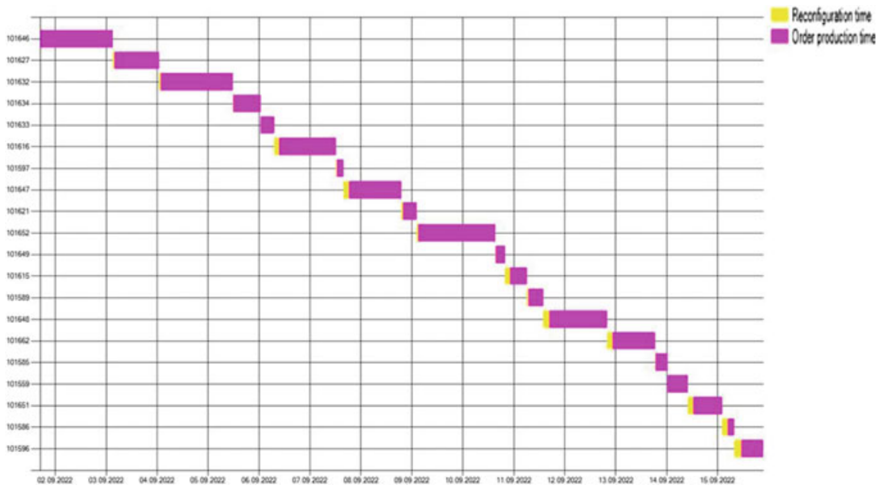



**Fig. 2** Gantt chart as a result of mastering the skill of building an optimal plan using a simulator-training system

## 4 Conclusion

A computer simulator and training system has been developed to train the production staff of multi-assortment industrial enterprises in the skills of solving the problem of optimal production planning. The introduction of the created computer system for the development of competence-based learning results allows to significantly improving the professional level of production personnel and students in the study of industrial optimal planning tasks.

The developed computer system has been successfully tested at Russian and foreign industrial enterprises, introduced into the educational process and can be used to improve the skills of management and production personnel of innovative industries.

The research was carried out at the expense of a grant from the Russian Science Foundation (project No. 21–79-30029).

## References

1. Shashikhina, O.E., Chistyakova, T.B., Kohlert Ch.: Software package and mathematical models for optimal planning of multi-assortment polymer films industrial production. In: 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency, SUMMA 2020, pp. 561–565 (2020)
2. Eremin, O., Stepanova, M.: A reinforcement learning approach for task assignment in IoT distributed platform. In: Studies in Systems, Decision and Control, vol. 350, pp. 380–389. Springer (2021)
3. Meshalkin, V.P., Stoyanova, O.V., Dli, M.: I: Project management in the nanotechnology industry: specifics and possibilities of taking them into account. Theor. Found. Chem. Eng. **46**(1), 50–54 (2012)
4. Amurrio, A., Azketa, E., Gutierrez, J.J., Aldea, M., Harbour, M.G.: Response-time analysis of multipath flows in hierarchically-scheduled time-partitioned distributed real-time systems. IEEE Access **8**, 196700–196711 (2020). https://doi.org/10.1109/ACCESS.2020.303346
5. Ershova, O., Chistyakova, T.: Methods and technologies of developing computer simulators for training resource-saving control of electrotechnological installations. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Society 5.0: Human-Centered Society Challenges and Solutions. Studies in Systems, Decision and Control, vol. 416, pp. 351–363. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-95112-2_29
6. Garey, M.R., Johnson, D.S., Sethi, R.: The complexity of flowshop and jobshop scheduling. Math. Oper. Res. **1**(2), 117–129 (1976)
7. Beasley, D., Martin, R.R., Bull, D.R.: An overview of genetic algorithms: Part 1, fundamentals. Univ. Comput. **15**(2), 58–69 (1993)
8. Oztemel, E., Gursev, S.: Literature review of Industry 4.0 and related technologies. J. Intell. Manuf. **31**, 127–182 (2020). https://doi.org/10.1007/s10845-018-1433-8
9. Hwang, J.S.: The fourth industrial revolution (industry 4.0): intelligent manufacturing. SMT Surf. Mt. Technol. Mag. **37**, 10–15 (2016)
10. Ziyatdinov, N.N.: Computer simulation and optimization in chemical technology. Theor. Found. Chem. Eng. **48**(5), 539–540 (2014)

11. Chistyakova, T.B., Novozhilova, I.V.: Methods and technologies of computer training complexes design for personnel in metallurgical production control. In: 2018 Third International Conference on Human Factors in Complex Technical Systems and Environments (ERGO)s and Environments (ERGO), pp. 132–136. IEEE (2018)
12. Machado, J.B., Isotani, S., Barbosa, E.F., Ribeiro, M.B.: Towards an ontological infrastructure for content modeling and personalization. In: 2012 Seventh International Workshop on Semantic and Social Media Adaptation and Personalization, pp. 107–112. IEEE (2012)
13. Marshall, K.S., Crawford, R., Jensen, D.: Analogy seeded mind-maps: a comparison of verbal and pictorial representation of analogies in the concept generation process. In: ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Charlotte, North Carolina, USA (2016)
14. Zhang, L., Lv, H., Tan, D., Xu, F., Chen, J., Bao, G., Cai, S.: Adaptive quantum genetic algorithm for task sequence planning of complex assembly systems. Electron. Lett. **54**, 870–872 (2018)
15. Fang, K., Uhan, N., Zhao, F., Sutherland, J.W.: A new approach to scheduling in manufacturing for power consumption and carbon footprint reduction. J. Manuf. Syst. **30**(4), 234–240 (2011)
16. Guo, Z.X., Wong, W.K., Leung, S.Y.S., Fan, J.T., Chan, S.F.: Mathematical model and genetic optimization for the job shop scheduling problem in a mixed-and multi-product assembly environment: a case study based on the apparel industry. Comput. Ind. Eng. **50**(3), 202–219 (2006)
17. Chaudhry, I.A., Khan, A.A.: A research survey: review of flexible job shop scheduling techniques. Int. Trans. Oper. Res. **23**, 551–591 (2016)
18. May, G., Stahl, B., Taisch, M., Prabhu, V.: Multi-objective genetic algorithm for energy—game theory approach. Appl. Mech. Mater. **809–810**, 1450–1455 (2015)
19. Krenczyk, D., Olender, M.: Simulation aided production planning and scheduling using efficient job shop scheduling. Int. J. Prod. Res, pp. 1–19 (2015)
20. Bolshakov, A., Slobodyanyuk, L., Shashikhina, O., Kovalchuk Ya.: A combined method for solving the problem of optimizing the production schedule of metal structure processing for use in a cyber-physical control system of a metallurgical enterprise. Studies in Systems, Decision and Control. Cyber-Physical Systems: Modelling and Industrial Application, vol. 418, pp. 243–260. Springer Nature Switzerland AG (2022)

# Intelligent Training System for Controlling Polymer Extrusion Processes in Emergency Situations

**Andrey Polosin, Tamara Chistyakova, Christian Kohlert, and Frank Kleinert**

**Abstract** An intelligent training system for controlling extrusion processes in productions of multi-assortment polymeric films in the event of emergency situations has been developed. Extrusion is one of the main stages of production, which determines the throughput of the system and the consumer characteristics of the products (the number of surface defects of various types, color). These characteristics depend on the thermal state, uniformity, and color of the extrudate. Therefore, emergency situations associated with deviations in the throughput of the extruder and the quality indices of the extrudate due to regulatory restrictions lead to a decrease in output and an increase in film defects. The training system includes a knowledge base of emergency situations, their reasons and recommendations for elimination, a library of mathematical models for calculating throughput and quality indices depending on control actions, databases of film types, polymer properties, geometric parameters of extruders and technological parameters of extrusion, a module for setting the training scenario, the module for handling emergency situations and the formation of control tips, the module for the formation of the training protocol, interfaces of the trainee (the extruder operator) and the instructor. The library of mathematical models allows you to adjust the system for training in the control of extruders of various types. The training scenario set by the instructor includes the type of film, the throughput of the system, the brand of the extruder, characteristics of emergency situation being simulated. The training protocol records the actions of the trainee and the achieved values of the output parameters. Testing according to the extrusion

A. Polosin (✉) · T. Chistyakova
Saint Petersburg State Institute of Technology (Technical University), Moskovsky Prospekt 26, Saint Petersburg 190013, Russia
e-mail: polosin-1976@sapr.lti-gti.ru

T. Chistyakova
e-mail: nov@technolog.edu.ru

C. Kohlert
Kohlert-Consulting, Rotsteinerstr. 14, 56414 Oberahr, Germany
e-mail: info@kohlert-consulting.com

F. Kleinert
Klöckner Pentaplast Europe GmbH & Co. KG, Industriestr. 3-5, 56401 Montabaur, Germany
e-mail: frank.kleinert@kpfilms.com



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_7

processes in calendering productions of pharmaceutical and food films at factories in Russia and Germany has confirmed the operability of the computer system.

**Keywords** Intelligent system · Expert knowledge base · Production-frame model · Databases of raw materials · Equipment · Products · Mathematical models · Control training · Emergency situations · Product defects · Extrusion · Production of polymeric films

## 1 Introduction

Modern production of polymeric films for packaging pharmaceutical and food products are continuous, multistage, large-tonnage (throughput up to 3000 kg/h) technological systems [1–3]. The complexity of these systems is due to the presence of recyclables, which are necessary for the return of waste (cut and crushed edges of films) to production as secondary raw materials [4]. The production of polymeric films is characterized by a variety of equipment, a wide range of products (more than 60 types of films) and, as a result, frequent changeovers to produce a new range of products (up to several times per shift). The consumer characteristics of the film depend in a complex way on the quality indices of the polymeric material at the production stages (the quality of intermediates is often not monitored), the characteristics of raw materials and apparatuses, the technological parameters of the stages. The number of such bonds in the calendered production of polymeric films exceeds 800. Appearance and color characterize the main value of films as packaging materials for consumers. Defects of various types on the surface of the film form its appearance. Defects caused by overheating and destruction of the polymer (black points, destructive stripes), incomplete plasticization of the bulk polymer (non-molten polymeric particles), insufficient mixing of the polymer melt (defects of the "fish eye" type) most often occur and negatively affect the applicability of the film as a package. The color defects of the film include mismatch of color coordinates with reference values and uneven coloring (insufficient mixing of the melt is also the cause of uneven coloring). The extrusion process plays a key role in ensuring the required consumer characteristics of the film (a limited number of surface defects, color compliance with the standard, uniform coloring). It is designed to produce a homogeneous viscous mass (extrudate) by melting the solid phase moving in the screw apparatus (extruder), heating, mixing and pushing the melt through the forming head of the extruder. Then (on the calender) the extrudate is formed into a film. Deviations of the extrusion regime from the regulations contribute to violations of hydrodynamic, thermal conditions and the mechanism of phase transition, the formation of stagnant zones in the extruder. As a result, the quality of the extrudate deteriorates greatly, which does not allow the use of the extrudate for further forming. For example, one of the main defects of the extrudate—thermal destruction (destruction of chains of polymer macromolecules [5])—is a consequence of excessive (due to stagnant zones) residence time of the polymer in the extruder at high temperature. Thermal degradation changes the color of the extrudate (from yellowing to darkening), increases its creep and adhesion to the conveyor between the extruder and the calender. The returnable waste fed into the extruder contributes to an increase in the color deviation of the extrudate from the standard in the production of colored polymeric films. The defect of the extrudate is not eliminated at the next stages of production (forming of the extrudate, fixing of the film structure) and leads to the appearance of corresponding defects on the surface of the film. Up to 20% of polymeric film defects are caused by violations of the extrusion regime. The lack of synchronization of the extruder and the calender can lead not only to a decrease in the yield of high-quality products, but also to the shutdown of the production system [6]. With excessive throughput of the extruder, the extrudate accumulates in the feeding gap of the calender, overheats due to a long (~5 min) stay in stock and collapses under the influence of high temperature. When the calender is "starved", the extrudate stock in the calender feeding gap is not formed, which leads to an emergency situation (collision of the calender feeding rolls). In order to ensure the absence of both defects and accidents, the operator must reduce (increase) the throughput of the extruder, based on a visual assessment of the level of the extrudate stock. In the production of colored films (using the method of staining the polymer melt with liquid colorants dosed into the extruder), such a throughput adjustment is a disturbance leading to a color deviation of the extrudate from the standard. This requires a corresponding change in the control actions, that is, the flow rates of colorants. If the consumer characteristics of the film deviate beyond the limit values according to the regulations, then the film cannot be used as packaging (since it does not ensure the safety and attractiveness of products for consumers) and is disposed of as non-returnable waste. An increase in the amount of defective films leads to a decrease in the competitiveness of production and an increase in its danger to the environment. Therefore, timely detection and correction of emergency situations associated with deviations in the throughput of the extruder, the quality indices of the extrudate beyond the threshold regulatory values, help to prevent the appearance of relevant film defects (reduce the amount of non-returnable waste) and increase the yield of quality products. However, the complexity of the hardware and technological design of the extrusion process [7–9], the lack of a system for monitoring the throughput of the extruder and the quality of the extrudate (as a rule, the quality is assessed visually by the appearance of the extrudate), many reasons of emergency situations significantly complicate the control of extrusion in emergency situations. This leads to operators making erroneous control decisions. In this regard, it is necessary to provide an opportunity for operational staff in advance, using a computer training system, to work out the skills of extrusion control in emergency situations for various types of polymeric films and options for hardware and technological design of the process.

The analysis of computer systems for training production personnel of high-tech facilities of the chemical, oil refining, metallurgical industry, energy has showed that the most effective way to solve the problems of training in controlling complex technological processes in various modes of their functioning (adjustment for a new task, routine mode, emergency situations) is the development and implementation of intelligent training systems [10–17]. These systems are based on multi-variate models of the description of control objects, adjusted to their variable characteristics. The training (simulator) models include:

- models for the representation of informal expert knowledge for the study of ways to control technological processes, ways to eliminate emergency situations and the formation of intellectual tips on control;
- mathematical models that allow calculating process efficiency indices (throughput, energy consumption, quality indices of intermediates and products) depending on control actions and providing opportunities for active training;
- information models of raw materials, equipment, technological modes, products that ensure the adjustment of training systems for various modifications of hardware and technological design of processes [18–20].

Thus, in order to maintain at the appropriate level and improve the professional competencies of the operators of extrusion processes in the production of polymeric films, the development of adjustable intelligent training system is relevant. Such a system allows (based on the model of representing expert knowledge about emergency situations of the extrusion process and the mathematical model for calculating the throughput of the extruder and the quality indices of the extrudate) to form extrusion controlling skills in emergency situations for various training scenarios (types of films, brands of extruders).

## 2 Formulation of the Problem of Extrusion Control Training

The analysis of the polymer extrusion process as an object of study and control has allowed forming its informational description as a set of vectors of input parameters $X$, control actions $U$, disturbances $F$ and output parameters $Y$ (state parameters $S$ and efficiency indices $Q$):

$$
\begin{aligned}
&Y = F(X, U, F),\\
&Y = \{S, Q\},\\
&S = \{v_x, v_z, P, T, \eta, \tau_{av}\},\\
&Q = \{G, Q_{ext}\},\\
&Q_{ext} = \{I_d, \varphi_s, \gamma_{av}, L_{ext}, a_{ext}, b_{ext}, \Delta E_{ext}\},\\
&X = \left\{T_{film}, H_{polym}, M_{extrud}, \Gamma_{extrud}\right\},\\
&M_{extrud} = \left\{T_{extrud}, D, L/D\right\},\\
&\Gamma_{extrud} = \{n_{scr}, \Gamma_{scr}, \Gamma_{die}, S_0, A_{scr}\},\\
&U = \{N, T_{bk}, k = 1 \ldots n_T, G_{ci}, i = 1 \ldots n_c\},\\
&F = \{G_{wast}\}.
\end{aligned}
$$

In the production of polymeric films characterized by the types $T_{film}$, extruders of various types $T_{extrud}$ are used (single-screw, reciprocating, twin-screw extruders [7–9]). Extruders of various types differ in the number of screws $n_{scr}$ and the nature of their movement (rotation or simultaneous rotation and axial reciprocation). Extruders of the same type differ in brands $M_{extrud}$, characterized by the diameter $D$ and the relative length $L/D$ of the screw/screws. The geometric parameters of the extruder of this brand include the geometric parameters of the screw/screws $\Gamma_{scr}$ and the forming head (die) $\Gamma_{die}$, as well as parameters characterizing the features of the type of extruder (the screw reciprocation amplitude of the reciprocating extruder $S_0$, the axial distance of screws of the twin-screw extruder $A_{scr}$). The type of film determines the type of film-forming polymer $T_{polym}$, processed in an extruder. The parameters of the thermos-physical properties of the solid phase and the melt $H_{polym}$ depend on the type of polymer. The intensity and duration of the thermal and mixing effects on the polymer in the extruder vary by varying the control actions. These include the speed of the screw/screws $N$, temperatures of the thermal zones of the barrel $T_{bk}$ and flow rates of liquid colorants $G_{ci}$. Returnable waste flow rate $G_{wast}$ is a disturbance that leads to deviations of the extrusion mode from the regulations. The state of the process is characterized by the velocities of the cross-channel and down-channel flows $v_x$, $v_z$, pressure $P$, temperature $T$, viscosity $\eta$ of polymer melt and average residence time in the extruder $\tau_{av}$. The throughput of the extruder $G$ and the quality indices of the extrudate $Q_{ext}$ are indices of the extrusion efficiency $Q$. The quality of the extrudate is assessed according to the following indices:

- thermal destruction index $I_d$ (mapping the dependence of the degree of destruction of the polymer on time under the temperature-time regime in the extruder on the dependence that was experimentally obtained under isothermal conditions corresponding to the beginning of color change due to destruction);
- characteristics of material homogeneity – the fraction of the solid phase $\varphi_s$, average mixing degree $\gamma_{av}$ (shear deformation accumulated during residence time in extruder);
- color coordinates (in CIELab space) $L_{ext}$, $a_{ext}$, $b_{ext}$, color deviation from the standard $\Delta E_{ext}$ [21].

Emergency situations of extrusion are primarily associated with extrudate defects (thermal destruction, material heterogeneity, color deviations). In order to exclude film defects resulting from poor-quality extrudate preparation, its quality indices must meet the requirement $Q_{ext} \subset [Q_{ext}{}^{min}; Q_{ext}{}^{max}]$. This requirement is formed from the condition of the suitability of the extrudate for forming on a calender. Here $Q_{ext}{}^{min}$, $Q_{ext}{}^{max}$ are vectors of threshold values of quality indices (for example, the maximum index of destruction $I_d{}^{max}$, minimum degree of mixing $\gamma_{av}{}^{min}$). They depend on the type of film $T_{film}$. For example, for pharmaceutical films, black points are the least acceptable, so the restriction on the destruction index is more stringent compared to films for the manufacture of plastic cards.

The problem of training to control the extrusion process in emergency situations is as follows: for the training scenario $S_{train} = \{T_{film}, G_0, M_{extrud}, i_{St}, i_{Rs}\}$ set by the instructor, it is necessary by varying the control actions in the regulatory ranges

$U \subset [U^{min}; U^{max}]$ in accordance with the control tip *Rc*, formed by an intelligent training system (based on the analysis of the reasons of an emergency situation $Rs_j$, $j = 1 \ldots n_{Rs}$ and determining the true reason $Rs^*$), to eliminate an emergency situation *St*, by entering the appropriate efficiency index into the range of acceptable values $Q_l \in [Q_l^{\ min}; Q_l^{\ max}]$.

Here $G_0$ is the system throughput; $i_{St}$ is the number of the simulated emergency situation; $_{Rs}$ $i_{Rs}$ is the number of the simulated reason of the emergency; $U^{min}$, $U^{max}$ are vectors of threshold values of control actions, depending on the type of film $T_{film}$ and the brand of the extruder $M_{extrud}$; *Rc* is the tip containing the direction of changing the control actions on the extruder; $Q_l$ is the index of the efficiency of the process, the deviation of which is prohibitive according to the regulations ($\Delta_l = Q_l^{\ min} - Q_l > 0 \vee \Delta_l = Q_l - Q_l^{\ max} > 0$) is a sign of an emergency situation *St*.

## 3 Structure of the Intelligent Training System

The structure of the intelligent training system developed to solve the training problem is shown in Fig. 1.

The intelligent training system includes a module for setting a training scenario, a bank of data and knowledge about the characteristics of the extrusion process, a subsystem for modeling the process, a module for handling emergency situation and forming a process control tip, a module for forming a training protocol, interfaces of the trainee (extruder operator) and the instructor. The system forms tips for the trainee on the direction of changing control actions in order to eliminate emergency situations set by the instructor. Emergency situations are associated with deviations of efficiency indices (throughput, extrudate quality indices) beyond the threshold values. Expert knowledge about emergency situations of the extrusion process is obtained by interviewing controlling production staff, studying technological regulations and instructions for controlling the production of polymeric films in emergency situations. This knowledge is complexly structured. Therefore, for their formalized representation, a knowledge base of emergency situations, their reasons and recommendations based on a production-frame model of knowledge representation has been developed [18, 19]. Subject knowledge about emergency situation (ES) is presented in the form of a prototype frame, which is described in the Backus–Naur notation: $Fr ::= \, < \text{ES}, Q_{Fr}, A_Q >$, $Q_{Fr} = \{q_1, q_2, q_3\}$, $A_Q = \{a_{1.1}, \ldots, a_{3.4}\}$, where $Q_{Fr}$ is a set of frame attributes that allow you to structure knowledge for emergency situations ($q_1$), the reasons for their occurrence ($q_2$) and recommendations for elimination ($q_3$), $A_Q$ is set of characteristics of attributes that make up their linguistic and informational description. In order to represent knowledge about specific emergency situations of the extrusion process, a transition from a prototype frame to a system of example frames (see Fig. 2) has been performed. To obtain an example frame of a specific emergency situation, the characteristics of each attribute of the prototype frame $a_{1.1}, \ldots, a_{3.4}$ values describing this situation are assigned.

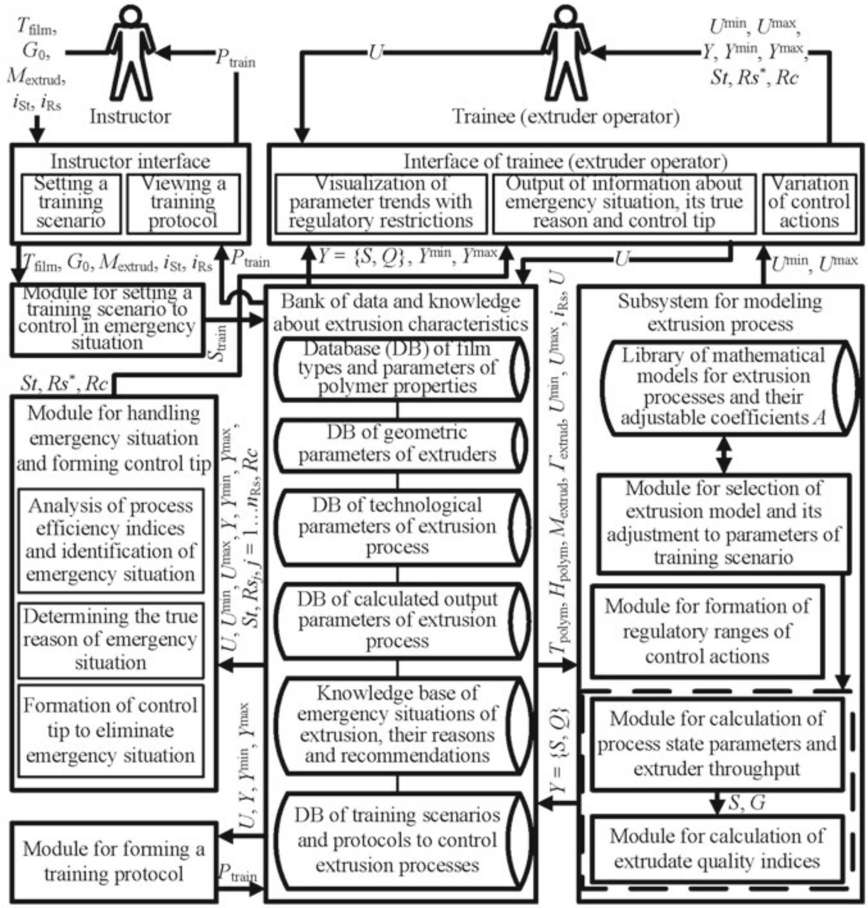



**Fig. 1** Structure of an intelligent training system for controlling extrusion processes in emergency situations

The knowledge base includes a description of emergency situations that are associated with extrudate defects (thermal destruction, non-melted inclusions, etc.) and deviations in the throughput of the extruder, leading to fluctuations in the extrudate stock when feeding the calender. These situations are fixable, as they are eliminated by purposefully changing the control actions without stopping the production system. Priorities of emergency situations are formed depending on the degree of negative impact of situations on the operation of the production system and consumer characteristics of polymeric films. These priorities determine the order of elimination of emergency situations in the event of several situations. Each emergency situation in specific production conditions occurs for one of several possible reasons. Therefore, one of the tasks of an intelligent training system is to determine the true reason of an emergency situation. Each reason is matched with a specific recommendation

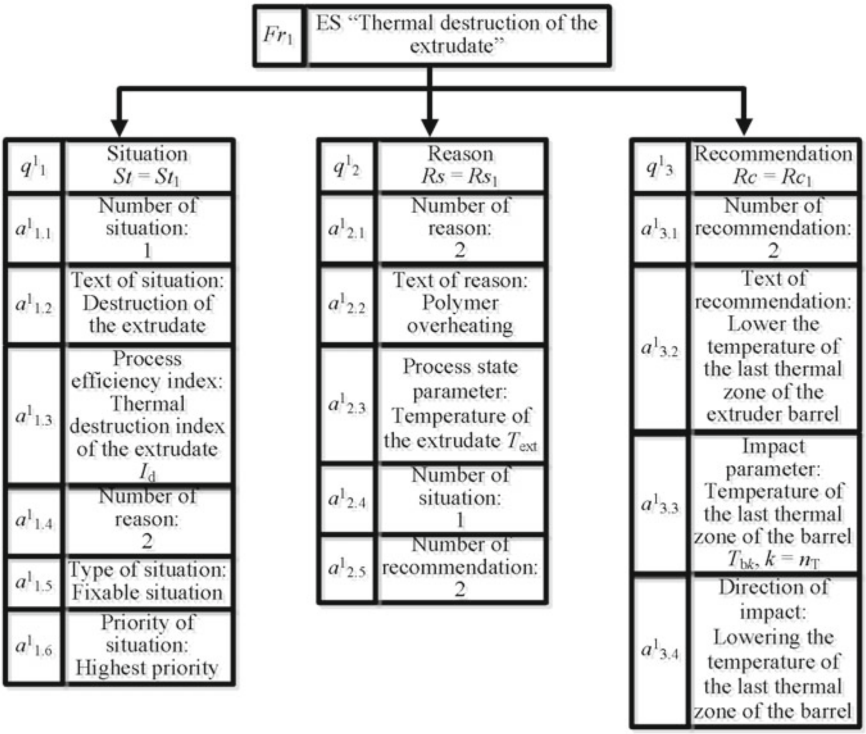



**Fig. 2** Information structure of the example frame Emergency situation "Thermal destruction of the extrudate"

describing the actions of the extruder operator that must be performed to eliminate the reason, and hence the emergency situation caused by it. The controlling knowledge necessary to identify emergency situations, determine their true reasons and form control tips are represented in the form of production rules: $Pr$:: = IF (condition), THEN (consequence). For example, to identify an emergency situation "Thermal destruction of the extrudate" ($St_1$) and the formation of the control tip the following rules are implemented: a simple rule IF ($I_d > I_d^{max}$), THEN ($St = St_1$); focusing rules IF ($St = St_1$) ∧ ($T_{ext} > T_{ext}^{max}$), THEN ($Rs^* = Rs_1$), IF ($St = St_1$) ∧ ($Rs^* = Rs_1$), THEN ($Rc = Rc_1$).

Simulation of the occurrence of emergency situations is carried out in the system by modeling the reasons of situations specified in training scenarios $S_{train}$. To do this, the values of the corresponding parameters of the mathematical model of the extrusion process (input parameters $X$, control actions $U$ or adjustable coefficients of the equations of model $A$) are changed in relation to their nominal values $X^{nom}$, $A^{nom}$ or regulatory ranges [$U^{min}$; $U^{max}$], stored in the data bank and model library. A variety of types of extruders and polymers, the complexity of hydrodynamic and heat and mass transfer processes occurring in the screw channel of the extruder, non-Newtonian behavior of melts require an integrated approach to modeling extrusion processes. At the same time, existing mathematical models describe processes in extruders of certain types [22, 23] or in separate functional zones of the extruder [24–32]. As a rule, they do not allow us to take into account all the features of extrusion (phase transition, viscosity anomaly, melt leakage through clearances, cuts in screw flights, energy dissipation, heat exchange of the polymer with the barrel, screw) and to give a comprehensive quantitative assessment of the characteristics of the process, taking into account the quality of the extrudate. Therefore, a library of mathematical models has been developed to adjust the system for training in the control of multi-assortment extrusion processes in various types of apparatuses. It includes models of polymer processing in single-screw, reciprocating and twin-screw extruders having screws of a typical configuration and equipped with dies [7–9]. The models are deterministic and are based on the laws of conservation of mass, momentum, energy (for constructing equations of phase and interphase balances), the laws of rheology (for constructing equations of viscosity dependence on shear strain rates and temperature) [21, 33]. When constructing models, the following general assumptions (for all types of extruders) are accepted: reversed movement of the barrel and screw; smallness of curvature and complete filling of the screw channel; steady-state nature of the process over time and along the length of the channel; constancy of polymer properties; absence of radial flow, inertial and mass forces, convective heat transfer across the channel width and heat transfer by thermal conductivity along the channel length; melt adhesion to the channel walls. The model of the extrusion process is selected from the model library depending on the type of extruder $T_{\text{extrud}}$. Adjusting the model to the polymer type $T_{\text{polym}}$ and the brand of the extruder $M_{\text{extrud}}$ consists in forming the nominal values of the parameters of the properties of the phases of the extruded polymer $H_{\text{polym}}$, geometric parameters of the extruder $\Gamma_{\text{extrud}}$, coefficients of the model $A$. The adjustable coefficients of the model are the coefficients of the rheological equation of the melt state (for example, the power law index $n$), heat transfer coefficients between the polymer phases, the barrel and the screw of the extruder, coefficients of the thermal destruction model (time $\tau_{\text{d}}$, temperature $T_{\text{d}}$ and activation energy $E_{\text{d}}$ of destruction), coefficients of regression models for calculating color coordinates ($c_p$, $p = 0 \ldots n_{\text{L}}$, $d_q$, $q = 0 \ldots n_{\text{a}}$). Due to the nonlinearity of the model equations, numerical methods (methods of the theory of plane asymmetric flows, finite difference methods [34]) are used to solve them. The adjusted extrusion model allows you to calculate distributions of flow velocities $v_{\text{x}}$, $v_{\text{z}}$, pressure $P$, temperature $T$, viscosity $\eta$ of melt through the screw channel, average residence time $\tau_{\text{av}}$ and throughput $G = \rho Q_{\text{E}}$. Here $\rho$ is the melt density, $Q_{\text{E}}$ is the volumetric down-channel flow rate through the "screw channel—die" system, calculated by iteratively solving the equation formed from the coupling condition of the screw and the die (equality of pressure differences in the screw channel and the die). So, for a reciprocating extruder, this equation has the following form:

$$(\pi N)^n \left(1 + \frac{1}{n}\right)^n \int_0^Z (D\cos\varphi + S_0 \sin\Phi_{osc} \sin\varphi)^n \frac{B_z \mu}{H^{n+1}} \frac{F_d - Q_E / Q_d}{(1 - Q_E / Q_d) F_P} dz - \mu_{die} \left(\frac{Q_E}{k_{die}}\right)^n = 0, \quad (1)$$

where $Z$, $H$ are the length and the depth of the screw channel; $\phi$ is the angle of inclination of the screw flights; $\varPhi_{osc}$ is the screw reciprocation phase; $B_z$ is the dimensionless pressure gradient in the down-channel flow of the melt; $\mu$, $\mu_{die}$ are the melt consistency coefficients in the screw channel and the die; $F_d$, $F_P$, $k_{die}$ are coefficients of the geometric shape of the screw channel and the die; $Q_d$ is the volumetric forced flow rate in the screw channel.

According to the parameters of the state $S$, the quality indices of the extrudate are calculated:

$$I_d = \frac{\tau_{av}}{\tau_d} \exp\left[\frac{E_d}{8.31(T_{ext} + 273)(T_d + 273)}(T_{ext} - T_d)\right]100, \quad T_{ext} = T|_{z=Z}, \quad (2)$$

$$\gamma_{av} = \tau_{av}(ZH)^{-1} \int_0^Z \int_0^H \sqrt{(\partial v_x / \partial y)^2 + (\partial v_z / \partial y)^2}\, dy dz, \quad (3)$$

$$L_{ext} = f_L(k_i,\ i = 1 \ldots n_c,\ c_p,\ p = 0 \ldots n_L), \quad a_{ext} = f_a(k_i,\ i = 1 \ldots n_c,\ d_q,\ q = 0 \ldots n_a), \quad (4)$$

$$\Delta E_{ext} = \sqrt{(L_{ext}{}^* - L_{ext})^2 + (a_{ext}{}^* - a_{ext})^2 + (b_{ext}{}^* - b_{ext})^2}, \quad k_i = G_{ci} / G, \quad (5)$$

where $L_{ext}{}^*$, $a_{ext}{}^*$, $b_{ext}{}^*$ are the color coordinates of the standard, depending on the type of film $T_{film}$.

Thus, the developed mathematical models allow us to calculate the parameters of the state $S$ and the efficiency indices $Q$ of the extrusion process depending on the control actions $U$ for various types $T_{extrud}$ and brands $M_{extrud}$ of extruders, types of processed polymers $T_{polym}$. To check the adequacy of the mathematical models, the measured and calculated values of the process throughput, temperature and color coordinates of the extrudate have been compared. The data has been obtained for reciprocating and twin-screw extruders of various brands during the processing of rigid polyvinyl chloride (PVC) and the implementation of different extrusion modes (by varying control actions). The results of statistical processing of the obtained data have confirmed the adequacy of the models according to the Fisher criterion, the coefficient of determination and the standard deviation, which does not exceed 3–5% for various output parameters of the process.

The adjustment of the intelligent training system to the parameters of the training scenario is carried out using an information model of the extrusion process. This model is implemented as a set of relational databases. The database of film types and parameters of polymer properties allows you to form nominal values of parameters of thermo-physical properties of polymer phases $H_{\text{polym}}$ (depending on the type of polymer $T_{\text{polym}}$) and threshold values of extrudate quality indices $Q_{\text{ext}}{}^{\min}$, $Q_{\text{ext}}{}^{\max}$ (depending on the type of film $T_{\text{film}}$). The database of geometric parameters of extruders allows you to form nominal values of geometric parameters $\varGamma_{\text{extrud}}$ (depending on the brand of the extruder $M_{\text{extrud}}$). The database of technological parameters of the process allows you to form regulatory restrictions on control actions $U^{\min}$, $U^{\max}$, process state parameters $S^{\min}$, $S^{\max}$, extruder throughput $G^{\min}$, $G^{\max}$ (depending on the type of film $T_{\text{film}}$, extruder brand $M_{\text{extrud}}$, system throughput $G_0$). Arrays of values of control actions are stored in the database of technological parameters, which are imported from the database of measured parameters of industrial production of polymeric films, which is included in the software package for big data processing and film quality control [35]. These arrays are used to simulate the extrusion process in time. According to the values of the control actions $U$, which change over time, the output parameters $Y$ are calculated and displayed (in the form of trends). The modeling results are stored in a database of output parameters. Training scenarios and results are stored in a database of scenarios and protocols. Training protocol $P_{\text{train}}$, generated by the system, contains the values of the control actions $U$, set by the trainee to eliminate an emergency situation, as well as the calculated (for them) and threshold values of the output parameters $Y$, $Y^{\min} = \{S^{\min}, G^{\min}, Q_{\text{ext}}{}^{\min}\}$, $Y^{\max} = \{S^{\max}, G^{\max}, Q_{\text{ext}}{}^{\max}\}$. The training goal is achieved if the trainee has set the extrusion mode $U \subset [U^{\min}; U^{\max}]$: $Y \subset [Y^{\min}; Y^{\max}]$.

## 4 Algorithm for Modeling Emergency Situations and Forming Tips for Extrusion Control

The algorithm for forming tips for eliminating emergency situations is shown in Fig. 3.

For a given training scenario $\text{S}_{\text{train}}$ threshold values of control actions $U^{\min}$, $U^{\max}$ and output parameters $Y^{\min}$, $\text{Y}^{\max}$ of extrusion process, an array of values of control actions $U$ (to simulate the process in time) are formed, an extrusion model is selected, and its parameters X, U, A are adjusted to simulate the occurrence of a given emergency situation $i_{\text{St}}$. The values of the process efficiency indices $Q$ calculated by the mathematical model are compared with the threshold values $Q^{\min}$, $Q^{\max}$. In case of deviation of the efficiency index $Q_l$ beyond the threshold values an emergency situation $St$ is identified, the information description of which is the index $Q_l$.

For an identified situation, a list of possible reasons of its occurrence $R_{sj}$, $j = 1 \ldots n_{\text{Rs}}$ is formed from the knowledge base. Each reason is described by one of the process parameters $(U, S)$. To determine the true reason $Rs^*$ when this situation occurs,

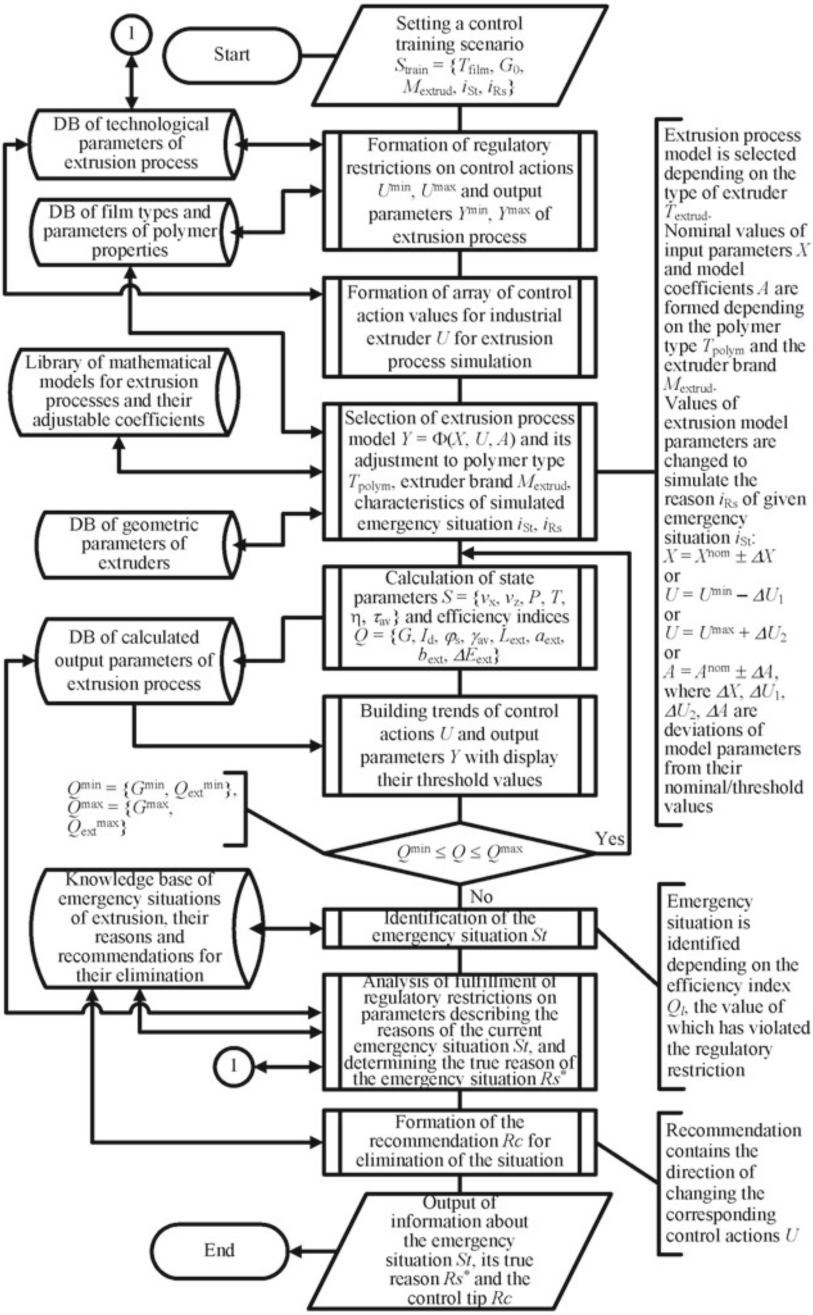



**Fig. 3** The algorithm for modeling emergency situations and forming control tips

the values of these parameters are compared with the regulatory threshold values, and the parameter whose value violated the restrictions is identified. To control the correctness of the system, the equality of the number of the true reason and the number $i_{\mathrm{Rs}}$ specified in the training scenario is checked. For the reason found, a recommendation *Rc* is extracted from the knowledge base. It contains the direction of changing the corresponding control actions *U* to eliminate this reason, and hence the emergency situation. The trainee receives a message about an emergency situation *St*, its true reason $Rs^*$ and the control tip *Rc*. By changing the control actions on the process accordingly, the trainee eliminates an emergency situation, transferring the process to a routine mode.

The proposed algorithm makes it possible to solve the problem of training to control the extrusion process in various emergency situations for different variants of the hardware and technological design of the process and types of polymers, since these characteristics are set in the training scenario for which the algorithm is adjusted.

## 5 Testing of the Intelligent Training System

The system has been tested using examples of training to control extrusion processes in emergency situations associated with various types of extrudate defects (thermal destruction, inclusions of non-molten polymer, insufficient mixing, color deviations) and arising for various reasons (excessive/insufficient residence time, overheating/insufficient heating, insufficient mixing intensity, etc.). Data of calendered production of pharmaceutical and food packaging films of various types (based on rigid PVC) at factories in Russia and Germany has been used for testing. Reciprocating extruders and counter-rotating twin-screw extruders have been used as the hardware design of the extrusion processes.

An example of the trainee's interface is shown in Fig. 4. Here the trends of the screw speed *N* of the PR-200 reciprocating extruder and the index of thermal destruction of the extrudate $I_{\mathrm{d}}$ are displayed with threshold values $N^{\min}$, $N^{\max}$, $I_{\mathrm{d}}^{\max}$. It also displays information about an emergency situation related to the destruction of the extrudate ($I_{\mathrm{d}} > I_{\mathrm{d}}^{\max}$), its true reason (excessive residence time of the polymer in the extruder: $\tau_{\mathrm{av}} > \tau_{\mathrm{av}}^{\max}$, where $\tau_{\mathrm{av}}^{\max}$ is the upper threshold value of the residence time), modeled by reducing the screw speed *N* beyond the lower threshold value $N^{\min}$, and the control tip (increasing the screw speed *N*). Implementing the control tip, the trainee enters the destruction index into the regulatory range (which is displayed on the trend of the destruction index), eliminating the emergency situation and preventing the corresponding surface defects of the polymeric film.

The test results have confirmed the operability of the system for the formation of skills for controlling extrusion processes in the event of emergency situations.

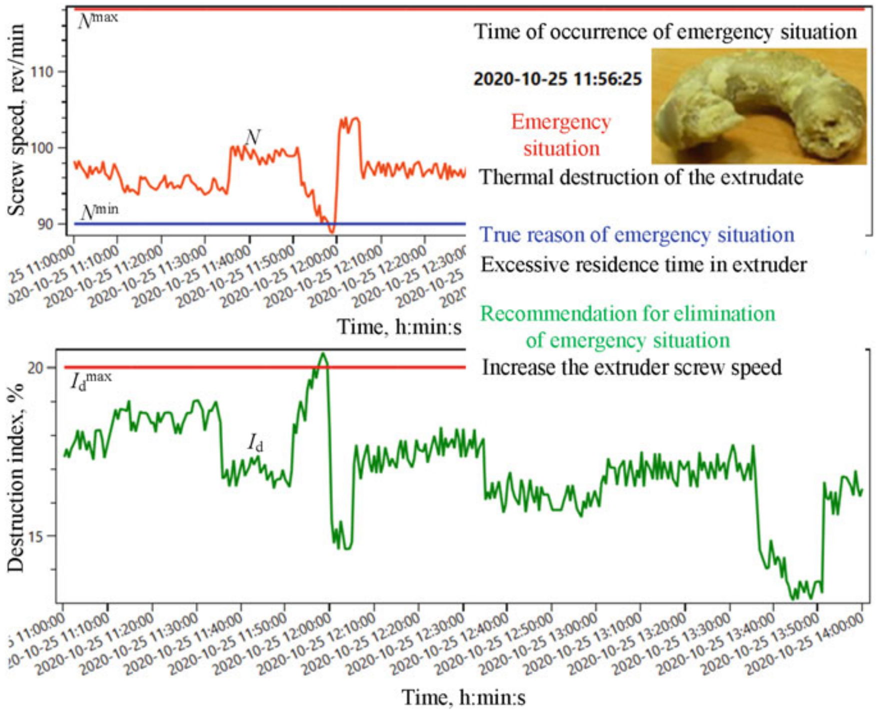



**Fig. 4** Trends in the parameters of the extrusion process and a control tip for the trainee

## 6 Conclusion

An adjustable intelligent training system has been developed, which is designed to train extruder operators of multi-assortment production of polymeric films. The system allows you, on the basis of multi-variate models for describing extrusion processes (a production-frame model for representing expert knowledge, functional mathematical models, relational information model), to solve the problem of forming skills for controlling extrusion processes in the event of emergency situations for various training scenarios (various types of films produced, hardware and technological design of the process, emergency situations and the reasons of their occurrence).

The results of testing according to the data of industrial production of polymeric films for packaging pharmaceuticals and food products have confirmed the operability of the system as a computer tool for improving the professional competence of extruder operators, contributing to reducing the time of control decision-making to eliminate emergency situations of the extrusion process. This leads to an improvement in the consumer characteristics of films, resource saving, an increase in the yield of quality products and an increase in the environmental safety of production (by reducing the amount of non-returnable waste following the timely elimination of emergency situations of the extrusion process and preventing the corresponding defects of the polymeric film).

## References


1. Gedde, U.W., Hedenqvist, M.S., Hakkarainen, M., Nilsson, F., Das, O.: Processing of polymeric materials. In: Applied Polymer Science, pp. 453–487. Springer, Cham (2021)
2. Rosato, D.V.: Calendering. In: Extruding Plastics, pp. 418–440. Springer, Boston (1998)
3. Saha, N.C., Anup, K.G., Meenakshi, G., Susmita, D.S.: Food Packaging: Materials, Techniques and Environmental Issues. Springer Singapore (2022)
4. Martens, H., Goldmann, D.: Recyclingtechnik. Springer, Berlin (2016)
5. Owen, E.D. (ed.): Degradation and Stabilisation of PVC. Springer, Dordrecht (1984)
6. Gartner, H., Schnabel, A., Kohlert, C.: Control systems for calenders—from mixer to take-off. Kunstst. Plast Eur. **87**(6), 24–27 (1997)
7. Rauwendaal, C.: Polymer Extrusion, 5th edn. Carl Hanser, Munich (2014)
8. White, J.L., Potente, H. (eds.): Screw Extrusion: Science and Technology. Carl Hanser, Munich (2003)
9. Kohlgrüber, K.: Co-Rotating Twin-Screw Extruders: Fundamentals, Technology, and Applications. Carl Hanser, Munich (2008)
10. Reinig, G., Winter, P., Linge, V., Nägler, K.: Training simulators: engineering and use. Chem. Eng. Technol. **21**(9), 711–716 (1998)
11. Gutierrez, J., Elopriaga, J.A., Fernandez-Castro, I., Vadillo, J.A., Diaz-Ilarraza, A.: Intelligent tutoring systems for training of operators for thermal power plants. Artif. Intell. Eng. **12**(3), 205–212 (1998)
12. Samarina, A.M., Ershova, O.V.: Architecture of a software training complex for aluminum industry personnel. Autom. Remote Control **71**(8), 1711–1717 (2010)
13. Dozortsev, V.M.: Methods for computer-based operator training as a key element of training systems. Autom. Remote Control **74**(7), 1191–1200 (2013)
14. Patle, D.S., Ahmad, Z., Rangaiah, G.P.: Operator training simulators in the chemical industry. Rev. Chem. Eng. **30**(2), 199–216 (2014)
15. Ahmad, Z., Patle, D.S., Rangaiah, G.P.: Operator training simulator for biodiesel synthesis from waste cooking oil. Process Saf. Environ. Prot. **99**, 55–68 (2016)
16. Marcano, L., Haugen, F.A., Sannerud, R.: Review of simulator training practices for industrial operators: how can individual simulator training be enabled? Saf. Sci. **115**, 414–424 (2019)
17. Dozortsev, V.M., Agafonov, D.V., Nazin, V.A., Novichkov, A.Y., Frolov, A.I.: Computerized operator training: continued importance, new opportunities, and the human factor. Autom. Remote Control **81**(5), 935–954 (2020)
18. Chistyakova, T.B., Novozhilova, I.V.: Intelligence computer simulators for elearning of specialists of innovative industrial enterprises. In: Proceedings of the XIX International Conference on Soft Computing and Measurements, pp. 329–332. IEEE (2016)
19. Chistyakova, T., Reinig, G., Novozhilova I.: Use of computer trainers for teaching management manufacturing personnel of chemical industries. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M.V. (eds.) Cyber-Physical Systems: Design and Application for Industry 4.0 2020, SSDC, vol. 342, pp. 371–382. Springer, Cham (2021)
20. Ershova, O., Chistyakova, T.: Methods and technologies of developing computer simulators for training resource-saving control of electrotechnological installations. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M.V. (eds.) Society 5.0: Human-Centered Society Challenges and Solutions 2021, SSDC, vol. 416, pp. 351–363. Springer, Cham (2022)
21. Chistyakova, T.B., Polosin, A.N.: Computer modeling system for energy- and resource-saving control of multi-assortment polymeric film production. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M.V. (eds.) Cyber-Physical Systems: Design and Application for Industry 4.0 2020, SSDC, vol. 342, pp. 119–131. Springer, Cham (2021)

22. Lee, K.Y., Han, C.D.: Analysis of the performance of plasticating single-screw extruders with a new concept of solid-bed deformation. Polym. Eng. Sci. **30**(11), 665–676 (1990)
23. Lewandowski, A., Wilczyński, K.J., Nastaj, A., Wilczyński, K.: A composite model for an intermeshing counter-rotating twin-screw extruder and its experimental verification. Polym. Eng. Sci. **55**(12), 2838–2848 (2015)
24. Basov, N.I., Volodin, I.N., Kazankov, Yu.V., Pervushin, V.E.: Hydrodynamics and heat exchange in melting in the screw channel of an extruder. Theor. Found. Chem. Eng. **17**(1), 61–67 (1983)
25. Elbirli, B., Lindt, J.T., Gottgetreu, S.R., Baba, S.M.: Mathematical modeling of melting of polymers in a single-screw extruder. Polym. Eng. Sci. **24**(12), 988–999 (1984)
26. Pervadchuk, V.P., Trufanova, N.M., Yankov, V.I.: Mathematical model and numerical analysis of heat-transfer processes associated with the melting of polymers in plasticating extruders. J. Eng. Phys. **48**(1), 60–64 (1985)
27. Potente, H., Melish, U.: Theoretical and experimental investigations of the melting of pellets in co-rotating twin-screw extruders. Int. Polym. Proc. **11**(2), 101–108 (1996)
28. Vergnes, B., Souveton, G., Delacour, M.L., Ainser, A.: Experimental and theoretical study of polymer melting in a co-rotating twin-screw extruder. Int. Polym. Proc. **16**(4), 351–362 (2001)
29. Lyu, M.-Y., White, J.L.: Simulation of non-isothermal flow in a modular Buss kneader and comparison with experiment. Int. Polym. Proc. **12**(2), 104–109 (1997)
30. Sardo, L., Vergnes, B., Valette, R.: Numerical modelling of the non-isothermal flow of a non-Newtonian polymer in a co-kneader. Int. Polym. Proc. **32**(4), 425–433 (2017)
31. White, J.L., Chen, Z.: Simulation of non-isothermal flow in modular co-rotating twin screw extrusion. Polym. Eng. Sci. **34**(3), 229–237 (1994)
32. Teixeira, C., Gaspar-Cunha, A., Covas, J.A.: Flow and heat transfer along the length of a co-rotating twin screw extruder. Polym.-Plast. Technol. Eng. **51**(15), 1567–1577 (2012)
33. Polosin, A.N., Chistyakova, T.B.: Mathematical modeling for resource and energy saving control of extruders in multi-assortment productions of polymeric films. J. Phys: Conf. Ser. **1015**(3), 032104 (2018)
34. Chiruvella, R.V., Jaluria, Y., Sernas, V.: Extrusion of non-newtonian fluids in a single-screw extruder with pressure back flow. Polym. Eng. Sci. **36**(3), 358–367 (1996)
35. Chistyakova, T.B., Kleinert, F., Teterin, M.A.: Big data analysis in film production. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M.V. (eds.) Cyber-Physical Systems: Advances in Design and Modelling 2019, SSDC, vol. 259, pp. 229–236. Springer, Cham (2020)

# Method for Constructing Virtual Reality Simulators for Turning and Milling for an Engineering Education System for Building Cyber-Physical Systems

**Roman Aslanov and Alexander Bolshakov**

**Abstract** The chapter describes the development and implementation of three-dimensional computer simulator for turning and milling in virtual reality, the architecture of hardware-software complex and its functionality based on the proposed methodology. The principles of user interaction with three-dimensional models, UI/UX design are considered. The simulator is designed to solve the tasks of visualization of three-dimensional models according to the specified scenarios during training and qualification assessment activities. Practical application in the implementation of effective training for industrial enterprises, optimization of physical and financial costs, minimizing the number of errors and accidents, absence of the possibility of damage to the equipment, the need for a special equipped room, depreciation accrual, maintenance costs and consumables: cutters, cutters, drills and workpieces. The variability of the scenario allows a diverse range of operations to be performed in the application. The novelty of the proposed project is: a model of professional training and qualification assessment capable of evaluating students' work processes and professional skills; a specialized software package with unique program modules; method of optimizing program modules and three-dimensional models of the program complex of industrial segment training in a single virtual space.

**Keywords** Immersive technologies · Virtual reality · Three-dimensional visualization · Computer simulators · Deep learning · Computer graphics

R. Aslanov (✉)
Moscow City Pedagogical University, Moscow 109028, Russia
e-mail: aslanov.boxing@mail.ru

A. Bolshakov
Peter the Great St. Petersburg Polytechnic University (SPbPU), Saint Petersburg, Russia 195251



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_8

## 1 Introduction

Improving the training process for high-tech industry requires special attention. In this regard, the study of methods and means of applying immersive technologies to improve the quality and accessibility of training is very important in today's society [1–6]. It should be noted that computer simulators are quite successfully used to train operators in various subject areas [7–12]. In this case, the effect of virtual reality is used [13–15]. Virtual reality systems are used in engineering education [16, 17], medicine [18–20], for people with disabilities [21], and driving education [22].

Virtual reality simulators are used more and more widely in professional training and qualification assessment. There are different opinions on whether the introduction of virtual reality technologies in training is appropriate or it is done for a beautiful visual effect and does not directly affect the educational activities.

Quality training requires effort, one of the ways to increase its effectiveness is the use of immersive technologies, namely conducting part of the practical training on virtual reality simulators. Currently, virtual reality technology is in demand not only in the entertainment industry: cinematography, games, attractions. In education, immersive technologies have been successfully used for many years to improve quality, economic efficiency, and accessibility [22–24]. The possibilities of virtual reality are quite wide, allowing to transfer the user to various locations and interactive scenarios. The advantages are many, so the possibility of starting the technological process again, without the loss of resources and the threat of real breakdowns in production, makes the use of virtual reality in training appropriate and necessary. Visibility, safety, immersion and involvement make virtual reality systems interesting and attractive for use both for trainees and teachers [25, 26].

Thus, the training of turners and millers on universal machine tools for industrial enterprises is a rather labor-intensive and economically costly process. This makes the present project to develop and implement a simulator for turning and milling in a virtual reality environment relevant.

This chapter describes the results of the development and implementation of three-dimensional computer simulator of turning and milling in virtual reality on the basis of the proposed methodology.

Problem Statement. The aim of the work is to increase the effectiveness of the educational process based on the use of virtual reality simulator, minimizing the cost of consumables in the process of training and assessment of qualifications, the availability of equipment. To achieve this goal in the development of the simulator, it is necessary to solve a number of tasks:

- perform an analysis of existing methods and means of developing virtual reality simulators;
- analyze similar developed systems and their impact on efficiency;
- to develop the concept of work, terms of reference, user scenario, UX/UI of the virtual reality simulator;
- formulate requirements for the system under development;
- develop the structure of the system and the algorithms of its components;

- design and create a database to store user and system information for the virtual reality simulator;
- conduct testing and perform comprehensive debugging of the system [27];
- develop technical documentation of the virtual reality simulator;
- evaluate the effectiveness of the developed virtual reality simulator.

Initial data. Let us formulate the functional requirements for the virtual reality system of training on milling and lathe universal machines. After analyzing similar systems, developed the concept of work, terms of reference, user scenario, UX/UI of virtual reality simulator, let us formulate the functional requirements for the system of virtual reality training on milling and turning universal machines.

The virtual reality application is to be implemented as a training and qualification assessment simulator for the field of "Mechanical Engineering". The purpose of the simulator "Laboratory of Universal Turning and Milling Machining in Virtual Reality (VR) Environment":

- increase the efficiency of the educational process;
- minimize the costs of consumables in the process of training and assessment of qualifications;
- ensure the availability of the equipment.

Let us formulate the requirements for the system:

- to the structure and functioning: stable and reliable functioning of the system modules to ensure the full performance of processes of uninterrupted power supply of technical means, the use of licensed software, time standards for service maintenance of software and hardware, protection of information from malicious programs;
- to the number and qualifications of the system staff and its mode of operation: to work with the software requires a computer operator with knowledge in the field of virtual reality simulators and the end user. If the required skills are available, the end user can work as a Computer Operator. The task of the Computer Operator is to maintain the operability of the hardware, install and maintain the operability of the software and other related programs, install programs and work with data, including backups;
- to reliability: the virtual reality simulator should automatically perform a check on the correctness of the data input, as well as on the compliance with the acceptable data ranges, including the handling of exceptional situations;
- to ergonomics and technical aesthetics: the virtual reality simulator must have an aesthetically pleasing interface with good object visualization (UI), as well as a user-friendly intuitive interface (UX);
- to the dialog mode: in case of errors in the software complex modules, error messages with recommendations for their elimination should be displayed;
- to protection of information from unauthorized access: access to the virtual reality application is provided by login and password and is delimited by administrator and user modes. Access to the application is carried out by a license key. The password has character length and naming requirements (from 3);

- to patent purity: conditions of license agreements shall be observed for all technical and software objects of the virtual reality simulator and ensured patent purity [28];
- to standardization and unification: when developing a virtual reality simulator, a standard methodology of information modeling must be used (GOST 34.601–90—Information Technology. Set of standards for automated systems), (GOST 19.102.77—Unified system of program documentation. Stages of development).

## 2 Methodology for the Development and Implementation of Simulators for the Selected Subject Area of Turning and Milling Development

The aim is to develop a simulator to improve the economic efficiency of educational processes and to solve other applied problems of the field of mechanical engineering.

Analysis and development of requirements. Performing the analysis of the subject area to justify the feasibility of development: defining the goals, issues, object and subject of research, as well as the relevance and objectives necessary to achieve the formulated goal. Further on the basis of the analysis the requirements to the simulator are created.

Designing. On the basis of results of the analysis and the set requirements to the simulator designing of the software product, including construction of architecture of the appendix, algorithms for realization of performance of demanded functions of simulator is carried out. Concept art is developed and sketches are drawn on the basis of the chosen references for visualization of the environment and models of the simulator under development.

Development. On the basis of the developed model of the simulator, 3D models of the environment and interaction objects used in the simulator are created. During modeling, program code is developed taking into account the proposed architecture and algorithms. In the process of realization of development blocks models are introduced in the development environment and united with programs for the stage-by-stage formation of the final product.

Testing. After the development stage of the simulator, various types of testing are conducted to check the fulfillment of the specified requirements for the product and to identify errors arising for the user of the simulator. Based on test results, the simulator is refined if necessary to prepare the project for commissioning for real users.

Implementation and maintenance. The final product (simulator) is installed on the computers and handed over to the customer for use. After installation, training of the customer's staff is organized to familiarize users with the management and various types of interactions with the system. If difficulties arise that were not identified during the testing phase, the developers fix the identified problems. Updates to the simulator are developed in the process of maintenance, if this process is defined in the requirements of the project.

The following describes the results of the project for the development and implementation of a 3D computer simulator for turning and milling in virtual reality based on the proposed methodology, from the design stage to the implementation and maintenance stage.

## 3 Description of the Design Process of the Virtual Reality Turning and Milling Simulator and the Architecture of the Hardware-Software Complex

The simulator design process includes the following steps:

1. analysis and immersion in the subject area;
2. development of the terms of reference with the description of the content processes, scenario and models;
3. development of sketches, schemes of screens with transitions, UI/UX design;
4. development of the system architecture.

The developed virtual reality system for milling and lathe machines has the following structure (see Fig. 1).

The resources who perform the work are teachers of special disciplines, students of the "Production and Engineering Technology" sector and production specialists working on universal machines.

In the process of activity, they work with objects that are used and transformed to produce a result. These include universal milling and turning machines, blanks for making parts on the machines. The activities performed by the user in the process of working on the universal machine tools. Parts are controlled and monitored on

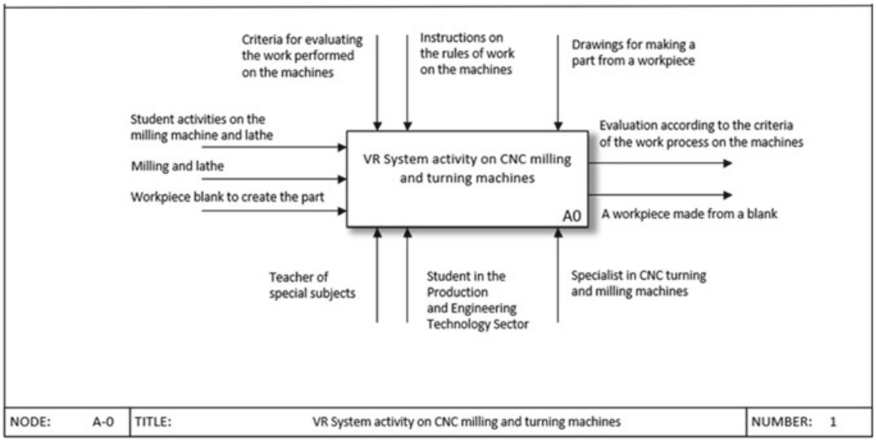



**Fig. 1** IDEF0-diagram of virtual reality system functioning for milling and turning machines

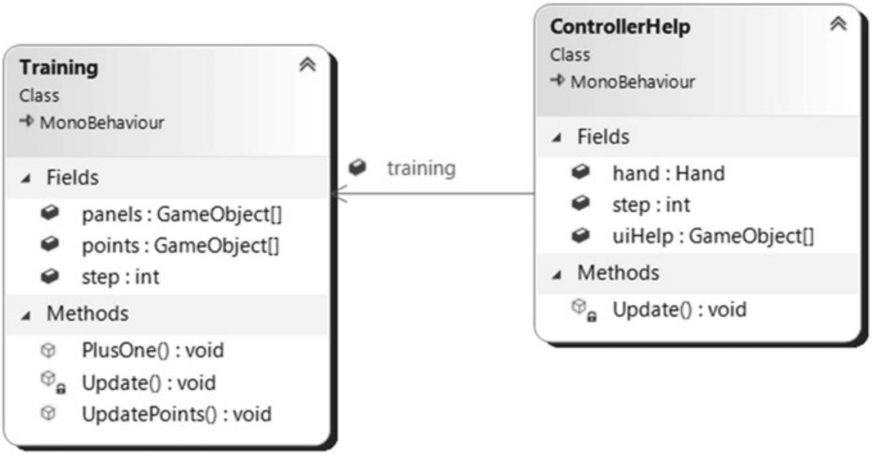



**Fig. 2** Classrooms for milling machine training

the machines by a special system, which monitors the machine tool process against specific criteria, based on the machine tool operating instructions and drawings of the parts to be produced.

These activities result in parts being produced on general-purpose machines with process quality scores based on specified criteria. The processes are indicated in the context diagram of the BP System Activity for work on milling and lathe universal machines.

Application architecture. Milling machine code architecture:

1. The classes responsible for teaching the user how to use the milling machine are shown in Fig. 2

   a. Training—the class responsible for training.
   b. ControllerHelp—the class responsible for help on the milling machine.

2. Classes responsible for interaction with the user interface of the milling machine and its operation. Below the graphic illustrations of the developed classes are omitted, only their descriptions are given

   a. MillingManager—a class responsible for the state of work of the milling machine.
   b. LockDoorButton—a class that is responsible for closing the door in the milling.
   c. ControllerTakerMilling—the class that is responsible for finding the workpiece in the milling machine.
   d. SwitchPos—the class responsible for the movement of the milling cutter.
   e. Generation—the class responsible for the generation of a workpiece.
   f. BindMilling—the class responsible for the interface of the milling machine.

   g. HandController—the class responsible for tracking the presses on the controller.
   h. SliderUp—the class responsible for the operation of the slider interface on the milling machine.
   i. SheetSwitch—the class responsible for changing the screen on the workpiece selection interface.
   j. ControllerToolMilling—the class responsible for the operation of the tool of the milling machine.

3. The classes responsible for tool operation, calculating the deflection distance of the cutter

   a. Jupka—the class responsible for tool work, measuring deviation.
   b. The other classes are needed to correct deviation from different sides.

4. Classes responsible for machining the part with the cutter

   a. ControllerCube—the class responsible for the dimension of the workpiece.
   b. MeshStudy—the class responsible for machining the workpiece.

Turning machine code architecture:

1. Classes responsible for the operation of the lathe

   a. ControllerLathe—the class responsible for the state of the lathe's work and storing the data about the lathe's settings.
   b. ControllerSupport—the class responsible for the operation of the slide of the lathe.
   c. ControllerSupportMove—the class responsible for the movement of the slide.
   d. Taker—the class responsible for the location of the part in the machine.
   e. ToolLathe—the class responsible for the tool of the lathe.
   f. ControllerDisplaySpeed—the class responsible for displaying the machine settings.
   g. TakeCutter—the class responsible for finding the tool in the lathe.
   h. ButtonCoordController—the class responsible for changing the coordinates of the slide on the display.
   i. CylPartController—the class responsible for setting the workpiece of the lathe.

2. The classes responsible for the work of machining parts on the lathe

   a. InsideTool—the class responsible for the processing tool inside the workpiece with the tailstock.
   b. InsideProcessing—the class responsible for the contact between the tool, for inside machining, and the workpiece.
   c. ControllerInsideProcessing—the class responsible for the processing inside the part.
   d. CylindController—the class responsible for the processing of the part.

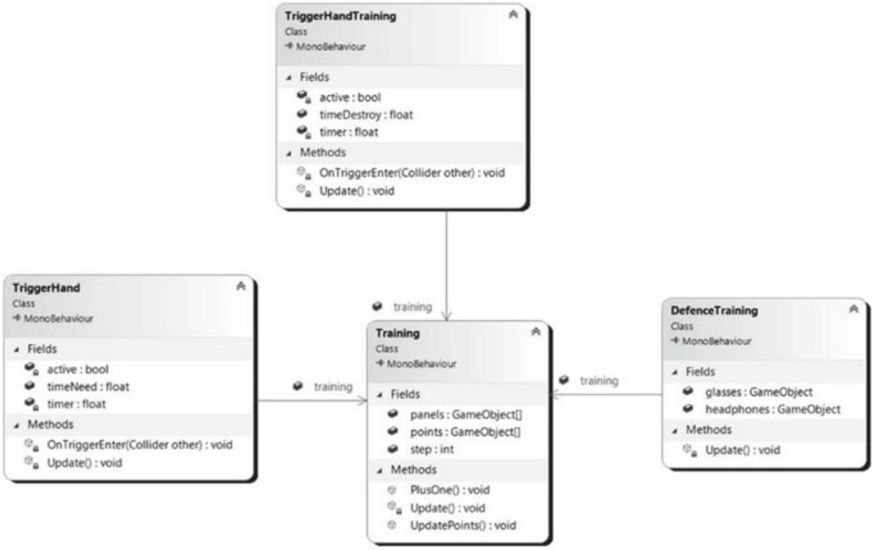



**Fig. 3** Classes responsible for learning

e. CylCamsController—the class responsible for finding cameras for saving the part as images.
f. BackForwardCamScreenshot—the class responsible for the work of the front and back cameras which store the part.

The architecture of the code of the general functionality:

1. The classes responsible for training are shown in Fig. 3.

   a. Training—the class responsible for changing the training panel.
   b. TriggerHand—a class responsible for tracking the execution of cues, through hand tracking.
   c. TriggerHandTraining—the class responsible for tracking cues through index finger tracking.
   d. DefenceTraining—a class that tracks putting on protection.

2. Classes responsible for the slider interface

   a. HandController—the class responsible for controllers pressing.
   b. SliderControllerTrigger—the class responsible for slider changes.

3. Classes responsible for the player's movement and laser

   a. Move—a class that moves the player smoothly.
   b. MoveController—The class that selects the type of movement.
   c. Laser—the class that controls the laser.

   d. SceneManage—the class responsible for the scene change.

4. The classes responsible for the error system
   a. ErrocChecker—a class that monitors errors.
   b. ErrorManager—the class which warns of an error.

5. The classes responsible for the micrometer
   a. MekrometrTrigger—the class responsible for the measurement.
   b. Micrometr—the class that handles measurements.

6. The classes responsible for saving the part
   a. SaveScrin—the class responsible for saving the image of the part.
   b. CylindrController—the class responsible for the part of the lathe.

## 4 Development of the Functional Support of the Software and Hardware Complex of the Stand

The main functionality of the system is to implement two modes: training processes on universal machine tools and execution of production tasks. The system, in its turn, controls the fulfillment of tasks by the Executor from switching on the universal machine until the part is produced. In case of incorrect actions of the Executor, the System displays the corresponding error notifications. Subsequently, the Observer can view the performed work and carry out measuring actions to evaluate the quality of the performed work. Then a report is formed and a grade for the executed work is given. Demonstration of the sequence of states and actions of the system with a description of the main functions of the virtual reality system on work on milling and lathe universal machines is shown in the activity diagram in Fig. 4.

The following support systems are used to solve the required tasks:

1. Mathematical (algorithmic) support—a set of mathematical methods, models and algorithms of information processing, used for the formation of management impacts on universal machines, as well as in solving problems to calculate the quality indicators of the process in the manufacture of parts.
2. Software—is a set of programs required to implement the functionality of the virtual environment.
3. Information software is a system of classification and coding of information, databases and data arrays, which are necessary to perform the functions realized in the project.

The proposed functional structure of the virtual reality system for milling and lathe universal machines is shown in Fig. 5.

The Virtual Reality Turning and Milling Simulator presents an application with two modes: exam and job training with pop-up prompts. The user specifies the data

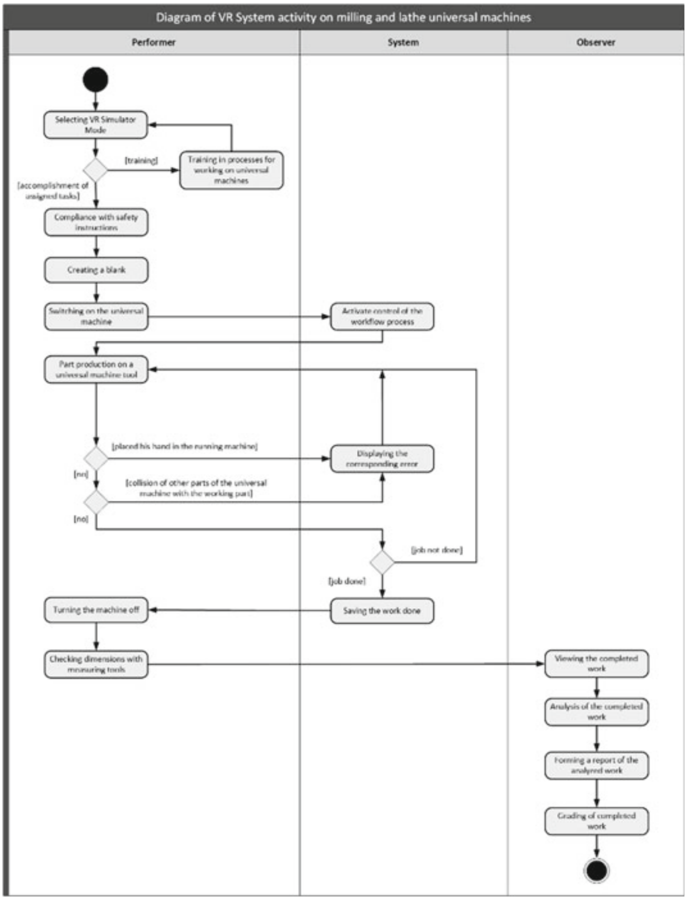



**Fig. 4** Diagram of the VR system's activity on milling and turning universal machines

of a three-dimensional workpiece, which will be further processed on the machines and measured using a special measuring tool.

Here the 3D models for the simulator scene are developed according to generally accepted principles, which are reflected below:

1. Development of the model concept, "sketch" sketch and design engineering in Adobe Photoshop.
2. Model development in specialized software Autodesk 3ds Max, Blender, Maya, ZBrush.
3. Retopology, model optimization, polygon reduction.
4. Creation of a sweep of all parts of the mesh for the correct texture application.
5. Fusing maps in Substance 3D Painter.
6. Texturing the model in Substance 3D Painter.
7. Animation and bone animation for some models including rigging and skinning for characters.

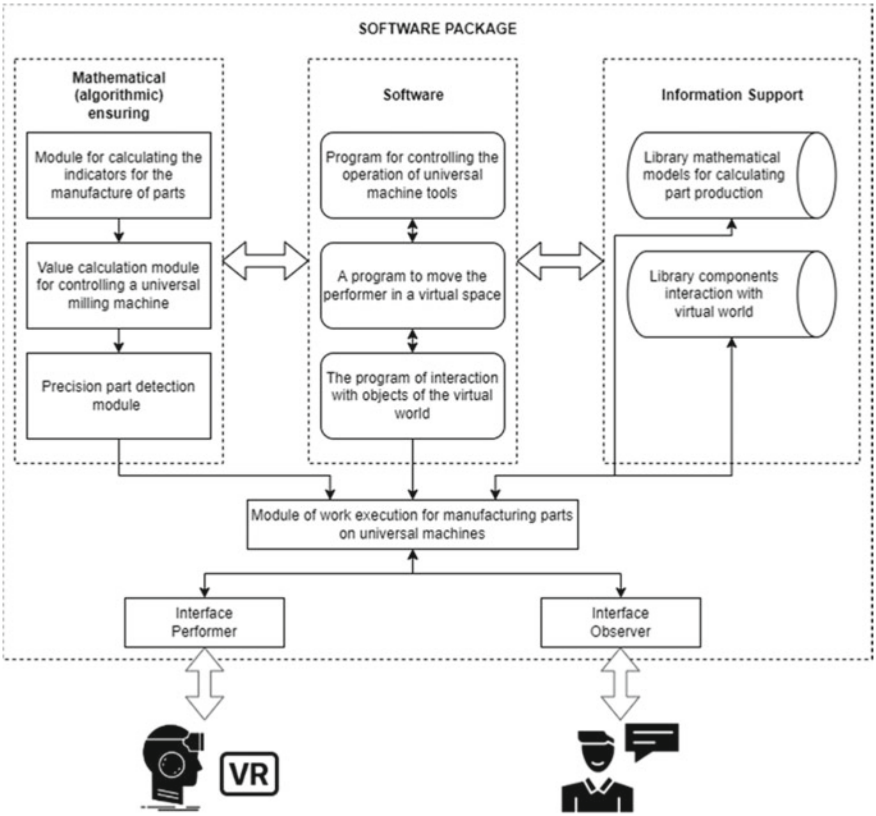



**Fig. 5** Functional structure of the VR System for milling and lathe work on universal machines

Optimization of the VR-application. For the correct operation of the application in the virtual reality environment, the work was carried out according to the following principles:

1. level geometry was optimized—the total volume of models on the scene does not exceed 100 000 polygons;
2. optimized static lighting—to reduce the load the light maps of static objects were calculated;
3. OcclusionCulling—only 3D models which fall within the user's line of sight are developed in the program;
4. ReflectionProbes—used for realistic reflection of objects on the scene;
5. Using GPU-Instancing for mass materials of 3D objects on stage;
6. Using Batching Calls—participation of the video card to reduce the load on the CPU.

Development tools and equipment. For the development of the simulator the cross-platform development environment Unity was used with the connection of the code editor Visual Studio in C# languages, synchronization with the virtual reality equipment based on the Steam VR tool. Graphic content and visual part were developed using Autodesk 3ds Max, Blender, Maya, ZBrush, Substance 3D Painter, Adobe Photoshop.

Content management is based on HTC Vive virtual reality equipment, interaction with objects using controllers.

During the testing phase, certain hidden defects in the software were identified, and appropriate changes were made to eliminate them.

UI/UX interface description. The user interface of the application is implemented and adapted to the virtual reality technology. In the main menu the user is at a comfortable distance from the interface and uses a laser pointer to interact with it. The interface predominantly uses clear icons, minimizing the text part. The main block of the app uses the index finger to interact with the buttons for more precise interaction. Animations are used to improve the visualization of the interface. The color of the interface is in blue tones. Interaction with machines is adapted and simplified to virtual reality technologies in order to comfortably operate the lathe and milling machine, i.e., the machine setup is performed using a clickable interface rather than twists/levers (see Figs. 6, 7, 8, 9).

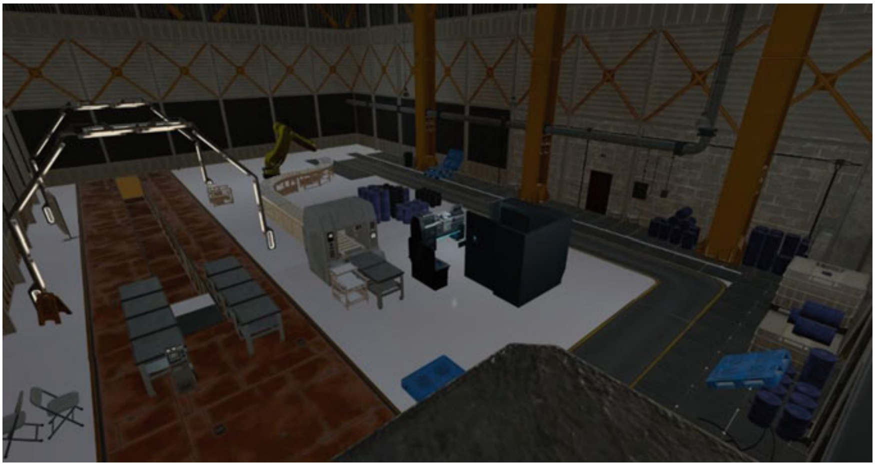


**Fig. 6** General view

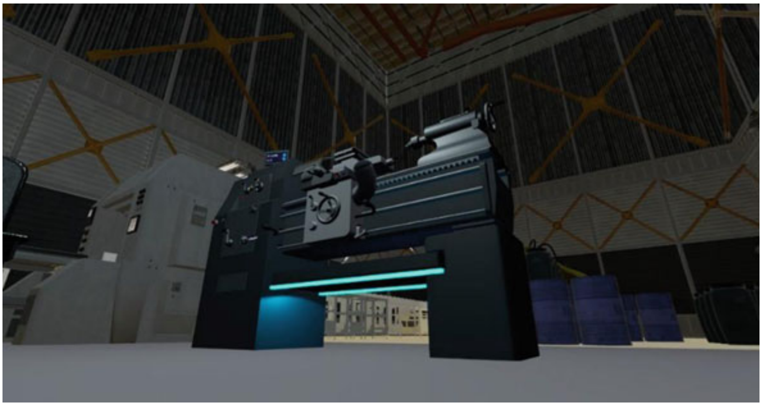


**Fig. 7** Lathe

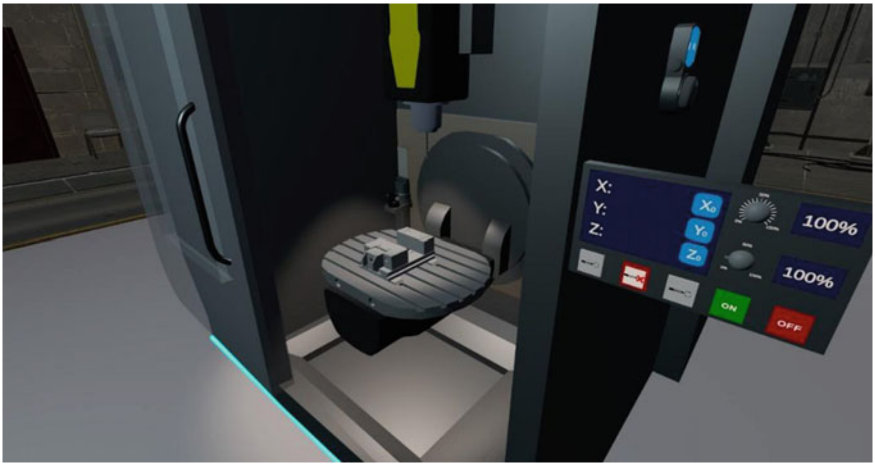



**Fig. 8** Milling machine

## 5 Conclusions

The developed automated training system for universal turning and milling in the virtual reality environment with the immersive effect has an important practical application in the implementation of effective training for industrial enterprises, optimization of physical and financial costs, minimizing the number of errors and accidents. The software product is a functional simulation of the operation of a universal lathe and milling machine with a full package of features that are identical to the original equipment.

A universal lathe with a digital display device and a tool set—cutters and drills, the software control of which provides movement of the tool holder by turning the

**Fig. 9** Measuring and hand tools

limbs due to the rotation of the controller. On this basis the movements used on the real lathe are simulated.

Universal milling machine with position control system and digital indication device with a set of cutting tools, having the mechanisms and methods of processing similar to the control with the real equipment.

The automated training system has no age restrictions on the tolerance, unlike physical universal milling and turning machines, which belong to the category of increased danger.

An important factor of economic efficiency is the absence of the possibility of damage to the equipment, the need for a special equipped room, depreciation accrual, maintenance costs and consumables: cutters, cutters, drills and workpieces. It should be noted that during the training practice “PM.03 Development of technological processes for the assembly of assemblies and products in mechanical assembly production” for one student the workpiece in the form of “rolled steel” with a diameter of 60 m and a length of 2 m is consumed. After the implementation of VR-simulator it was possible to reduce the consumable workpiece from 2 to 0.5 m. Thus, it was possible to reduce the expendable billets by 4 times. The variability of the scenario allows a diverse range of operations to be performed in the application.

Automated training system for universal turning and milling in virtual reality environment has been implemented and is used in the State Budgetary Professional Educational Institution of the City of Moscow “Moscow State Educational Complex”.

The novelty of the proposed project is:

- a model of professional training and qualification assessment capable of evaluating students’ work processes and professional skills;
- a specialized software package with unique program modules;

- method of optimizing program modules and three-dimensional models of the program complex of industrial segment training in a single virtual space.

## References


1. Qingyang, I., Qian, L., Ziwei, L., Shen, J.: Virtual reality or video-based self-instruction: comparing the learning outcomes of cardiopulmonary resuscitation training. Bull. S. Ural State Univ. Ser. Educ. Health Phys. Cult. **13**(2), 53–62 (2021)
2. Collaço, E., Kira, E., Sallaberry, L.H., Queiroz, A.C.M., Machado, M.A.A.M., Crivello Jr, O., Tori, R.: Immersion and haptic feedback impacts on dental anesthesia technical skills virtual reality training. J. Dent. Educ. **85**(4), 589–598 (2021). https://doi.org/10.1002/jdd.12503
3. Labedan, P., Darodes-De-Tailly, N., Dehais, F., Peysakhovich, V.: Virtual reality for pilot training: study of cardiac activity. In: VISIGRAPP 2021—Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, vol. 2, pp. 81–88 (2021)
4. Angelov, V., Petkov, E., Shipkovenski, G., Kalushkov, T.: Modern virtual reality headsets. In: HORA 2020—2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications Proceedings (2020). https://doi.org/10.1109/HORA49412.2020.9152604
5. Kwon, K., Park, J.S., Shin, B.-S.: Virtual anatomical and endoscopic exploration method of internal human body for training simulator. J. Korean Med. Sci. **35**(12), e90 (2020). https://doi.org/10.3346/JKMS.2020.35.E90
6. Kandi, V.R., Brittle, P., Castronovo, F., Gaedicke, C.: Application of a virtual reality educational game to improve design review skills. In: Construction Research Congress 2020: Project Management and Controls, Materials, and Contracts—Selected Papers from the Construction Research Congress 2020, pp. 545–554 (2020)
7. Dozorcev, V.M.: Newindustrial safety challengeswill computer simulators help? Bezopasnost'truda V Promyshlennosti **9**, 31–38 (2019)
8. Petrov, D., Chistyakova, T., Charykov, N.: Computer system of visual modeling in design and research of processes of carbon nanocluster compounds synthesis. In: Kravets, A.G. et al. (eds.) Cyber-Physical Systems: Design and Application for Industry 4.0, Studies in Systems, Decision and Control, vol. 342, pp. 181–195. Springer Nature Switzerland AG (2021). https://doi.org/10.1007/978-3-030-66081-9_14
9. Ershova, O., Chistyakova, T.: Methods and technologies of developing computer simulators for training resource-saving control of electrotechnological installations. In: Kravets, A.G. et al. (eds.) Society 5.0: Human-Centered Society Challenges and Solutions, Studies in Systems, Decision and Control, vol. 416, pp. 351–364. Springer Nature Switzerland AG (2022). https://doi.org/10.1007/978-3-030-95112-2_29
10. Chistyakova, T.B., Novozhilova, I.V.: Intelligence computer simulators for elearning of specialists of innovative industrial enterprises. In: 2016 XIX IEEE International Conference on Soft Computing and Measurements (SCM), St. Petersburg, pp. 329–332 (2016). https://doi.org/10.1109/SCM.2016.7519772
11. Kamada, Y.: Recent trends of dynamic simulator applications and prospects for omegaland. Yokogawa. Tech. Rep. (English edition), vol. 56, issue no. 1 (2013)
12. Kohlert, M., Konig, A.: Advanced multi-sensory process data analysis and on-line evaluation by innovative human-machine-based process monitoring and control for yield optimization in polymer film industry. tm—Technisches Messen **83**(9), 1953–1961 (2016). https://doi.org/10.3233/978-1-61499-105-2-1953
13. van den Berg, M., Hartmann, T., de Graaf, R.: Supporting design reviews with pre-meeting virtual reality environments. J. Inf. Technol. Constr. **22**, 305–321 (2017)

14. Bouchlaghem, D., Shang, H., Whyte, J., Ganah, A.: Visualisation in architecture, engineering and construction (AEC). Autom. Constr. **14**(3), 287–295. (2005). https://doi.org/10.1016/j.autcon.2004.08.012
15. Mäkinen, H., Haavisto, E., Havola, S., Koivisto, J.-M.: User experiences of virtual reality technologies for healthcare in learning: an integrative review. Behav. Inf. Technol. **41**, 1–17 (2020). https://doi.org/10.1080/0144929X.2020.1788162
16. Alhalabi, W.: Virtual reality systems enhance students' achievements in engineering education. Behav. Inf. Technol. **38**(2), 216 (2019). https://doi.org/10.1080/0144929X.2018.1519895
17. Selzer, M.N., Gazcon, N.F., Larrea, M.L.: Effects of virtual presence and learning outcome using low-end virtual reality systems. Displays **59**, 9–15 (2019). https://doi.org/10.1016/j.displa.2019.04.002
18. Bracq, M.-S., Michinov, E., Arnaldi, B., Caillaud, B., Gibaud, B., Gouranton, V., Jannin, P.: Learning procedural skills with a virtual reality simulator: an acceptability study. Nurse Educ. Today **79**, 153–160 (2019). https://doi.org/10.1016/j.nedt.2019.05.026
19. Butt, A.L., Kardong-Edgren, S., Ellertson, A.: Using game-based virtual reality with haptics for skill acquisition. Clin. Simul. Nurs. **16**, 25–32 (2018). https://doi.org/10.1016/j.ecns.2017.09.010
20. Kern, F., Winter, C., Gall, D., Kathner, I., Pauli, P., Latoschik, M.E.: Immersive virtual reality and gamification within procedurally generated environments to increase motivation during gait rehabilitation. In: 26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019—Proceedings (2019).https://doi.org/10.1109/VR.2019.8797828
21. McCarthy, C., Lai, T.D., Favilla, S., Sly, D.: Towards an immersive auditory-based journey planner for the visually impaired. In: ACM International Conference Proceeding Series, pp. 387–391 (2019). https://doi.org/10.1145/3369457.3369499
22. Riegler, A., Riener, A., Holzmann, C.: Virtual reality driving simulator for user studies on automated driving. In: Adjunct Proceedings—11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019, pp. 502–507 (2019). https://doi.org/10.1145/3349263.3349595
23. Dawley, L., Dede, C.: Situated learning in virtual worlds and immersive simulated. In: Spector, M.J. et al. (eds.) Handbook of Research on Educational Communications and Technology, 4th edn, pp. 723–734. Springer, New York (2017). https://doi.org/10.1007/978-1-4614-3185-5_58
24. Mmaki, J., Trevor, M., Ronel, M.: Experiential learning through virtual and augmented reality in higher education. In: ICETM '18: Proceedings of the 2018 International Conference on Education Technology Management (ICETM 2018), pp. 42–45. Association for Computing Machinery (2022). https://drive.google.com/file/d/1B5iDxLP4FFMudwGGXjzBCmjcMUsyVmCY/view?usp=sharing (or) https://doi.org/10.1145/3300942.3300956
25. Bolshakov, A.A., Klyuchikov, A.V., Kovylov, N.V.: Building a system architecture for displaying data in a complex of output devices. In: 2020 International Conference on Actual Problems of Electron Devices Engineering (APEDE), pp. 302–304. IEEE (2020). https://doi.org/10.1109/APEDE48864.2020.9255414. ISBN 9781728143088
26. Bolshakov, A.A., Klyuchikov, A.V.: Decision support system for selecting designs of autostereoscopic displays. In: Kravets, A.G. et al. (eds.) Cyber-Physical Systems: Design and Application for Industry 4.0. Studies in Systems, Decision and Control, vol. 342. pp. 73–88. Springer Nature Switzerland AG (2021). https://doi.org/10.1007/978-3-030-66081-9_6. ISSN 2198–4182
27. Chugunkov, I.V., Kabak, D.V., Vyunnikov, V.N., Aslanov, R.E.: Creation of datasets from open sources. In: Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), pp. 295–297. Institute of Electrical and Electronics Engineers Inc. (IEEE) (2018)
28. Aslanov, R.E., Artemiev, I.A., Frolov, A.A., Khabbatullin, R.R., Shishkin, I.N.: Laboratory of universal turning and milling in a virtual reality environment (VR). Certificate of Registration of the Computer Program № 2021660861. Request № 2021615879, Registered 02.07.2021

# Development of Software for the Organization of Training of TPP Workers on the Basis of Digital Twins of Equipment

**Vladimir Agibalov, Michael Belov, Aleksey Dolgushev, and Ivan Shcherbatov**

**Abstract** The chapter considers the task of developing software that uses digital duplicates of equipment to organize training of thermal power plant (TTP) personnel. This solution can also be used in the educational process of higher education institutions for engineering specialties. To solve the problem of implementing a prototype software developed mathematical and algorithmic support for the construction of mathematical models of various pieces of equipment for use in the digital twins of thermal power plant equipment and thermal schemes. In the course of software development the following mathematical models have been created: steam turbine cylinder model, surface-type low-pressure heater model, condenser model. Algorithmic support for building models was synthesized for use in plant personnel training systems and organization of educational processes in institutions of higher education for engineering specialties. The simulation of negative phenomena in accordance with a pre-designed scenario allows the educational process to be implemented with high efficiency. The efficiency is conditioned not only by the possibility of generation of different equipment operation modes, but also by taking into account the influence of different negative factors on the operability. The tools to create scenarios of abnormal situations: fixing the values of certain parameters, single-step, step change of parameters, creating a linear function of change of physical parameter of the object in time. The programming language for realization of the simulator software is chosen.

V. Agibalov (✉) · M. Belov · A. Dolgushev · I. Shcherbatov
National Research University "MPEI", 14 Krasnokazarmennaya, Moscow 111250, Russia
e-mail: AgibalovVA@mpei.ru

M. Belov
e-mail: BelovMK@mpei.ru

A. Dolgushev
e-mail: DolgushevAN@mpei.ru

I. Shcherbatov
e-mail: ShcherbatovIA@mpei.ru

 
A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_9

**Keywords** Digital twin · Thermal power plant · Simulation model · Thermal scheme · Staff training · Emergencies · Adaptation

## 1 Introduction

One of the problems of the modern energy sector is the reduction of energy efficiency of generation facilities during their operation [1]. Therefore, solving the problem of training personnel to work in non-standard, abnormal modes of equipment operation can significantly improve the energy efficiency of TPPs.

Due to the fact that in the process of operation power facilities are exposed to many negative influences, their parameters change greatly over time [2–7]. This phenomenon leads to a change in facility behavior over time, which can be confusing to the operational staff. Failure to make technically correct decisions can lead to accidents and pre-accidents [8–12].

It is too expensive to train personnel on real equipment, so it makes sense to introduce digital models that can simulate equipment operation under different modes of operation with the required accuracy.

Therefore, the task of development and application of digital simulation systems in order to create simulators, based on digital twins, for training students at educational institutions and specialists at energy facilities is becoming increasingly important.

## 2 Algorithms Development

The software for construction of digital twin equipment models is based on appropriate algorithmic and mathematical methods. Before construction of mathematical models of TPP equipment units and their implementation in software product, it is necessary to develop algorithms that would allow to implement the concept of construction of digital twins in TPP equipment when they are used as a part of the software product under development. Algorithms ensuring of the solving task consists of the following set of algorithms: A1—algorithm for building mathematical model of the object (equipment unit); A2—algorithm for building functional dependencies for digital twin of the object (equipment unit); A3—algorithm for building scheme of the technological process; A4—training scenarios development; A5—scenario realization (emergency, emergency, pre-accident situations etc.).

Generalized scheme of conceptual approach of construction and application of digital twin models as a part of software is presented in Fig. 1.

Part of the key algorithms (A1 and A2) are implemented in the preliminary phase (analysis phase), and their result is the basis for the construction of the digital twin in the software. In addition, these algorithms will be refined as the work progresses. In the current phase, they are conceptual in nature and are represented by generalized structural diagrams.

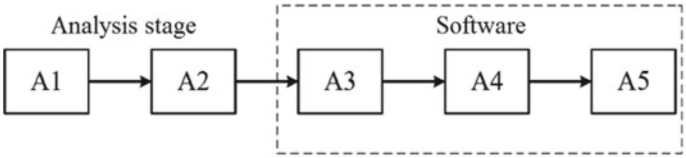



**Fig. 1** Construction and use of digital twin models

At the initial stage, it is necessary to build a mathematical model of a specific piece of equipment. For this purpose, an algorithm is synthesized (Fig. 2), which implements the classical approach for building a mathematical model of the research object in the form of thermal and material balance equations, differential equations and others.

Block 1 selects sources of information about the object, for which a mathematical model is to be built, providing the implementation of the concept of building digital twins of equipment. The sources may include both specific technological and thermal schemes of plants and other sources, which provide various descriptions of the technological process.

Based on the analysis of sources, a structural scheme (block 2.1) is selected, for which its digital copy will be implemented in the software, as well as all mathematical

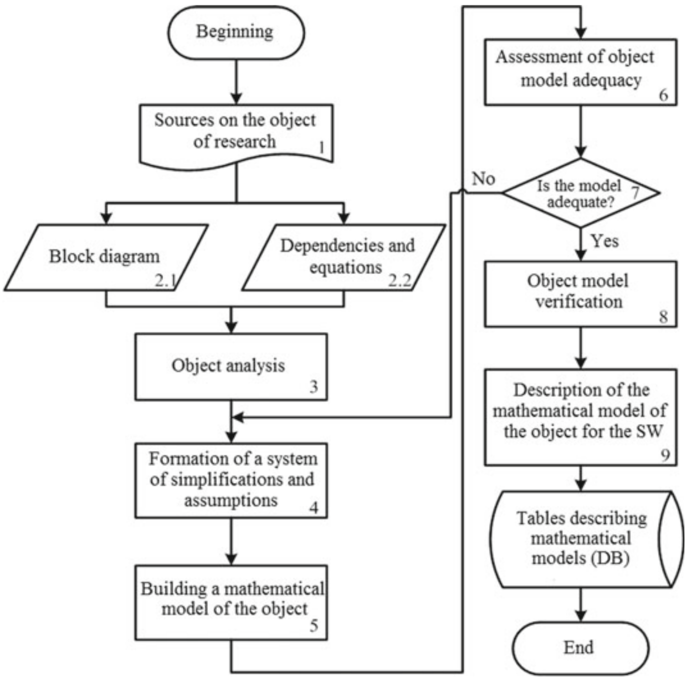



**Fig. 2** Algorithm of building a mathematical model

dependencies, energy characteristics and other functional dependencies (block 2.2), which ensure the construction of a digital twin of the object in relation to the nature of the problem to be solved.

Block 3 is designed to analyze the object in terms of the possibility of mathematical, functional or expert description of processes and objects of the thermal scheme in order that in block 4 was formed a system of simplifications and assumptions, based on which the construction of the object model, which will later be implemented in the software product and will be the mathematical basis of the digital twin.

In block 5, based on the adopted system of simplifications and assumptions, the available scheme, as well as mathematical and functional dependencies, the mathematical model of the object is built. Based on classical statistical methods (using Fisher's criterion, the approach will be described further during the stages of model adequacy assessment), the adequacy of the constructed mathematical model is assessed. If the model is not adequate to the object described (check in block 7), the system of simplifications and assumptions as well as the dependencies included in the model are checked (transfer to block 4). Otherwise, the model is verified and described in the required format, which can easily be implemented in a programming language.

Methods used in the blocks of the synthesized algorithm: block 1—selection of sources on the topic of research; blocks 2–3—textual analysis of selected sources; block 4—textual analysis, expert survey; block 5—methods for building mathematical models, thermal and material balance equations, differential equations, energy characteristics; block 6—statistical methods for assessing adequacy (evaluation based on Fisher criterion); block 8—evaluation of the root mean square error, accuracy of model calculation; block 9—building a description based on the template generated.

This chapter adopts the following format (template) of the mathematical model of the object, which will be the result of the described algorithm A1:

(1) input parameters: $X1$ - …; $X2$ - …; …; $Xn$ - …
(2) output parameters: $Y1$ - …; …; $Ym$ - …
(3) to calculate the parameter $Y1$ we use the following mathematical dependence: $Y1 = \mathrm{f}(X1, \ldots., Xn)$.

Based on the described algorithm, models of equipment units of the software prototype under development will be built in the future.

## 3 Development of Mathematical Models for the Personnel Simulator

As the power system contains a set of separate elements connected in a certain way, the mathematical model of the system must reproduce all the relations and connections within the object, concerning the relationship of all elements or selected groups of elements, considered in this case as subsystems [13].

At modeling of power systems, a distinction is made between cases, when similarity is established for all elements influencing studied functions, which appear both in time and space (full similarity), and cases, when similarity is established only for a part of processes or studied system functions (partial similarity), for example, when change of process parameters only in time is studied without considering corresponding changes in space [14].

Full similarity and, accordingly, full simulation of power systems is implemented mainly in the study of systems or individual elements, the action of which is significantly related to the propagation of electromagnetic energy in space (design and study of such system elements as electrical machines, transformers, waveguides, long power transmission lines, etc.). Incomplete modelling is usually implemented in the study of power system modes [15–17].

Mathematical modeling of power systems is practically realized by making the system of equations adapted for solving on computer, represented in the form of algorithms and programs, by means of which the numerical characteristics of processes (in the form of graph or table), occurring in studied power system are received on computer [17].

Mathematical modeling of power systems is widely used in simulators and simulations, which is especially important in the study of processes and personnel training [18].

The main advantage of using simulation in simulators is that it is possible to obtain the necessary data about the object of study, process or phenomenon without working with it directly [19]. This not only saves time, financial and labor costs, but can also help to protect the subject or the people working on it from unnecessary risks.

When developing a simulator, one can incorporate a structure for further functionality expansion (scalability of the virtual simulator), create a system of virtual system simulation (a system of interdependent objects), functionality for prompt substitution of elements of the "investigated" object, design (download files from an external media, without changing the program running the simulator) [20].

Thus, the simulator allows simulating different scenarios, such as, for example, the scenario of normal operation of equipment or the emergence of emergency situations. Moreover, using data on possible abnormal situations, when running the simulation in automatic mode, the simulator program can create a unique scenario for passing the training.

Since the simulator performs an educational function, the question arises as to how to evaluate the actions of personnel in emergency situations. Criteria must be selected for evaluating the actions of personnel during emergencies. In our simulator we propose the following criteria: time of operator's actions; number of mistakes made by the operator; determination of the cause of an emergency situation.

At the end of the training session the operator will be informed about the evaluation of his skills and/or in case of a negative experience, the consequences which resulted in his incorrect actions.

In the course of software development, mathematical models of the steam turbine cylinder, condenser and low pressure heater of the surface type (LPH) have been created. The generalized structures of condenser and LPH are shown in Figs.3 and 4 respectively.

The input parameters for the condenser are $p_S$, $t_S$—parameters of heating steam, $t_{sat}$—temperature of saturation in the heater, $p_c$—pressure of heated condensate, $t_{input}$—condensate temperature at the heater inlet, $G_c$—condensate flow rate through heater, $\theta$— underheating of water in the heater to the heating steam saturation temperature, $k$—heat transfer coefficient for heaters with u-shaped brass tubes, $d_{in}$—inner

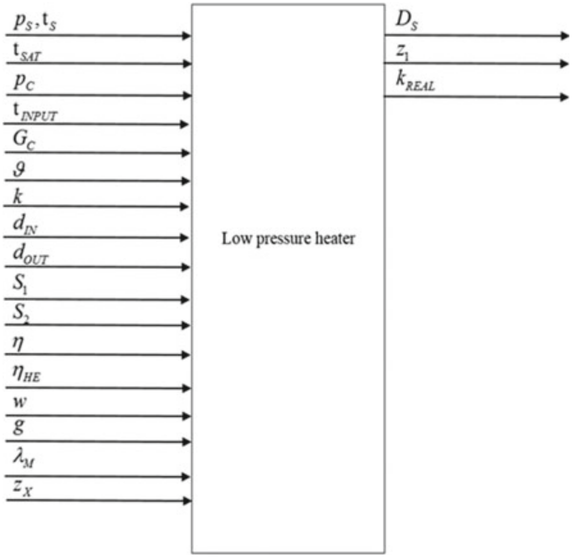



**Fig. 3** Generalized structural diagram of the low-pressure heater

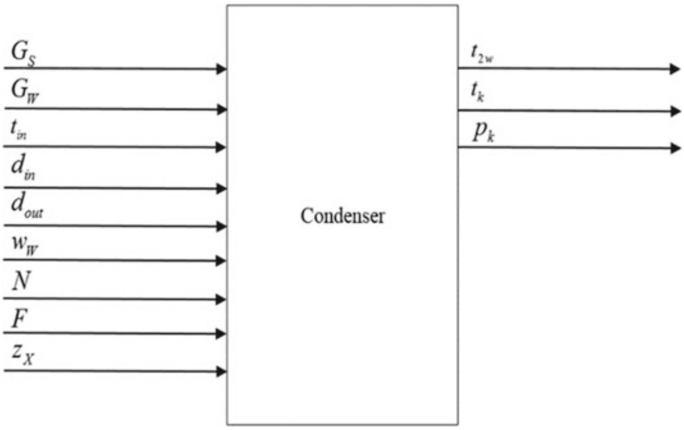



**Fig. 4** Generalized structural diagram of the condenser

diameter of u-tubes, $d_{out}$—outer diameter of u-tubes, $S_1$—staggered tube arrangement size, $S_2$—staggered tube arrangement size, $\eta$—filling factor of the tubes, $\eta_{he}$—heater efficiency, $w$—velocity of water movement in pipes, $g$—free fall acceleration, $\lambda_m$—metal heat transfer coefficient, $z_x$—number of water passes.

The output parameters for the condenser are $D_S$—flow rate of steam to the heater, $z_1$—number of pipes along the water course, $k_{real}$—true heat transfer coefficient.

The input parameters for the LPH are $G_S$—steam flow rate in the condenser, $G_w$—cooling water flow rate, $t_{in}$—water temperature at the condenser inlet, $d_{in}$—the inner diameter of the tubes, $d_{out}$—outer diameter of tubes, $w_w$—water velocity in the tubes, $N$—number of moves, $F$—heat exchange surface area, $z_x$—number of tubes.

The output parameters for the LPH are $t_{2w}$—condenser outlet water temperature, $t_k$—value of the saturation temperature in the condenser, $p_k$—saturation pressure in the condenser.

## 4 Development of Emergency Situations Simulations

The properties of the objects in the mathematical model and the relationships between them must be used as a guide for the development of the contingency simulation. For example, to create abnormal situations, we can change the parameters of an object at one time, fix certain values or create a linear function of the change of a physical value over time that will simulate the wear and tear of the equipment.

So, fixing a variable can be used to simulate the malfunctioning of a control or shut-off valve, the position of the RO stays constant (simulating a stuck valve), a burst water pipe, the water level in an element should increase and it stays constant.

Abrupt change of parameters can also describe burst pipes, breach of tightness of objects. But the application of a change in a physical quantity over time, which must remain constant, can describe metal corrosion, deposition of salts, thinning of metal, etc.

All these tools allow you to describe defects that lead to equipment failures and emergencies.

## 5 Software Implementation

The high-level programming language Python has been chosen for the software implementation of the simulator; it allows both creating an interface and creating mathematical models. Because of its versatility and ease of learning, it is perfectly suited for creating a training simulator [21]. Trainees can try to create their own contingency if they wish and, thanks to the Python maintain library function, they can load this situation into the software module without changing the main program code. Fragment of the program is shown in Fig. 5.

```python
delta = 5  # начальная мера ошибки

tk = 30 + 273.15  # в первом приближении принимаем температуру насыщения в конденсаторе равной 30°C
while delta > 0.001:
    r = HSST(tk) - HSWT(tk)  # теплота парообразования
    t2в = t1в + ((Gп * r) / (Gв * 4.19))  # температура воды на выходе из конденсатора
    delta_t = (t2в - t1в) / (log((tk - t1в) / (tk - t2в)))  # среднелогарифмическая разность температур
    tср_в = (t1в + t2в) / 2  # средняя температура в трубках
    tп = (tk + tср_в) / 2  # температура пленки конденсата на трубках

    # по температуре пленки конденсата определяем:

    # коэффициент теплопроводности
    lambda_п = 0.045 * log(1 + tп - 273.15 - 0.845) + 0.47
    # плотность плёнки конденсата
    po_п = DSWT(tп)
    # коэффициент динамической вязкости
    mu_п = (1.78 * 10 ** (-6)) / (1 + 0.0337 * (tп - 273.15) + 0.00021 * (tп - 273.15) ** 2) * DSWT(tп)
    # теплота парообразования
    delta_h_п = HSST(tп) - HSWT(tп)
    # коэффициент теплоотдачи при конденсации практически неподвижного пара на горизонтальных гладких трубках
    alfa_п = 0.728 * (
            delta_h_п * 1000 * lambda_п ** 3 * po_п ** 2 * 9.81 / (mu_п * (tk - tср_в) * dн * 10 ** (-3))) ** 0.25

    # по средней температуре воды в трубках определяем:

    # коэффициент кинематической вязкости
    nu_в = (1.78 * 10 ** (-6)) / (1 + 0.0337 * (tср_в - 273.15) + 0.00021 * (tср_в - 273.15) ** 2)
    # коэффициент теплопроводности
    lambda_в = 0.045 * log(1 + tср_в - 273.15 - 0.845) + 0.47
    # число Прандля
    Pr = 11.228 * 2.71828 ** (-0.032 * (tср_в - 273.15)) + 1.212  # Спорный момент
    # плотность воды
    po_в = DSWT(tср_в)

    Re = W * (dвн * 10 ** (-3) / (nu_в))  # число Рейнольдса
    alfa_тр = 0.023 * Re ** 0.8 * Pr ** 0.4 * (lambda_в / (dвн * 10 ** (-3)))  # коэффициент теплоотдачи
    Rст = 1.15 * ((dвн * 10 ** (-3)) / 29.4) * log10(dн / dвн)
    k = ((1 / alfa_тр) * (dн / dвн) + (1 / alfa_п) + Rст) ** (-1) # коэффициент теплопередачи с гладкими трубками
```

**Fig. 5** Mathematical model of the LPH

# 6 Conclusion

Software for training engineers was developed. Mathematical models of low-pressure heater, condenser, turbine cylinder was developed. Analysis of tools to create scenarios of abnormal and emergency situations, hidden and obvious defects of equipment was carried out. Programming language for software creation has been chosen. In the future, the developed software will become the basis for a set of training simulators for TPP personnel.

**Acknowledgements** The work was performed within the framework of the project "Development of mathematical, algorithmic and software for building simulation models of digital twins of TPP equipment and thermal circuits for use as part of diagnostic systems and predictive analytics" with the support of the grant of NRU MPEI for implementation of research programs "Power Engineering", "Electronics, Radio Engineering and IT" and "Industry 4.0 Technologies for Industry and Robotics" in 2020–2022.

# References

1. Energy Efficiency and Energy Development: State Program of the Russian Federation, Ministry of Energy of the Russian Federation, p. 235 (2016)

2. Protalinsky, O.M., Shcherbatov, I.A., Stepanov, P.V.: Identification of the actual state and entity availability forecasting in power engineering using neural-network technologies. J. Phys: Conf. Ser. **891**, 012289 (2017)
3. Protalinskiy, O.M., Shcherbatov, I.A., Khanova, A.A., Protalinskiy, I.O.: Adaptive system of forecasting the reliability of technological equipment of power facilities. Inform. Control Syst. **1**, 93–105 (2019)
4. Arakelian, E., Pashchenko, A., Shcherbatov, I., Tsurikov, G., Titov, F.: Creation of predictive analytics system for power energy objects. In: 2019 Twelfth International Conference Management of Large-Scale System Development (MLSD), pp. 1–5. IEEE (2019)
5. Shcherbatov, I.A., Turikov, G.N.: Determination of power engineering equipment's defects in predictive analytic system using machine learning algorithms. J. Phys. Conf. Ser. **1683**(4), 042056. IOP Publishing Ltd. (2020)
6. Andryushin, A., Shcherbatov, I., Dolbikova, N., Kuznetsova, A., Tsurikov, G.: Outlier detection in predictive analytics for energy equipment. In: Cyber-Physical Systems: Advances in Design & Modelling, pp. 193–203. Springer, Cham (2020)
7. Haque M.T., Kashtiban A.M.: Application of neural network in power system a review. World Acad. Sci. Eng. Technol. **6**, 53–57 (2005)
8. Voropai, R., Shcherbatov, I., Agibalov, V., Belov, M.: Repair program formation on the basis of the technical condition classifiers. In: Cyber-Physical Systems: Design and Application for Industry 4.0, pp. 107–116. Springer, Cham (2021)
9. Radomsky, V.M.M.: Search of latent defects, prediction of undesirable phenomena in technical objects and technologies. Izvestia Samara Sci. Cent. Russ. Acad. Sci. **7**, 155–162 (2005)
10. Kuzmin, V.V., Kosov, D.S., Novikov, A.L., Ivashchenko, A.V.: Failure prediction system for equipment of industrial enterprises. Reliab. Qual. Complex Syst. **3**(11), 87–89 (2015)
11. Fedorov, E.E.: Development of defect identification technique for machines and equipment. Math. Mach. Syst. **1**(2), 152–157 (2008)
12. Selcuk, S.: Predictive maintenance, its implementation and latest trends. In: Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture **231**(9), 1670–1679 (2016)
13. Tetelbaum, I.M., Schneider, J.R.: Practice of Analog Modeling of Dynamic Circuits (1987)
14. Agibalov, V., Belov, M., Dolgushev, A., Shcherbatov, I.: Building a digital capacitor model for diagnostic and predictive analytics applications. In: AIP Conference Proceedings, vol. 2402, p. 040034. AIP Publishing LLC (2021)
15. Shcherbatov, I., Agibalov, V., Dolgsuhev, A., Belov, M.: Subsystem for building a digital twin of the main and auxiliary equipment of thermal scheme of thermal power plant. In: Society 5.0: Human-Centered Society Challenges and Solutions, pp. 239–249. Springer, Cham (2022)
16. Kravets, A.G., Egunov, V.: The software cache optimization-based method for decreasing energy consumption of computational clusters. Energies **15**(20), 7509 (2022). https://doi.org/10.3390/en15207509
17. Gorushkin, V.I.: Making Energy Calculations with Computational Machines, Moscow (1962)
18. Azarjev, D.I.: Mathematical Modeling of Electrical Systems, Moscow (1962)
19. Novoseltsev, V.N.: Advantages and disadvantages of mathematical modeling. Fundam. Res. **6**, 121–122 (2004)
20. Viet, N.T., Kravets, A.G.: The new method for analyzing technology trends of smart energy asset performance management. Energies **15**(18), 6613 (2022). https://doi.org/10.3390/en15186613
21. Dymora, P., Paszkiewicz, A.: Performance analysis of selected programming languages in the context of supporting decision-making processes for Industry 4.0. Appl. Sci. **10**(23), 8521 (2020). https://doi.org/10.3390/app10238521

# Heterogeneous Information System for the Integration of Departmental Databases on the State and Development of Human Capital

**Alexandr Ivanov and Alexey Bogomolov**

**Abstract** The results of the development of the architecture of a heterogeneous information environment for the integration of information resources of enterprises that solve the problems of investing in human capital, providing the availability of up-to-date data on the state and history of the development of human capital both at the scale of enterprises and at the federal level, are presented.

**Keywords** Information environment · Human capital · Integration of information resources · Integration of databases

## 1 Introduction

The development of the economy, society and family, which combines labor resources, knowledge, tools for intellectual and managerial work, the environment and labor activity, necessitates the development and implementation of information technologies for managing the development of human capital [1–3].

The progress in the development of information technologies, the digitalization of many spheres of life determines the progress in the development and implementation of information technology technologies for managing the development of human capital [4, 5].

At the same time, it is critically important to have up-to-date data on the state and history of the development of human capital not only on the scale of enterprises (organizations, institutions), but also at the federal level, ensuring the implementation of a single information space for the formation and development of the innovation economy and the knowledge economy [6, 7].

A. Ivanov (✉) · A. Bogomolov
Federal Research Center "Computer Science and Control", Russian Academy of Sciences, 44/2 Vavylova Street, Moscow 119333, Russia
e-mail: aivanov@isa.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_10

This, in turn, requires the synthesis of a heterogeneous information environment for the integration of information resources of enterprises that solve the problems of investing in human capital [8–10].

Numerous formulations of solving the problem of creating an information basis for monitoring and managing the development of human capital are known [2–8]. Many of the tasks of human capital analysis could be solved on the basis of a comprehensive population register, which would contain information about each person, covering all aspects of his life that are of social significance.

Ideas for creating such a register were put forward as early as the 1950s, and the first decisions on its implementation at the federal level date back to the 1970s. Subsequently, such decisions were made at the highest level in various forms repeatedly, but were not fully implemented. The main reason for this is that various departments that have personalized information about the population did not agree to give it "into the wrong hands". Therefore, insurmountable obstacles of a technological, legal and financial nature always arise in the way of creating a population register.

At the same time, various departmental databases are actively developing very successfully, possessing in total almost all the information necessary for analyzing the country's human capital in various sections. For example, the Ministry of Internal Affairs has databases of vehicle owners, the Pension Fund has a database of all employees with their earnings, the tax department has databases on income, property ownership and firms, the Ministry of Health has medical databases, etc. At the same time, in most cases, mechanisms are provided for obtaining information from databases by external users (of course, if they have the appropriate rights). Moreover, in a number of cases, interdepartmental electronic interaction has been established with the transfer of the necessary information.

Thus, the absence of a comprehensive registry required for human capital management at the federal level does not mean that departments are not ready to share the information they have, but precisely that they are not ready to give it away in its entirety and want to remain its monopoly owners.

From this it is clear that in order to solve the monitoring and development of human capital, it is necessary to create a system for extracting the necessary information from a large number of different (heterogeneous) databases. This, in turn, requires the solution of a number of technological and legal issues.

## 2 Features of Organizing Data Centers

To monitor human capital, according to existing practices for assessing human development indices, it is necessary to obtain information in at least three main areas: the life expectancy index takes into account the components of health (activities to ensure the physical and mental health of society, on which human performance de pends) and demography (longevity, indicators of average life expectancy at birth, etc.) [1–4]; the education index is based on the educational component (activities for the production of intellectual and qualified resources, through the assimilation and use of acquired knowledge, skills and experience, access to education, the average expected duration of school-age children, the average duration of education of the adult population, etc.) and the cultural component (interaction with the world of material and spiritual values, the ability to understand and evaluate different forms of culture, as well as the ability to create one's own values) [1–8]; gross national income index: a decent standard of living, measured by the value of gross national income per capita at purchasing power parity [5–7]. Monitoring systems should also consider the following components: civil component—implies the fulfillment of the duties of a citizen of one's country and to act as a participant in administrative, legal and political relations. The spiritual and moral component is based on historically established, socially approved moral norms and values that ensure personal autonomy and social solidarity. The labor component is defined as the need, ability and readiness to carry out professional activities, adhere to the norms of labor relations, and are enthusiastic about their labor duties. The optimal use of the labor component in the modern world economy is determined by motivational and value aspects, it is they that lead to the mobilization of the existing labor potential.

It is obvious that the relevance of evaluating all these indices directly depends on the amount of information on which the calculations are based. Quite a lot of data has already been accumulated in the existing information systems of public and private structures, which can significantly improve the quality of calculations.

The rapid accumulation of information provides more and more opportunities for statistical and analytical research, but often the necessary information is fragmented in various heterogeneous sources, and access to its fullest volume is difficult.

The idea of solving special analytical problems by extracting information in a heterogeneous distributed environment, where several independent databases are integrated, has been relevant for more than three decades. Simultaneously with the advent of the World Wide Web, the variety of computing systems and software for building local databases and information systems has grown. In turn, this gives rise to the task of combining into a single system different in terms of technical implementation and different types of computer resources (computing resources, resources for storing and transmitting information) and conveying the total resource to the consumer.

The first attempts to create heterogeneous distributed databases were based on the unification of relatively similar database management systems (at least in terms of data organization). So, for example, back in 1985, 12 Codd's rules for relational data base management systems were formulated, which would allow building distributed databases quite easily. However, in practice, developers of relational database management systems do not fully follow such strict rules. And the task of combining local databases into a distributed one at the transaction level has become less relevant with the advent of database management systems such as ORACLE, DB2 and the like, on the one hand, and ERP systems, on the other.

Further development of methods for integrating information systems in a distributed heterogeneous environment follows the path of a service approach. In this case, each of the local systems becomes a service for servicing requests of a certain type. Distributed systems are built using gateways in the form of WEB services, using service-oriented architecture (SOA) technology, based on GRID technology, etc. In all these cases, the medium of exchange is the Internet environment, and the interaction protocols are tied to network transactions.

Virtualization and open standards allow services to go beyond the boundaries of the enterprise (organization). And communication service providers and service providers play a very significant role in blurring the boundaries between businesses and shaping global ecosystems.

When building heterogeneous distributed systems, there are acute problems of identifying similar objects when integrating independent systems in a single environment—problems of semantic certainty and semantic conflicts.

The trend of recent years is the use of messaging systems and/or electronic document management systems for linking independent databases into a heterogeneous environment. At the same time, such an exchange environment also provides the function of monitoring the execution of pending transactions.

With this approach, a special role is given to the forms of documents as the basis for interaction with databases and knowledge. Forms allow you to generate database schemas, systematize the location and processing of data in database management systems, which greatly simplifies the creation of databases and analytical research. Form is a structure of many frames, elementary portions of knowledge. Databases and knowledge are built on the basis of the structure of objects and concepts presented in database schemas; templates, layouts, display forms and procedures for processing knowledge presented in forms, as well as declarative knowledge, dictionaries and classifiers stored in database management systems.

An important aspect in the construction of distributed information systems is the problem of ensuring the security and integrity of documents circulating in such systems. Modern digital signature technologies and blockchain technologies should be kept in mind when considering these issues.

An integrated heterogeneous system can combine various systems of one industrial enterprise, government agency or industry, while the problems of identifying objects, controlling information flows and accessing data are solved quite easily due to the presence of a single administrative control.

In the overwhelming majority of cases, the methods used for database integration are not instrumental in the sense that they do not allow describing and implementing the process of including new data from some abstract database into the system. These systems are built on the basis of specific data structures, specific specialized models of objects and processes, and do not require the necessary level of access rights differentiation. In the case of working with personal data, the latter aspect becomes not only important, but legally necessary.

An important element of the information system for integrating departmental databases on the state and development of human capital is data processing centers [11]. In order for such centers to meet the needs of managers and the tasks they solve, the development of databases requires a systematic approach. The fulfillment of this requirement is achieved by choosing a methodology for building databases (in terms of composition, volume and updating of data), which would adequately take into account the needs of their users. At the same time, the information unity of the collected and stored data provides for ensuring the unambiguity of the presentation and understanding of all the data used [12, 13].

To develop and improve information support for solving the problems of human capital development, it is necessary:

- creation of a promising service of information resources for the management of human capital, determination of its structure, functions and tasks to be solved; development of a single information space (SIS), determination of its composition, as well as sources and mechanisms for filling the data exchange center, the structure and hierarchy of stored information in order to rationally organize information and analytical support for the activities of officials;
- development of a system concept for automation of management, informatization and digitalization of tasks to be solved, as well as the stages of its implementation to increase the efficiency of solving problems of human capital development;
- elaboration of issues of theory and practice of application and development of digital information technologies, clarification of their place and role in the created SIS, management systems for the development of human capital;
- clarification of the conceptual apparatus of the digitalization being implemented, the tasks and problems of implementing the SIS, as well as the functions that ensure the creation of promising regional and departmental information systems as part of the SIS for the development of human capital;
- development of modern subsystems for quality control of the activities of organizations and officials for the development of human capital;
- providing support for decision-making on managing the development of human capital;
- automating the management of the daily activities of specialists in accounting, re porting, grouping and summarizing information, planning the preparation of reporting and information, information and analytical and administrative documents.

One way or another, the solution of these tasks is impossible without the use of modern digital technologies, which involve the personal technical equipment of specialists and their workplaces, as well as representatives of various socio-professional groups of the population. This, in turn, involves the development of methods and means of storing information, Internet technologies, related software, sensors and M2M technologies—modern interaction of sensors and actuators (interaction "from machine to machine") [14–16].

M2M technologies are designed to interact or combine telecommunication and information technologies. Their application ensures the automation of various processes, as well as the creation of more advanced service packages in all areas related to the development of human capital. At the same time, there is a steady trend in the development of electronic intelligent devices that can interact with each other. This suggests that human capital development activities, which are interdisciplinary in nature, will in the near future be associated with the use of M2M technologies, corporate cloud technologies, etc.

The foregoing allows us to consider the possibility of developing and creating services for managing the development of human capital, which is understood as a set of managerial, organizational, methodological and clinical processes that ensure the coordinated actions of specialists from various organizations and automated elements of sensors and actuators (as stationary, and mobile) [17, 18].

The foregoing allows us to conclude that there is already a social order for the creation of a heterogeneous information system for integrating departmental databases on the state and development of human capital. It is assumed that in the coming years, the issue of mandatory monitoring of the quality of life of a significant part of the population with the mandatory prompt provision of the necessary data for the implementation of human capital development management at the level of organizations, regions, ministries (departments), as well as at the federal level [19–21]. An analysis of the literature, a generalization of world and domestic experience in the development of information technologies, as well as our own data, make it possible to identify the following trends: in the near future, the world is expected to massively equip organizations and social professionals with mobile devices that act as personal digital assistants containing "wired" processes, work regulations and job or professional responsibilities of specific specialists in the development of human capital.

Integrated personal electronic card—a set of electronic personal records relating to one person, collected, transferred and used by several organizations. The records included in the integrated personal electronic cards can be stored both centrally and distributed. In both cases, individual electronic personal records are accessed through a personal identification number. The issues of managing integrated personal electronic cards, storing information in it (in fact, in data exchange centers), access rights and standards for information exchange and data information compatibility (interoperability) are currently being worked out. General requirements for integrated personal electronic cards should be formulated in a separate national standard. At the same time, despite the emergence of mobile applications and gadgets, as well as the relevant experience in implementing sensor technologies in a number of sectors to ensure the socio-economic well-being of the population, there are few implementations of the latter in existing information systems, and there are practically no complex system solutions. A serious reorganization of the human capital monitoring system is proposed by many authors, firms and development organizations. The proposals include the development (creation and maintenance), experimentation, practical testing and implementation in daily activities of the following tools, [10–17]:

- digital technologies (network-centrism, data processing centers, mobility of doctors and patients, cyber-physical systems as an ideology for the interaction of sensors and actuators) and the constant development of digital services;
- devices for registration, collection and processing of personal data (introduction into practice of modern sensors, actors, controllers, etc.);
- a repository of digital models of services of the ecosystem for managing the development of human capital, as well as the stand of the chief designer; ensuring

the mobility of the population by equipping them with mobile devices and software agents connected to the "cloud" of the ecosystem for managing the development of human capital;

- implementation of the ideas of network-centrism and ensuring the rapid migration of the center for making operational decisions on the nodes of the ecosystem of services for managing the development of human capital;
- approbation and activation of the service in order to create groups of corporate SIM card users according to institutional, professional, functional, technological, territorial and/or other criteria;
- means of communication and communication (mobility of participants in the process of human capital development).

Further development of interaction and the introduction of M2M technologies, the connection of software agents and actuators to the "cloud" of the eco-system for man aging the development of human capital, the creation of a multi-agent communication platform for the development of human capital necessitate the consideration of the problem of automated collection and storage of factographic information.

Thus, the introduction of modern digital technologies will allow not only to organize work on the development of human capital at the level of world standards, but also to better use the accumulated information in order to improve the quality of life of the population at the federal level (on a national scale).

## 3 Architecture of a Heterogeneous Information System for the Integration of Departmental Databases

As a result of complex interdisciplinary research in several fields of knowledge (system analysis, computer science, jurisprudence, economics, sociology, psychology, management), the architecture of such a system, which includes five subsystems, as well as the rules for the interaction of these subsystems, has been substantiated [22–25].

1. The subsystem for input and recognition of non-textual information is designed to input and analyze non-textual information using a scientific and methodological apparatus of technical vision that implements a complex technological chain for obtaining frames from a video stream, extracting and analyzing significant information. The algorithms implemented in the construction of this subsystem provide a solution to the problems of registering a video stream containing non-text information; adaptive filtering of noise that occurs when registering a video stream; video stream recognition (object detection, object tracking, object classification); primary processing of the video stream (preparation of images for the detection of moving objects on them) [22, 25].
2. Subsystem for input and processing of text information, which provides input of text documents in paper form and import of text documents in electronic form.

3. A subsystem for automating work with documents and monitoring their execution, which provides viewing and printing of documents; recognition of graphic images of documents, including in the mode of streaming scanning; recognition of the barcode of the entered documents and the formation of a description file; report generation; morphological analysis of the text representation of the document; file con tent indexing; support for hierarchical dictionaries and directories. This subsystem should support work with a geographically distributed structure, ensuring the transfer of information between remote workplaces in an encoded form [22, 25].
4. A subsystem for operational and archival storage, which solves the problems of long-term storage of scans of original documents. This subsystem implements secure database integration using distributed registries, consistent with the Peppol project, and also supports a consistent data exchange format, a standard for encryption based on public keys, and uses a self-contained infrastructure of certificate authorities. Secure integration of databases is provided with the help of the Russian system of inter departmental electronic document management and the system of interdepartmental electronic interaction [13, 17].
5. Data and messaging subsystem that supports electronic data interchange using applications supported by electronic data interchange, an electronic data interchange interface, and a subsystem. An application supported by EDI must process transactions received through an EDI interface. The subsystem must convert EDI messages that comply with international standards (EDIFACT or X.12) into an EDI interface and vice versa, and transmit messages through the exchange network.

To access data in real time through an indirect scheme of integrated databases, a single universal query interface is implemented in the data and messaging subsystem, which allows you to extract information directly from source databases in accordance with a modular approach to software development approach.

The following levels of database integration are implemented in the data and messaging subsystem:

- data layer for which unstructured, semi-structured, hard-structured representations are allowed using [10];
- the level of standard query languages and protocols (for example, SQL, JDBC, OLE DB, etc.) [11, 12];
- a superset level of metadata dictionaries;
- data warehouse technology level;
- the level of software and user interfaces;
- level of corporate information systems;
- the level of web services, which is based on providing a standard interface for accessing data from all databases of the merged organizations using web services based on the standard protocol for exchanging structured messages in a distributed

computing environment (SOAP) and a web browser that provides access to corporate applications of different organizations, access to which is carried out and controlled by the corresponding web services [11, 22, 25].

In the data and messaging subsystem, it is necessary to implement the Z39.50 distributed resource access protocol, which supports a single hierarchical representation of distributed information resources in a client–server architecture environment and provides the user with a single interface for accessing them, providing access to data bases in terms of the SQL-language.

The proposed descriptions and characteristics of the functional subsystems of the general architecture of the heterogeneous information environment for the integration of information resources of enterprises that solve the problems of investing in human capital are new, providing the integration of large both test and non-text data.

Solving the problem of database integration for building heterogeneous information systems is an urgent task that is at the forefront of modern world science, which is confirmed by research on database integration methods when building systems from components of various architectures, including the integration of partially homogeneous databases and cloud database integration.

## 4 Conclusion

One of the applications of the results of the study is their application as part of the information support of health centers—units working in three main areas:

- medical care for the correction of risk factors for chronic non-communicable dis eases to patients;
- organizational, methodological and analytical work to coordinate the activities of similar centers at the regional level;
- field events to conduct a comprehensive survey of employees of enterprises and various organizations.

The proposed solutions ensure the interface of information systems of health centers with information systems of federal and municipal authorities—this makes it possible to take into account, when developing and implementing a set of measures to preserve the health of the population, a set of characteristics of social status, professional training, characteristics of living conditions and, accordingly, improve the quality and validity of decisions taken in the interests of human capital development.

**Acknowledgements** This work was supported by a grant from the Russian Foundation for Basic Research (project no. 19-29-07271).

## References


1. Makovskaya, N.V.: Human capital development in Belarus via in-house investment. Inf. Sci. **75**, 128–139 (2016)
2. Baranov, A.O., Slepenkova, I.M., Tagaeva, T.O.: Improvement of statistics on the re production of human capital. Stud. Russ. Econ. Dev. **31**(1), 17–23 (2020)
3. Dudukalov, E., Kuznetsova, I., Poltarykhin, A., Mandrik, N., Aleshko, R., Poletaev, S.: Modeling human capital dependence and production with a high level of automation. Int. Rev. **1**, 69–79 (2021)
4. Ushakov, I.B., Bogomolov, A.V.: Diagnostics of human functional states in priority studies of Russian physiological schools. Med.-Biol. Socio-Psychol. Probl. Saf. Emerg. Situations **3**, 91–100 (2021)
5. Asmolov, A.G., Guseltseva, M.: Education as a space of opportunities: from human capital to human potential. In: The European Proceedings of Social & Behavioural Sciences EpSBS, pp. 40–45 (2019)
6. Eskindarov, M.A., Gruzina, Yu.M., Firsova, I.A., Melnichuk, M.V.: Human capital competencies in high-tech and knowledge-intensive sectors of the economy. Econ. Soc. Changes Facts Trends Forecast **13**(6), 199–214 (2020)
7. Fomin, V.P., Igoshina, N.A.: Assessment of the intercompany relationship between business interests and human capital. Lect. Notes Netw. Syst. (LNNS) **160**, 390–397 (2021)
8. Mirgorodskaya, E.O., Zabolotnyaya, E.A., Magomadov E.M.: Competitiveness of human capital of youth: determinants and evaluation model. In: The European Proceedings of Social & Behavioural Sciences EpSBS, pp. 2299–2305 (2019)
9. Teslenko, V., Melnikov, R., Bazin, D.: Evaluation of the impact of human capital on innovation activity in Russian regions. Reg. Stud. Reg. Sci. **8**(1), 109–126 (2021)
10. Zaborovskaia, O., Nadezhina, O., Avduevskaya, E.: The impact of digitalization on the formation of human capital at the regional level. J. Open Innov. Technol. Market Complex. **6**(4), 1–24 (2020)
11. Stolyar, V.P., Krainyukov, P.E., Rybakov, Yu.L., Gukasov, V.M.: Digital medicine: theory, state, prospects for implementation and practical application. Med. High Technol. **4**, 5–16 (2018)
12. Orekhov, V.D., Prichina, O.S., Loktionova, Y.N., Yanina, O.N., Gusareva, N.B.: Scientific analysis of the happiness index in regard to the human capital development. J. Adv. Res. Dyn. Control Syst. **12**(4), 467–478 (2020)
13. Krainyukov, P.E., Stolyar, V.P. Military medicine and digital technologies: theory, practice, problems and prospects. Mil. Med. J. **340**(6), 9–19 (2019)
14. Soldatenko, S., Bogomolov, A., Ronzhin, A. Mathematical modelling of climate change and variability in the context of outdoor ergonomics. Mathematics **9**(22), 2920 (2021)
15. Bogomolov, A.V.: Information technologies of digital adaptive medicine. Inform. Autom. **20**(5), 1154–1182 (2021)
16. Ushakov, I., Dragan, S., Bogomolov, A., Soldatov, S.: Technology for predictive monitoring of the performance of cyber-physical system operators under noise conditions. In: Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, vol. 333, pp. 269–280. Springer, Cham (2021)
17. Kuandykov, M.G., Krainyukov, P.E., Stolyar, V.P., Lim, V.S.: Unified military medical information system of the medical service of the armed forces: opportunities for creation and development strategy. Mil. Med. J. **341**(12), 4–19 (2020)
18. Kharlamov, A., Kharlamova, T.: Human capital development in the digitalization risk management process. In: Atlantis Highlights inComputer Sciences. Proceedings of the International Conference on Digital Technologies in Logistics and Infrastructure (ICDTLI 2019), pp. 199–203 (2019)
19. Bubeev, Yu.A., Vladimirskiy, B.M., Ushakov, I.B., Usov, V.M., Bogomolov, A.V.: Mathematical modelling of spread COVID-19 epidemic for preventive measures to protect life and health of elderly. Bull. S. Ural State Univ. Ser. Math. Model. Program. Comput. Softw. **14**(3), 92–98 (2021)

20. Iankovets, T.: Background of human capital development in the context of forming the economy of knowledge. Eureka Soc. Humanit. **2**, 3–9 (2018)
21. Mathews, S.C., McShea, M.J., Hanley, C.L., Ravitz, A., Labrique, A.B., Cohen, A.B.: Digital health: a path to validation. Digit. Med. **2**, 38 (2019)
22. Bulatov, K., Arlazarov, V., Chernov, T., Slavin, O., Nikolaev, D.: Smart IDReader: document recognition in video stream. In: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), pp. 39–44 (2017)
23. Mkrttchian, V., Chernyshenko, S.: Digital intelligent design of avatar-based control with application to human capital management. Int. J. Hum. Capital Inf. Technol. Prof. **12**(1), 19–32 (2021)
24. Ushakov, I.B., Bogomolov, A.V.: Informatization of programs of personalized adaptation medicine. Bull. Russ. Acad. Med. Sci. **69**(5–6), 124–128 (2014)
25. Bulatov, K., Fedotova, N., Arlazarov, V. Fast approximate modelling of the next combination result for stopping the text recognition in a video. In: 25th International Conference on Pattern Recognition (ICPR), pp. 239–246 (2020)

# Society 5.0: Socio-economic Systems Modelling

# Application of Quantum-Like Mathematical Models Based on Status Functions for the Analysis of Socio-Economic Systems. Part 1. Introduction

**Irina Veshneva and Alexander Bolshakov**

**Abstract** The basic concepts of classical probability, negative probability, fuzzy sets are considered with appropriate analogies of the quantum mechanical approach to create mathematical models for describing social phenomena. It also briefly describes the features of the application of the theory of intentionality, which uses the concept of institutionality for social phenomena. At the same time, institutionality is understood as a property of consciousness directed at objects and states of affairs in the real world. The concept of collective institutionality is singled out, which forms social facts arising from the collective impact of members of society. The concept of status is given, which denotes a set of characteristic values of a subject or object that determine its position in the system. The features and scope for fuzzy models are described, which are used for the mathematical description of linguistic terms and intentionality in estimating the probabilities of events in fuzzy set theory. In addition, the concept of negative probability and the possibility of using it to describe social events are described. The following is an example illustrating the concept of a complex-valued function and its use to measure the characteristics of socio-economic processes. In conclusion, the concept of status functions is introduced for comparison with membership functions and wave functions.

**Keywords** Quantum-like mathematical model · Status function · Socio-economic system · Intentionality theory · Institutionality

I. Veshneva
Saratov State University, 83, Astrakhanskaya Street, Saratov 410012, Russia
e-mail: veshnevaiv@gmail.ru

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



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_11

## 1 Introduction

The task of creating mathematical models that describe socio-economic processes is becoming very relevant [1–6]. It is possible to single out the main aspects that provide the need for mathematical models. First, it is an increase in the speed of socio-economic processes [7–9]. At the same time, the risks of management decisions increase. Secondly, this is a significant development of information technology. On the one hand, they provide an acceleration of socio-economic processes; on the other hand, they require an increase in computing power resources to solve many problems. Thirdly, this is the emergence of synergetic as a science about the general patterns of development of living and non-living systems, which makes it possible to create new models of socio-economic systems [10–12].

At first, attempts to describe the properties of real events in human activity led to the emergence of probability theory in the seventeenth century. The analysis of such sections as probability theory, mathematical statistics, set theory led to the creation of the theory of fuzzy sets. At the same time, Fuzzy Sets allows you to introduce a measuring scale for socio-economic phenomena and regulate the rules for making decisions. It should be noted that the possibilities of using Fuzzy Sets are limited. First, a large amount of information is contained in the intuitive preferences, which are difficult to formalize, of a person who forms lists of estimated parameters and constructs membership functions. Secondly, the introduced fuzzy evaluation corridors are static.

To understand the mechanisms that regulate the behaviour of a complex socio-economic system, it is necessary to measure its state variables and mathematically model the dynamics of each of the system components [13]. The creation of models can be focused on various mathematical apparatus. One of the key issues is the complexity of considering the mutual influence of individual processes. Therefore, it is difficult to determine how to achieve the state of the overall system that is being measured. The solution to the problems of describing phenomena in socio-economic systems should not be sought in the mathematical apparatus of quantum mechanics.

Formulation of the problem. We will discuss the concepts of classical probability, negative probability, Fuzzy Sets and the quantum mechanical approach to creating mathematical models of social phenomena.

## 2 Probabilistic Models

This class of models is widely used when the factors under study are uncertain. Such situations are typical for different areas of human activity. Examples are the roll of a specific number on dice (Fig. 1), the demand for a certain product, the result of passing tests, the results of voting in an election, the political situation in a given country, etc. To understand probabilistic methods, it is useful to keep in mind that uncertainty can be of quite different nature.

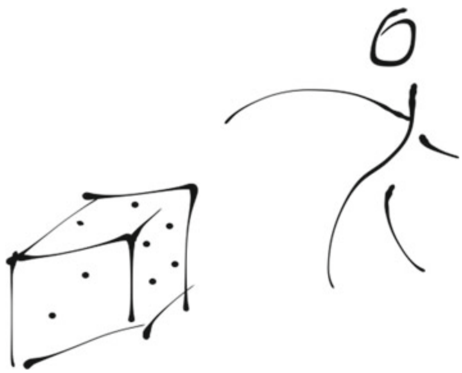


**Fig. 1** Classical probability—simple events

For classical probability, the concept of a random event is fundamental. This is an event that may or may not occur because of a particular test. In this case, the test can be both a purposeful action and a phenomenon that occurs independently of the observer. As a result of the test, one of some finite set of equally possible outcomes can occur. For example, some dice is thrown. Classical probability is defined as the ratio of the number of outcomes that favor the occurrence of events $m$ to the total number of outcomes $n$. When the number of outcomes is infinite, then the concept of geometric probability is introduced. If the numbers $m$ and $n$ are not known in advance, then new methods have to be invented.

A statistical probability is introduced for an event that may or may not occur because of an experiment. Therefore, it is possible to conduct a certain number of experiments $n$ and determine how many times the event A occurred. Then the ratio $m/n$ will be called the relative frequency of occurrence of the event $A$ in $n$ trials, it is also $\in$ [0; 1].

Subjective probability in many real situations is associated with determining the probability of events when one of the above methods is impossible to do. Then the above-mentioned understanding of probability as a measure of the reliability of a certain event comes to the fore. In this case, an expert survey should be conducted and based on its results, a subjective probability of an event should be obtained.

The earliest theory of intentionality is concerned with the principles of distinguishing between objects that exist in understanding and objects that exist in reality [14]. Currently, the concept of intentionality is being actively discussed in the humanities. For example, in [15], intentionality is understood as the power of minds to be about, to represent, or to stand for, things, properties and states of affairs.

Unlike physics or chemistry, the subject area of the social and economic sciences depends on institutional factors. This is because some properties of the world around us depend on the observer, and some do not. For example, stars and planets, the laws of their motion, their lifetime do not depend on the observer. Ownership of industrial enterprises depends on the system of social production relations. Paper money, being the paper on which it is printed, becomes a means of payment only by virtue of the legal norms regulating socio-economic processes. In the absence of appropriate agreements, money is worth no more than the paper on which it is printed. Thus, it can be assumed that the problems discussed in the humanities are dependent on conventions established by social institutions.

## 3 Institutionality

Searle in [16] introduces the concept of institutionality and distinguishes between the concepts of objectivity and subjectivity as properties of statements. Objective statements should be considered independent of the feelings and opinions of those who make or interpret such judgments. Subjective judgments, on the contrary, depend on the feelings and opinions of the participants in the discussion.

The concept of institutionality covers a range of issues related to formal social structures such as the state structure and informal ones, for example, customs, behaviour patterns, family.

The construction of social reality [17, 18] can be represented based on three simple principles: collective intentionality, attributed functions, status functions.

Collective intentionality [19] includes collective intentions, collective beliefs and desires. A person can have intentionality, that is, intentions, beliefs, and desires, such collective. Intentionality should be understood as the property of consciousness to be directed towards objects and states of things in the real world. Therefore, beliefs, desires, hopes, fears, and emotions can usually be technically characterized as intentional. The concept of collective intentionality includes other forms of intentionality, such as collective beliefs and desires. Collective intentionality shapes the social facts that emerge from the collective action of the members of society.

The assigned functions depend on the observer. People often attribute functions to objects that have no function in themselves but acquire it only through such an attribution. For example, tools perform functions based on their actual physical properties.

It is possible to assign functions to objects and phenomena that are different from their physical properties (physical structure) as real objects. Then attributing functions to them endows these objects with properties that they do not actually possess. Assigned functions that endow an object or phenomenon with properties not specified by their physical structure give objects (subjects or phenomena) a status. The status denotes a set of characteristic values inherent in the subject (or object) that determine its position in the system. Searle [18] calls the status functions performed by an object or subject because of existing institutional agreements in the system.

Assigning status functions usually has the following form: "*X* counts as *Y* in context *C*" ("*X* counts as *Y* in *C*"). In all these cases, *X* denotes some property of an object, individual, or situation, and *Y* endows the individual, object, or state of affairs with a special status. Here *C* represents the construction rules. For example, money, university degrees, marital relations, the game of chess, which cannot exist without the rules and social conventions that create them. Status functions of the form "*X* counts as *Y* in *C*" form the connecting foundation for the construction of social relations.

## 4 Fuzzy Sets

The mathematical description of linguistic statements and intentionality in the estimation of the probability of some event was developed by Zadeh [20] in the theory of fuzzy sets (Fuzzy Sets). Zadeh introduced the assumption that the characteristic function of the set can take any value in the interval [1]. This function is called the membership function for a fuzzy set, it is the basic concept for fuzzy logic. The ability to create mathematical models of exemplary human reasoning, which has the most amazing property of the ability to make the right decisions in an environment of incomplete and fuzzy information, has provided a huge increase in interest in the application of the Fuzzy Set in almost all branches of science and technology. The works of L. Zade laid the foundations for modeling human intellectual activity.

A typical example of a system that lends itself well to implementation using fuzzy logic is ABS—anti-lock braking system. There are many implementations of ABS, however, in the general case, control is carried out according to two input variables: wheel slip (the ratio of the vehicle speed to the instantaneous linear speed of a point on the outer radius of the wheel relative to its center) and the radial acceleration of the wheel. Both variables are represented as logical variables with a set of 5–8 terms each, for example, "absent", "weak", "medium", "strong", "very strong", etc., based on which the calculator, using the set rules (their number is equal to the product of the number of terms of the input variables), obtains the pressure value in the brake cylinder, striving to maintain optimal slip. A similar problem, however, is also solved by classical calculators using three-dimensional tables that describe the plane of the output value depending on the two input values.

Thus, Lotfi A. Zadeh introduced the membership function, which can be considered a weakly formalized analogue of the probability of occurrence of events. In this case, the probabilities are extended to some sets, called fuzzy ones. These sets are defined by functions and have a geometric domain.

## 5 Negative Probability

Let us complicate the considered case of occurrence or non-occurrence of the event. If we use the classical definition of the probability or frequency of the occurrence of an event when measuring a state based on some sample from the general set of events, then the probability of the occurrence of an event $p \in [0;1]$. Accordingly, the probability of the event not occurring is $1 - p$.

Could there be something more than just the non-occurrence of the event? Success, failure, anti-success: (1 0 − 1)? Could there be something more than just the non-occurrence of the event? For example, person A does not have his own opinion on a certain issue. Person B tries to convince A to accept his view of the situation but gets the opposite effect of communication. Like it's a negative probability?

If we use the classical definition of the probability or frequency of the occurrence of an event when measuring a state based on some sample from the general set of events, then the probability of the occurrence of an event is $p \in [0;1]$. Accordingly, the probability of the event not occurring is $1 - p$.

Dirac [21] associated negative probability with the amount of money. For example, he was waiting for a salary, he borrowed money. He didn't get paid and was left in debt. Then it turns out that the universal set of possible events must be expanded. Such events that create the probability of the event not occurring (is not a negation of the occurrence of the event). Or characterize the factors associated with it from the other side.

In the experiments of Belinsky [22] showed the need to introduce a negative probability. An interferometer with parametric radiation sources for two observers was used. Correlated photons are sent to two observers in two modes, one of which is delayed. Then the modes are mixed by 50% in beam splitters and are detected by four receivers, two for each of the observers. The random process is described by four dichotomous variables taking single values $\pm 1$. It is suggested that there exists a positive-definite normalized probability distribution function. For them, $2^4 = 16$ joint probabilities are formed for the discrete probability distribution function. The formation of the probability function consists of terms, half of which are negative. As a result, the existence of a positive-definite normalized probability distribution function is proved.

However, in practice the result is different. Due to the interference nature of the experiment, the probabilities on the detectors depend on the phase delays of a harmonic form. The joint probabilities determined from the experiment give a result that corresponds to the quantum description. The only explanation for the existence of the joint probability function obtained in the experiment is the non-observance of the obvious rule of positivity of probability. Belinsky A.V. considers that a negative probability means that the probability of the expected event is decreasing, while the probability of the opposite event is increasing. Negative [23] and complex-valued [24, 25] probabilities have been repeatedly discussed in the literature.

## 6 Complex Probability

The concept of negative probability was obtained in quantum mechanics. However, even in quantum mechanics, the concept of negative probability is not obvious, and negative probability is introduced as a necessity for interpreting some experiments. In this case, in fact, the probability is determined by the complex-valued wave function $\psi$, being complex. Experimentally, only the probability density $|\psi|^2$ can be

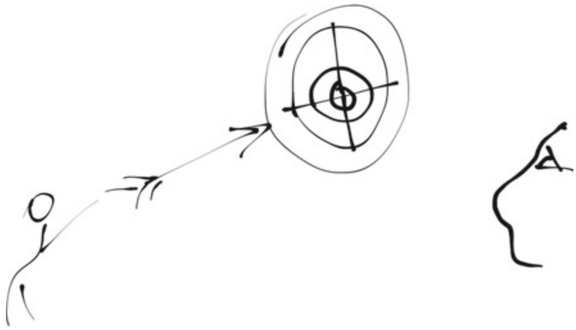


**Fig. 2** Classical probability and its perception by the observer

measured. In the orthodox understanding of quantum mechanics, negative probabilities do not exist due to the impossibility of conducting an experiment to directly measure them. Thus, the negative probability is contrary to orthodox probability theory and orthodox quantum mechanics, but not contrary to several experiments and common sense.

We will assume that the result of applying this operator to the assigned random function will allow us to calculate the intentional characteristic of the function attributed to the subject.

Let us imagine the meaning of using complex quantities to assign them to random functions of estimates. The set of possible states of the system is characterized by the introduction of random variables to describe the variables that form this set. A random variable is random in the sense of corresponding to the real state of the system. For example, let's describe shooting at a target. A random event is a hit in the center of the target. Let us assume that any random variable (Fig. 2) attributed to the state of an object (or subject) consists of the fact that it is $Z$ and its perceptions $Z'$:

$$Z = Z + Z'. \tag{1}$$

Then, to introduce a system of measurements of states, an object (subject) needs to introduce an ordered pair of real random variables $S = \{S1, S2\}$ belonging to different sets of estimates of observed "events" corresponding to the measurement of the object (subject). Here $S1$ is a random variable of the object (subject) evaluation that can be measured in practice, $S2$ reflects the perception of the event $S1$ corresponding to the imaginary part of the event.

For example, a binary value takes random values 0 and 1. Let's attribute to them the linguistic terms of the semantic interpretation "True" and "False". Consider the event of hitting the target $S1 = 1$, miss—$S1 = 0$. For the value $S2 = 1$, the event is taken as the true value of $S1$, $S2 = 0$, the event is taken as the false value of $S1$. On pairs $S = \{S1, S2\}$:

(1,1)—a hit is taken as a hit;

(1,0)—a hit is considered a miss;
(0,1)—a miss is taken as a miss;
(0,0)—a miss is treated as a hit.

Thus, a simple and visual introduction of complex random variables attributed to the state of the object (subject) is possible. Note that the functions attributed to the state of the system always have a component of perception of the observation event. Let us show the expediency of introducing complex quantities to measure the states of an object (subject).

In quantum mechanics, the problem of introducing a complex probability follows from the impossibility of determining which way the particle has passed, and the occurrence of interference. "Moving" the problem of determining the path to classical physics, you can try to determine which of the parallel conductors passed the electric current. In everyday interpretation, one can ask "how did the information that influences the decision was obtained": the Internet, banner advertising, from acquaintances, or otherwise? It is often impossible to choose one of them. Thus, the question of "what way?" may also be relevant for decision support.

Suppose the state of the object (subject) is assigned a random function $f$, which is described by the probability distribution $\rho(f)$. It is often sufficient to know the characteristic value of $f$ and the range in which its random values are most likely distributed. To get an estimate for the center value and range width, the mean $f$ and the standard deviation from the mean σ are usually used:

$$\langle f \rangle = \int f\rho(f)\mathrm{df} \Big/ \int \rho(f)\mathrm{df}, \tag{2}$$

$$\sigma = \sqrt{\langle f^2 \rangle - \langle f \rangle^2}. \tag{3}$$

An object (subject) can be in states $A$ and $B$, which are determined by the attributed functions $\psi_1$ and $\psi_2$. If these states interact, then the probability distribution of the object (subject) is described as follows:

$$P = \psi_1 + \psi_2. \tag{4}$$

Such a mathematical representation is quite consistent with many interrelated processes in the internal structure of the object (subject). To describe the behaviour of social groups in the humanities, the term "group synergy" is used. Such a mathematical description corresponds to an intuitive understanding of the processes taking place in socio-economic structures. For example, aspects of an enterprise's balanced scorecard (internal business processes, customer relationships, financial performance, and employee learning and development) that are internally related. The traditional measurement of states allows us to obtain the following probability distribution of state estimates:

The estimation methods used introduce errors into the measured states, which inevitably affects the principles of fact-based management. The introduction of complex evaluation functions will overcome this discrepancy.

## 7 Status Functions

In [26, 27], status functions are introduced and described to create a mapping of the institutional characteristics of socio-economic processes into the problems of constructing mathematical models of socio-economic processes.

We will call a status function (SF) an operation that establishes a rule for determining the correspondence of a certain ordered pair of arguments and a certain value corresponding to them. Such a function can be represented as

$$(r, k) = A(r)e^{i2\pi kr} \tag{5}$$

or

$$\begin{aligned}(r, k) &= A(r)\cos(2\pi kr) + iA(r)\sin(2\pi kr)\\ &= A(r)\left(\cos^2(\pi kr) - \sin^2(\pi kr)\right)\\ &\quad + 2iA(r)\cos(\pi kr)\sin(\pi kr).\end{aligned} \tag{6}$$

Status functions are an extension of the theory of fuzzy sets, forming a quantum–mechanical analogue of membership functions.

### *7.1 Comparison Status Functions with Membership Functions and Wave Functions*

Let us compare membership functions, status functions, and quantum mechanical wave functions [28].

(a) Functions. Status functions (SF) $\psi(r)$. Basic SFs are the orthonormal basis of the studied socio-economic system. The type of SF is selected based on the formation of an orthonormal basis of the assessment system, provided that the corresponding linguistic variable of the measured characteristic is adequate to the process of assessing the state of the control object.
(b) Meaning/physical meaning. The SF has the physical meaning of the wave function. The square of the SF modulus is defined as the probability that the value of the estimated characteristic corresponds to the given coordinates.
(c) Normalization. SF $\psi$ must satisfy the so-called normalization condition, for example, in the coordinate representation having the form:

$$\int_{V} \psi * \psi \mathrm{dV} = 1. \tag{7}$$

This condition expresses that the probability of estimating a system with a given SF in the introduced rating system is equal to one. In the general case, integration should be performed over all variables on which the SF in each representation depends.

(d) The principle of superposition/union. If the system can be in the states described by SF $\psi_1$ and $\psi_2$, then it can also be in the state described by WF $\psi_{\sum} = \sum_{i=1}^{n} c_i \psi_i$ for any complex $c_1$ and $c_2$. The superposition of any number of possible states is described by the SF $\psi_{\sum} = \sum_{i=1}^{n} c_i \psi_i$. In such a state, the square of the modulus ci determines the probability that, during the measurement, the system will be found in the state described by the SF $\psi_i$.

(e) Two variables are used for the real and imaginary parts. The variable for the real part of the SF is an independent dimensionless variable or variables of the introduced coordinate system. The variable for the imaginary part is the intentionality score. It is a latent assessment of the direction of the orientation of the state change, and the energy expended to achieve the result. For personalities, this is a direct assessment of intentionality; for socio-economic systems, it is the correspondence of the observed state to the expected one.

(f) Condition measurement process. Obtaining information about the state of the system by carrying out assessment activities. The measurement results are interpreted as the values of the estimated indicators. Measurement is done through interaction. During the interaction, the measured object is affected, therefore, the measurement results are distorted and the state of the measured system changes.

## *7.2 Comparison Fuzzy Set Theory with the Method of Status Functions and Quantum Mechanical Wave Functions*

Let us compare Fuzzy set theory, status functions, and quantum mechanical wave functions.

(a) Functions. Membership functions (FC) $\mu(r)$ are chosen based on the simplicity of their representation and calculation, provided that the corresponding linguistic variable is adequate to the process. Analogy FP—SF: used as a mathematical expression instead of a variable (for example, linguistic), which does not have it.

(b) Meaning/physical meaning. FP is “an improbable subjective measurement of inaccuracy”. Moreover, the FP is different from the probability density and from the probability distribution function. Sometimes interpreted as the possibility or usefulness of an event.

(c) Normalization. A fuzzy set $A$ is called normal if there is such an $r$ that $max\ \mu_A(r) = 1$. Otherwise, it is called subnormal. It is important to note that $\mu(r) \in [0;1]$.

(d) The principle of superposition/union. In terms of meaning, it is closest to the operation of combining fuzzy sets, for example, $A$ and $B$, and the fuzzy set $A \vee B$ is called, the FP of which is equal to: $\mu A \vee B(r) = \mu A(r) \vee \mu B(r)$, where $\vee$ the sign of the maximum operation.

(e) The variable is entered only for the real part—this is the basic variable; it can be chosen based on the conditions of the problem. There is no analogue for the variable complex part of the SF or WF.

(f) Condition measurement process. The fuzzification procedure (introduction of fuzziness) is the establishment of a correspondence between the numerical value of the input variable of the fuzzy inference system and the value of the FP.

## *7.3 Comparison Quantum Mechanical Wave Functions with the Method of Status Functions*

Let us compare status functions and quantum mechanical wave functions.

(a) Functions. Wave functions (WF) $\psi(r, t)$ or $\psi(r)$—in the stationary state, corresponds to the energy levels. Here WF is a complex function describing the state of a quantum mechanical system. Allows you to get the most complete information about the system, fundamentally achievable in the microworld. Analogy WF—SF: are harmonic, satisfy the Laplace equation.

If the socio-economic system has well-defined states and one can choose states that are defined as basic, then there is everything to use the mathematical formalism of quantum theory. Thus, the SFs correspond to the WFs of pure states. The pure states of a system are a fully described state. A state is a complete description of a closed system in a chosen basis.

(b) Meaning/physical meaning. The WF has no physical meaning. The physical meaning of the WF modulus squared is interpreted as a probability density (for discrete spectra, simply a probability) to detect a system in a position described by given coordinates at a given time.

(c) Normalization. The WF $\psi$ must satisfy the so-called normalization condition, for example, in the coordinate representation having the form (7).

This condition expresses the fact that the probability of finding a particle with a given wave function in a space of volume $V$ is equal to one. In the general case, integration should be carried out over all variables on which the WF in a given representation depends.

(d) The principle of superposition/union. If the system can be in the states described by the WF $\psi_1$ and $\psi_2$, then it can also be in the state described by the WF $\psi_\sum = c_1\psi_1 + c_2\psi_2$ for any complex $c_1$ and $c_2$. The superposition of any number of quantum states is described by the WF $\psi_{\sum} = \sum_{i=1}^{n} c_i \psi_i$. In such a state, the square of the modulus ci determines the probability that, during the measurement, the system will be found in the state described by the WF $\psi_{\mathrm{i}}$. Therefore, for normalized WFs $\sum_{i=1}^{n} |\mathrm{c}_i|^2 = 1$.

(e) Variables. For the real part, these are independent WF variables, particle coordinates, or momenta. For the imaginary part, this is the wave number directly proportional to the momentum. The choice of a system of units makes it possible to obtain the exact equality of the wave number and momentum. For light in a vacuum and other massless field, it is proportional to energy.

(f) Condition measurement process. Obtaining information about the state of the system by conducting a physical experiment. The measurement results are interpreted as the values of the measured physical quantity. If the measurement result remains unknown, then the quantum system goes into a state that is described by the density matrix. Measurement is done through interaction. During the interaction, there is an impact on the measured object, therefore, the measurement results are distorted and the state of the quantum system changes.

## 8 Conclusion

Thus, the chapter substantiates the use of quantum-like mathematical models based on status functions for the analysis of socio-economic systems. At the same time, an epistemological analysis of the concepts of classical and negative probabilities, membership functions for fuzzy sets, as well as quantum mechanical theory was carried out. The main mathematical relations for status functions are given for comparison with membership functions and wave functions. Moreover, the status functions are proposed to be used to display the institutional characteristics of socio-economic processes in the construction of mathematical models of socio-economic processes. Features of fuzzy set theory are also described for comparison with the method of status functions and quantum mechanical wave functions. In addition, the characteristics of quantum–mechanical wave functions are given for comparison with the method of status functions.

**Acknowledgements** This research was partially supported by the Russian Fund of Basic Research (grant No. 22-010-00465).

## References

1. Boldyrev, Y., Chernogorskiy, S., Shvetsov, K., Zherelo, A., Kostin, K.: A mathematical model of regional socio-economic development of the russian arctic zone. Resources **8**(1), 45 (2019). https://www.mdpi.com/2079-9276/8/1/45 (or) https://doi.org/10.3390/resources8010045

2. Bertotti, M.L., Modanese G.: Mathematical models for socio-economic problems. In: Mathematical Models and Methods for Planet Earth. Part of the Springer INdAM Series book series SINDAMS, vol. 6, pp. 123–134. Springer, Cham (2014). https://www.researchgate.net/publication/312825656 (or) https://doi.org/10.1007/978-3-319-02657-2_10
3. Davtian, A., Shabalina, O., Sadovnikova, N., Berestneva, O., Parygin, D.: Principles for modeling information flows in open socio-economic systems. In: Society 5.0: Human-Centered Society Challenges and Solutions. Studies in Systems, Decision and Control, vol. 416, pp. 167–173. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-95112-2_14
4. Protalinskiy, O., Khanova, A., Shcherbatov, I.: Simulation of power assets management process. In: Dolinina, O., Brovko, A., Pechenkin, V., Lvov A., Zhmud, V., Kreinovich, V. (eds.) Recent Research in Control Engineering and Decision Making. ICIT 2019. Studies in Systems, Decision and Control, vol. 199, pp. 488–501. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12072-6_40
5. Bondareva, I., Khanova, A., Khanova, Y.: Configuring systems based on petri nets, logic-probabilistic, and simulation models. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov M. (eds.) Cyber-Physical Systems: Modelling and Intelligent Control. Studies in Systems, Decision and Control, vol. 338, pp. 257–266. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66077-2_21
6. Davtian, A., Shabalina, O., Sadovnikova, N., Berestneva, O., Parygin, D.: Cyber-social system as a model of narrative management. In: Society 5.0: Cyberspace for Advanced Human-Centered Society, Studies in Systems, Decision and Control, vol. 333, pp. 3–14. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-63563-3_1
7. Abalkin, L.: The acceleration of socioeconomic development: its essence and sources. Probl. Econ. **29**(6), 50–62 (1986). https://doi.org/10.2753/PET1061-1991290650
8. Demchenko, S., Melnikova, T.A.: The methodology of developing the system of indicators to evaluate the socio-economic development efficiency. J. Siberian Fed. Univ. Humanit. Soc. Sci. **11**, 2356–2384 (2015). https://doi.org/10.17516/1997-1370-2015-8-11-2356-2384
9. Khairullina, R.E., Lushchik, V.I., Lampsi, E.O., Romanov, V. I., Burovkina, A.L., Krupnov, A.Y., Butyrin, Yu.A.: Specific features of regulating the forecasting system for the socio-economic development in Russia. Rev. Gênero Interdisciplinaridade, **2**(01) (2021). https://doi.org/10.51249/gei.v2i01.136
10. Xie, J., Sreenivasan, S., Korniss, G., Zhang, W., Lim, C., Szymanski1, B.K.: Social consensus through the influence of committed minorities. Phys. Rev. E **84**, 011130 (2011). http://arxiv.org/PS_cache/arxiv/pdf/1102/1102.3931v2.pdf (or) https://doi.org/10.1103/PhysRevE.84.011130
11. Furnham, A., Chamorro-Premuzic, T.: Personality: intelligence and general knowledge. Learn. Individ. Differ. **16**, 79–90 (2006)
12. Ackerman, P.L., Beier, M.E.: Intelligence, personality, and interests in the career choice process. J. Career Assess. **11**(2), 205–218 (2003)
13. Yang-Yu, L., Albert-László, B.: Control principles of complex systems. Rev. Mod. Phys. **88**, 035006–1–035006–58 (2016). https://doi.org/10.1103/RevModPhys.88.035006
14. Chisholm, R.M.: The encyclopedia of philosophy. In: Intentionality, 4 edn. (2001)
15. Jacob, P.: Stanford encyclopedia of philosophy. In: Intentionality (2012). https://plato.stanford.edu/entries/intentionality/. Accessed 21 December 2012
16. Searle, J.R.: What is an institution? J. Inst. Econ. **1**(1), 1–22 (2005)
17. Smith, B., Searle, J.: The construction social reality: An exchange. Am. J. Econ. Sociol. **2**(1), 285–309 (2003)
18. Searle, J.R.: The Construction Social Reality. Free Press, New York (1995)
19. Searle, J.R.: Collective intentions and actions. In: Cohen, P., Morgan, J., Pollack, M.E. (eds.) Intentions in Communication. Cambridge, MA: MIT Press (1990) (Reprinted in: Searle, J.R.: Consciousness and Language. Cambridge University Press, Cambridge, pp. 80–105 (2002)
20. Zadeh, L.A.: Fuzzy sets, fuzzy logic, and fuzzy systems. Inform. Control **8**, 238–353 (1966)
21. Dirac, P.A.M.: The physical interpretation of quantum mechanics. Proc. Roy. Soc. Lond. Ser. A **180**, 1–39 (1942). https://royalsociety.org/journals/ (or) https://doi.org/10.1098/rspa.1942.0023

22. Belinsky, A.V.: On the violation of causality in quantum experiments. Moscow Univ. Phys. Bull. **73**, 252–262 (2018). https://doi.org/10.3103/S0027134918030037 (or) http://www.scopus.com/inward/record.url?eid=2-s2.0-85050643036&partnerID=MN8TOARS
23. Klyshko, D.N.: Quantum communication with entangled photons. Adv. At. Mol. Opt. Phys. **42**, 489–533 (2000). https://doi.org/10.1016/S1049-250X(08)60192-5
24. Home, D., Selleri, F.: Bell's theorem and the EPR paradox. Riv. Nuovo Cim. **14**(9), 1–95 (1991)
25. Sokolovski, D.: Are the 'weak measurements' really measurements? Quanta **2**, 50–57 (2013). https://arxiv.org/pdf/1305.4809.pdf (or) https://doi.org/10.12743/quanta.v2i1.15
26. Veshneva, I.V., Singatulin, R.A.: The status functions application for multispectral data images processing in virtual reality systems. In: Cyber-Physical Systems. Studies in Systems, Decision and Control, vol. 350, pp. 225–235. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67892-0_19
27. Veshneva, I., Bolshakov, A., Fedorova, A.: Analysis of the competitiveness risks of food production enterprises using mathematical modelling methods. In: Creativity in Intelligent Technologies and Data Science. CIT&DS 2021. Communications in Computer and Information Science, vol. 1448, pp. 238–249. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87034-8_18
28. Sanz, A.S., Miret-Arte, S.: Quantum phase analysis with quantum trajectories: a step towards the creation of a bohmian thinking. Am. J. Phys. **80**(525), 525–533 (2011). https://doi.org/10.1119/1.3698324

# Application of Quantum-Like Mathematical Models Based on Status Functions for the Analysis of Socio-Economic Systems. Part 2: Modelling Based on Status Functions

**Irina Veshneva and Alexander Bolshakov**

**Abstract** It is proposed to use quantum-like models with the use of status functions for mathematical modelling and subsequent analysis of complex socio-economic systems. The limitations of the methods of classical probability theory and mathematical statistics, as well as the theory of fuzzy sets, algorithms Mamdani, Suzuki and others for solving similar problems are described. A description of the main assumptions that are used in the mathematical modelling of socio-economic objects based on status functions is given. Examples are considered that describe the features of transition paths through intermediate states. An operator for the transition of a socio-economic system to various states, similar to the Hamiltonian, is presented. A spectrum of possible virtual trajectories is introduced to describe transitions to different states. A mathematical model based on status functions is proposed to describe the transition of the system to a measurable state. In the proposed Hamiltonian, the first term represents a subsystem of indicators, the second is an analogy of the energy of indicators in the information environment. At the same time, terms are distinguished that are analogues of the energies of the system of two controlled indicators: interaction, kinetic and potential. The description of the results of mathematical modelling and the analysis of the interaction of two hypothetical indicators of the socio-economic system are given. The indicators are taken from the statistics of innovation indicators of one of the regions of the Russian Federation.

**Keywords** Status function · Mathematical model · Socio-economic system · Quantum-like model

I. Veshneva
Saratov State University, 83, Astrakhanskaya Street, Saratov 410012, Russia
e-mail: veshnevaiv@gmail.ru

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



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_12

## 1 Introduction

There are various approaches to describe by mathematical methods the processes and phenomena that occur in socio-economic systems [1–7]. Historically, the methods of probability theory were the first to be used to understand the essence of occurring phenomena and attempts at predictive assessments of processes to support decision-making. The development of methods of probability theory and mathematical statistics has led to the emergence of the theory of fault tolerance, the creation of decision support algorithms, etc. It is necessary to single out the theory of fuzzy sets, the algorithms of Mamani, Sugeno, etc. On the other hand, the development of physics and the emergence of quantum physics led to the creation and development of mathematical approaches for describing phenomena that could not previously be explained by the methods of earlier mathematical models.

When describing socio-economic phenomena, such as decision-making by an individual or managing the competitiveness of regions and countries, one has to make decisions under conditions of “deep uncertainty”, act when there is deep uncertainty both in relation to the possible actions of other decision-makers and the surrounding complex information environment. The information environment includes the objects under study, as well as the information conditions of their existence and channels of information exchange. This information environment may also include decision confidence states. The last statement is very important when making managerial decisions and is not considered in most of the mathematical models used. Confidence in decision-making should be attributed to hidden factors that affect the results of an event occurrence experiment but are not directly measured. For example, as the phase of the oscillation when measuring the interference of light. Consequently, it becomes possible to use mathematical models of quantum mechanics [8–10] in the study of phenomena and processes occurring in the socio-economic and information environments.

In the study of decision-making processes in the socio-economic information environment, criteria for primary official information, information channels, and secondary information arising from the processing of primary information can be separately introduced and used. It is possible to use the concepts of “degree of information”: first-order information, which links information with an event, and “higher-order” information, which refers to information itself [11]. The notion of “common knowledge” or “institutional knowledge”, which is used when introducing status functions, is an example of such higher order information.

The model presented in this chapter uses quantum-like models for assessing indicators of the socio-economic environment based on status functions. Status functions were used to evaluate the states of the indicators. The indicators are interconnected through the information environment. To describe the process of adaptability, the mathematical apparatus of quantum field theory is used. The process of adapting indicator estimates to the environment is represented by the dynamics of quantum operators. Certain values are extracted from the general assessment of the state of the environment, which are the probabilities of the state of a given set of indicators of the state of objects in the system. In this case, the medium is represented as a quantum field. Scenarios of system evolution are presented considering the degrees of freedom of the environment. The advantage of the approach used is that it is possible to describe the complex dynamics of a system solely in terms of pure states. To represent pure states, the status functions introduced in the works of the authors of the chapter are used.

Status functions represent the maximum information about the state of the described system. They describe the indicator at the initial moment of time. Then this state is changed under the influence of interaction with other indicators and the environment. The superposition of different states encodes the hidden deep uncertainty of the system, which cannot be modelled within the framework of classical probability and mathematical statistics.

## 2 Socio-Economic Objects and Their Modeling Based on Status

We will define a socio-economic system $S$ as a set of social objects and relationships between them. In this case, the following assumptions were made.

1. Here $s_1$, …, $s_N$ are socio-economic objects that are subsystems; each subsystem is described by its intrinsic characteristics. The results of measuring these characteristics can be described based on the basis vectors of the space. For this, a rule is indicated that allows you to determine for any two elements of the space their scalar product.
2. The state of a socio-economic object can be described by a set of characteristics at the time of measurement, which are ordered and form a vector $|x\rangle$. The basic "pure states" of the system are determined only by the direction of the vector $|x\rangle$, and not by its length. These states form an orthonormal basis of the system, or the state of the system can be described based on them. In this case, there are no states, the description of which is impossible based on these vectors. Therefore, the evolution of a system can be described by a unitary operator, one that describes only a change in the direction of the system's state vector. The evolution of the system in time is also described by a unitary operator:

$$|x(t)\rangle = \hat{U}(t, t_0)|x(t_0)\rangle. \tag{1}$$

3. If the measured state of the system $|y_i\rangle$ can be described based on the characteristic vectors of the state $|x\rangle$, then this state should be represented as a superposition:

$$|x\rangle = \sum_i c_i|y_i\rangle. \tag{2}$$

This is a key difference from traditional measurements, which are subjected to statistical processing based on the mathematical apparatus of classical probability theory.

In this case, the transition from one state to another $|x\rangle \rightarrow |y_i\rangle$ is built based on mappings that establish connections between the internal characteristics of the described subsystems $s_i \rightarrow s_j$ $(i, j = 1, \ldots, n)$.

The transition from the state $|x\rangle$ to the state $|y_i\rangle$ corresponds to the probability amplitude:

$$c_i = \langle y_i \mid x \rangle, \tag{3}$$

which is a complex number, and the square of its modulus is the transition probability $|x\rangle \rightarrow |y\rangle$:

$$P(x \rightarrow y) = \left|c_{xy}\right|^2 \equiv |\langle y \mid x \rangle|^2. \tag{4}$$

In this way, projection operators are formed that describe the mapping of the observed state of the system onto the pure states of the system.

4. If the measured system was in the state $|x\rangle$ and after the measurement it goes into the set of states $|y_i\rangle$, represented by the combination $|x\rangle = \sum_i c_i |y_i\rangle$, then the choice of $|y_i\rangle$ can be done quite by accident. In this case, the conditional probability of finding a system in the state $|y_i\rangle$ :

$$\begin{aligned} \mathrm{w}_i &= |\mathrm{c}_i|^2 = \langle x \mid y_i \rangle \langle y_i \mid x \rangle \\ &= \left\langle x \middle| \widehat{P_{y_i}} \middle| x \right\rangle, \end{aligned} \tag{5}$$

where $\hat{P}_{y_i}$ is a projector, that is, an operator mapping the measured state $|y_i\rangle$ onto pure states $|x\rangle$. It is important to note that according to the mathematical model of the state, the basis is chosen randomly, the measurement result may be random. The observed patterns are the result of the manifestation of the hidden parameters of the system. This is a manifestation of randomness—one of the fundamental laws of Nature. Thus, the values observed in practice, which are the result of measuring statistical data on the parameters of the socio-economic system, should be placed in the form of a projector-operator that maps the measured state $|y_i\rangle$ to pure states $|x\rangle$.

5. If the transition from the state $|x\rangle$ to the state $|y_i\rangle$ is possible along two different paths path 1 and path 2, then

$$\langle y_i \mid x \rangle = \langle y_i \mid x \rangle_{\text{path 1}} + \langle y_i \mid x \rangle_{\text{path 2}} \tag{6}$$

See Fig. 1.

If paths 1 and 2 relate to "intermediate" states $|1\rangle$ and $|2\rangle$, respectively, then formula (1) can be rewritten as

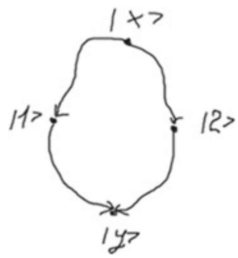



**Fig. 1** Transition paths path 1 path 2 $|x\rangle \rightarrow |y_i\rangle\rangle$ through intermediate states $|1\rangle$ and $|2\rangle$

$$\langle y_i \mid x\rangle = \langle y_i \mid 1\rangle\langle 1 \mid x\rangle + \langle y_i \mid 2\rangle\langle 2 \mid x\rangle, \tag{7}$$

This means that the probability $P_{y_i}$ of the transition $|x\rangle \rightarrow |y_i\rangle$ is generally not equal to the sum of the probabilities $P_{y_i}^{\text{path 1}} + P_{y_i}^{\text{path 2}}$ of the transitions $|x\rangle \rightarrow |y_i\rangle$, through 1 and $|x\rangle \rightarrow |y_i\rangle$ through 2, as described, for example, in the experiments of Belinsky [12–14] and Sect. 3 of this chapter. The probability contains an interference term, which is responsible for the possibility of not realizing some event and the possibility of interaction between subsystems. In the social sciences, this phenomenon is well known and is called synergy.

System status after measurement

$$|y_i\rangle = \frac{\hat{P}_{y_i}}{\sqrt{\mathrm{w}_i}} \mathrm{e}^{i\varphi}. \tag{8}$$

Changing the phase difference affects the interference pattern. The result of the interaction of subsystems and the ways of transition of these subsystems from one state to another becomes available for understanding, observation, and control [15, 16].

## 3 Jump Statement Representation

We will denote the state of the system by a node. Let $|x\rangle$, $|y\rangle$, … be the state of the system at the nodes, described by a set of characteristics now of measurement. When the impact $\mathrm{u}(\mathrm{u}_1, \mathrm{u}_2, \ldots)$ is implemented on the system, it passes from one state to another $|x\rangle \rightarrow |y\rangle$. We will assume the time of measurement to be fixed. This is impossible to implement. However, suppose that the real time interval is divided into $N$ intervals and $N \rightarrow \infty$. Let us assume that the energy operator is like the Hamiltonian $\hat{H}$. Then we can write the transition amplitudes as a sum over all possible virtual paths, which we denote by $\hat{A}$. We get:

$$\left\langle x \left| -\frac{i\hat{H}t}{h} \right| y \right\rangle = \lim_{N\to\infty} \sum_{k_1, k_2, \dots k_{N+1}} \langle x | a_{k_{N+1}} \rangle \left\langle a_{k_{N+1}} \left| -\frac{i\hat{H}t}{h} \right| a_{k_N} \right\rangle \langle a_{k_N} | \dots | a_{k_2} \rangle \left\langle a_{k_2} \left| -\frac{i\hat{H}t}{h} \right| a_{k_1} \right\rangle \langle a_{k_1} \mid y \rangle = \sum_{\text{path}} P^{x\leftarrow y}, \tag{9}$$

where $h$ is a constant following from the normalization of the scalar product of vectors in a given Hilbert space, $a_k$ and $|a_k\rangle$ are the eigenvalues and eigenvectors of the variable $\hat{A}$ describing the transition:

$$f^{\text{path}} = \int_0^t \beta(t')a(t')dt', \tag{10}$$

where $\beta(t)$ is a known function, the choice of which is arbitrary, and it is the basis function of the canonical expansion. In our case and in what follows, we will assume that these are functions of the orthogonal basis of the system, represented by status functions [17, 18].

Let a grouping of possible paths be carried out and a function of paths be formed for which the value $f^{\text{path}}$ is equal to some $F$:

$$\phi^{x\leftarrow y}(t|F) = \sum_{\text{path}} \delta\left(F - f^{\text{path}}\right) A^{\text{path}}_{x\leftarrow y}. \tag{11}$$

Then we smooth this group with a certain function, for example, a Gaussian distribution, and obtain the scattering amplitude

$$\psi^{x\leftarrow y}(t|F) = \sum G\left(F - F'\right)\phi^{x\leftarrow y}(t|F)dF'. \tag{12}$$

If we use the basis $\{\beta(t)\}$ containing the observed state $|\psi(t, F)\rangle$ [19], then we can use an equation like:

$$i\partial/\partial t|\psi(t|F)\rangle = \left[\hat{H}\beta(t)\hat{A}|\psi(t|F)\rangle\right] - ih\partial/\partial F\beta(t)\hat{A}|\psi(t|F)\rangle. \tag{13}$$

Thus, the movement of the state of the subsystem along possible trajectories can be described. Thus, R. Feynman carries out similar arguments on the integration of close trajectories [20].

## 4 Model for Describing the Transition of a System to a Measurable State

Model for describing the transition of a system to a measurable state The initial general idea was to formally use status functions [21], instead of membership functions in decision support algorithms based on the Mamdani algorithm. The advantage of the status functions introduced by the authors is that they can be used like the building blocks of quantum field theory in macroscopic systems, so that dynamical systems can be considered based on quantum-like mathematical models. Status functions are analogous to the system state vector and can be used to represent very complex dynamic processes. These states can be of any nature: physical, cognitive, technological, socio-economic. This is the minimum interpretation that can be used to move to quantum field theory. From this formal point of view, the application of the mathematical formalism of quantum theory is simply a matter of convenience, and in fact it has proved useful in several applications in various contexts [22–24].

Note that one can start constructing the space of system state vectors starting from any vector belonging to the introduced Hilbert space. These vectors make it possible to describe the system of methods for introducing operators of the type of the Hamilton operator and to apply the matrix form for describing states.

Application of an operator like the Hamiltonian. Assume that the two measurable indicators of the socio-economic system do not interact directly. Let there be an infinite number of ways of transition from one measurable state to another. However, due to the interaction of hidden parameters, this transition is carried out along selected close trajectories. Moreover, $\sum_{\text{path}} P^{x \leftarrow y}$ can be represented as an integral over the space of close trajectories. Let's build a model in which the control of the state of the socio-economic system is carried out according to two controlled groups of indicators, represented by the states $\psi_1$ and $\psi_2$. Note that the model does not reflect complex internal processes in the system that lead to transitions between states. It provides only a formal operator representation of these transitions. Let us assume that the region can redistribute management resources between these groups of indicators, which in this context turn out to be competing through the information and communication environment. This is how indicators interact.

We will represent the Hamiltonian in the form:

$$\hat{H} = \hat{H}_1 + \hat{H}_2. \tag{14}$$

Here $\hat{H}_1$ is the Hamiltonian of the indicator subsystem; $\hat{H}_2$ is the Hamiltonian of the indicator energy analogue in the information environment.

Let the operators be used in the traditional form:

$$\hat{H}_1 = \frac{\partial}{\partial t} - ih\nabla^2; \tag{15}$$

$$\hat{H}_2 = \begin{pmatrix} 0 & 0 \\ 0 & \text{inter}_{22} * \psi_1 \end{pmatrix} + \begin{pmatrix} t_{11} * \psi_1 & 0 \\ t_{21} * \psi_2 & 0 \end{pmatrix} + \begin{pmatrix} \text{v}_{11} & \text{v}_{12} \\ \text{v}_{21} & \text{v}_{22} \end{pmatrix}, \tag{16}$$

here $\frac{\partial}{\partial t}$ is the time derivative, $i$ is the imaginary unit, $h$ is the normalization constant of the space dimension, $\nabla^2$ is the divergence gradient. In our case, this is the second derivative of the status function of the indicator with respect to the base variable. Equation (18) contains terms that are analogues of the energies of the system of two controlled indicators: interaction, kinetic and potential, respectively. We will assume that the energy analogue coefficients $\text{inter}_{22}$, $t_{11}$, $t_{21}$, $\text{v}_{12}$, $\text{v}_{21}$ are constants. We define $|\psi\rangle$ based on status functions in the form:

$$|\psi\rangle = \begin{pmatrix} \psi_1(r) \\ \psi_2(r) \end{pmatrix} = \begin{pmatrix} 2.26\exp\left(-2i\pi r - 41.32(0.13 + r)^2\right) \\ 2.33\exp(2i\pi r)\begin{pmatrix} \exp\left(-41.32(-0.13 + r)^2\right) \\ -0.24\exp\left(-41.32(0.13 + r)^2\right) \end{pmatrix} \end{pmatrix}, \tag{17}$$

where $r$ is the introduced basic variable. The real parts of the status functions are shown in Fig. 2.

In the operator of the potential energy analog, we will change the values of $\text{v}_{11}$ and $\text{v}_{22}$ in such a way as to compare the abstract case and the case that we will try to associate with statistical data. In addition, we will change the coefficient $\text{inter}_{22}$ in the interaction operator and analyze the manifestation of the interaction result in the results of a numerical experiment.

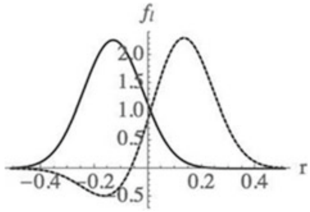



**Fig. 2** Real parts of status functions

## 5 Simulation Results and Discussion

Let's consider the calculation results for $t_{11} = 0, 1$, $t_{21} = 0, 25$, $v_{12} = 0, 04$, $v_{21} = 0, 05$, $v_{11} = 1$ и $v_{22} = 1$ . We assume that an analysis of the interaction of two hypothetical indicators of the socio-economic system is being carried out.

We will use status functions to create a model of an infinite set of trajectories along which the system passes from one state to another. Using (9), we obtain the following expressions for calculating the probability of transition from one state to another (fragment):

$$\begin{aligned}
&\frac{\partial \psi_1[t]}{\partial t}\\
&= \int_{-1}^{1} \mathrm{dr}\Big(0.0566182e^{-41.3223r^2-2i\pi r-10.7438r-0.698347}\psi_1[t]\\
&+1.0258\mathrm{e}^{-82.6446r^2-21.4876r-1.39669}\psi_1[t]^2\\
&+0.5129\mathrm{e}^{-82.6446r^2-4i\pi r-21.4876r-1.39669}\psi_1[t]^2\\
&+0.578304\mathrm{e}^{-41.3223r^2-2i\pi r-10.7438r-0.698347}\psi_2[t]\Big);
\end{aligned} \tag{18}$$

$$\begin{aligned}
&\frac{\partial \psi_2[t]}{\partial t}\\
&= \int_{-1}^{1} \mathrm{dr}\Big(2.26473\mathrm{e}^{-41.3223r^2-2i\pi r-10.7438r-0.698347}\psi_1[t]\\
&+0.0144576\mathrm{e}^{-41.3223r^2-2i\pi r-10.7438r-0.698347}\psi_2[t]\\
&+0.058435\mathrm{e}^{-41.3223r^2-2i\pi r+10.7438r-0.698347}P2(t)\\
&-2.64679\mathrm{e}^{-82.6446r^2-1.39669}\psi_1[t]\psi_2[t]\\
&+1.90965\mathrm{e}^{-82.6446r^2-21.4876r-1.39669}\psi_1[t]\psi_2[t]\\
&+\Big(0.540691\mathrm{e}^{-82.6446r^2-1.39669}\\
&+0.0668871\mathrm{e}^{-82.6446r^2-21.4876r-1.39669}\\
&+1.09269\mathrm{e}^{-82.6446r^2+21.4876r-1.39669}\ldots\Big)\psi_2[t]^2.
\end{aligned} \tag{19}$$

Let us integrate over the basic variable and obtain expressions for $\hat{H}|\psi\rangle$ after integrating over the basic variable, which erases the trajectories of the interaction of indicators in the information and communication environment.

The joint solution of Eqs. (20), (21) is shown in Fig. 3. There is a moment in time at which the "interaction" occurs. An increase in the interaction coefficient leads to a shift in the moment of interaction towards a shorter time. At the same time, the "strength of interaction", with which the maximum value of the peak of conditional probabilities of status functions can be compared, does not directly depend on the

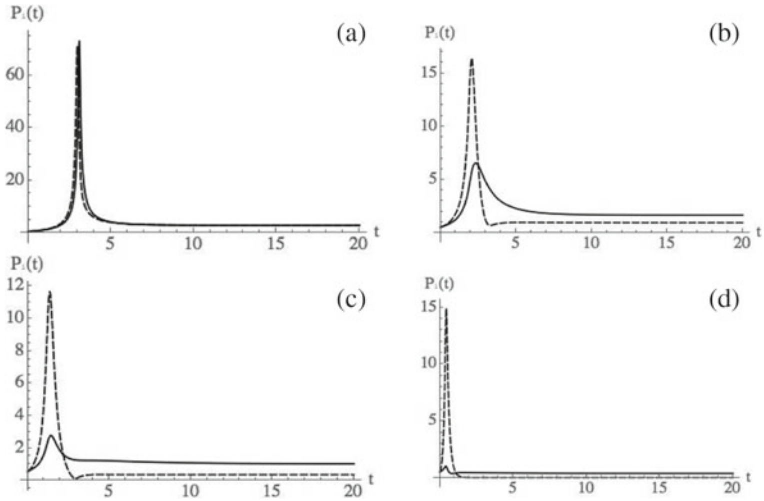



**Fig. 3** The results of numerical simulation of solutions to the system of Eqs. (20), (21) for the values $t_{11} = 0, 8$; $t_{21} = 0, 025$; $\mathrm{v}_{11} = 0, 04$; $\mathrm{v}_{22} = 0, 04$; $\mathrm{v}_{12} = 1$; $\mathrm{v}_{21} = 1$; **a** $\mathrm{inter}_{22} = 0, 05$; **b** $\mathrm{inter}_{22} = 0, 5$; **c** $\mathrm{inter}_{22} = 5$; **d** $\mathrm{inter}_{22} = 10$

interaction coefficient. The rest of the coefficients in the simulation of the interaction remain unchanged. The initial values of status functions at the start of modeling are set based on official statistics for assessing the competitiveness of Russian regions [25]: 〚∂ψ〛 _1 [0]= 0.51; 〚∂ψ〛 _2 [0]= 0.58.

Let us assume that the results of statistical data testify to a certain path passed by the subsystem of the measured indicator. Then the statistical data is related to the "potential" energy of the system. Let's take the statistics data [26] and add them to the coefficients $\mathrm{v}_{12}$ and $\mathrm{v}_{21}$. To perform this procedure, let's take the statistics data for two indicators included in the group of innovation indicators [25] (marked with dots in Fig. 4). The group indicators are normalized to 1. As a result, we use the data for the indicator $I_1\{0.5832, 0.6024, 0.5928, 0.6584, 0.6224, 0.844, 0.8536, 0.7952, 0.8112, 0.7512\}$ and for the indicator $I_2\{0.512, 0.44, 0.50, 0.56, 0.56, 0.384, 0.9432, 0.892, 0.488\}$. The interpolation results are shown in Fig. 4 solid curve. Relevant expressions:

$$\begin{aligned} I_1[t] &= \mathrm{v}_{12} \\ &= 111.823\mathrm{ie}^{-0.02\mathrm{i}t} - 111.823\mathrm{ie}^{0.02\mathrm{i}t} \\ &-219.357\mathrm{ie}^{-0.01it} + 219.357\mathrm{ie}^{+0.01\mathrm{i}t} \\ &+0.524102\mathrm{e}^{-0.1t} + 0.0257454\mathrm{e}^{-\mathrm{i}t} + 0.0257454\mathrm{e}^{\mathrm{i}t}; \end{aligned} \tag{20}$$

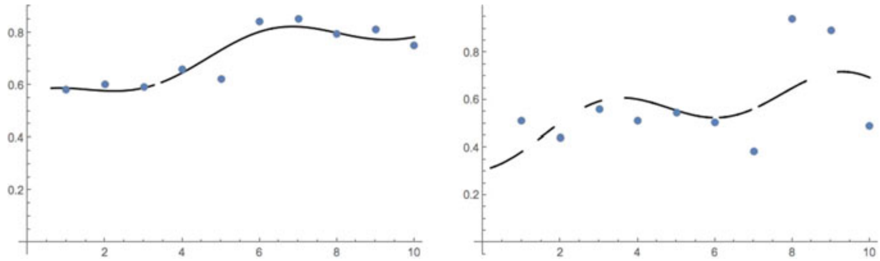



**Fig. 4** For the indicator $I_1(t)$ (left) and for the indicator $I_2(t)$ (right), the time dependence of the normalized statistical data (marked with dots) and the results of interpolation by trigonometric functions (solid line). The indicators are taken from the statistics of innovation indicators of one of the regions of the Russian Federation

$$\begin{aligned} I_2[t] &= \mathrm{v}_{21} \\ &= \left(137.058\mathrm{i}e^{-0.02\mathrm{i}t} - 137.058\mathrm{i}e^{0.02\mathrm{i}t} - 270.354\mathrm{i}\right)e^{-0.01\mathrm{i}t} \\ &+ 270.354\mathrm{i}e^{0.01\mathrm{i}t} + 0.39192e^{-0.1t} \\ &- 0.0429692e^{-\mathrm{i}t} - 0.0429692e^{\mathrm{i}t}; \end{aligned} \tag{21}$$

We will assume that the measured indicators at some point in time interact through the information and communication environment. After integration over trajectories, we obtain the expressions:

$$\begin{aligned} &\frac{\partial\psi_1(t)}{\partial t} \\ &= -3.69228\mathrm{i}e^{-0.01\mathrm{i}t} - 1.88223\mathrm{i}e^{0.02\mathrm{i}t} \\ &+ 1.88223\mathrm{i}e^{-0.02\mathrm{i}t} + +3.69228\mathrm{i}e^{0.01\mathrm{i}t} \\ &+ 0.192211\psi_1(t)^2 + 0.522969\psi_2(t) \\ &+ 0.00882184e^{-0.1t} + 0.000433354e^{-\mathrm{i}t} + 0.000433354e^{\mathrm{i}t} \end{aligned} \tag{22}$$

$$\begin{aligned} &\frac{\partial\psi_2(t)}{\partial t} \\ &= (0.564687 - 0.230465i)\psi_1(t)\psi_2(t) \\ &- 182.027\mathrm{i}e^{(0.-0.01\mathrm{i})t} - 92.2798\mathrm{i}e^{0.02\mathrm{i}t} \\ &+ 92.2798\mathrm{i}e^{-0.02\mathrm{i}t} + 182.027\mathrm{i}e^{0.01\mathrm{i}t} \\ &+ 0.175019\psi_2(t)^2 + 0.0130742\psi_2(t) \\ &+ 0.263877e^{-0.1t} - 0.0289308e^{-\mathrm{i}t} - 0.0289308e^{\mathrm{i}t} \end{aligned} \tag{23}$$

The calculation results are shown in Fig. 5 for different coefficients $\mathrm{inter}_{22}$. In this case, an increase in interaction $\mathrm{inter}_{22}$ leads to an earlier interaction time and does not have an explicit dependence in the maximum amplitude of the interaction peak.

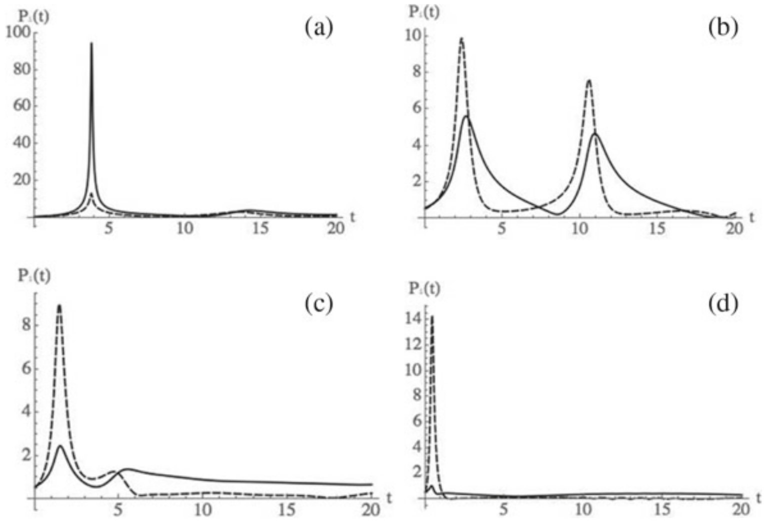



**Fig. 5** The results of numerical simulation of solutions to the system of Eqs. (22), (23) for the values $t_{11} = 0, 8$; $t_{21} = 0, 025$; $\mathrm{v}_{11} = 0, 04$; $\mathrm{v}_{22} = 0, 04$; $\mathrm{v}_{12} = I_1(t)$; $\mathrm{v}_{21} = I_2(t)$; **a** $\mathrm{inter}_{22} = 0, 05$; **b** $\mathrm{inter}_{22} = 0, 5$; **c** $\mathrm{inter}_{22} = 5$; **d** $\mathrm{inter}_{22} = 10$

For a certain range of $\mathrm{inter}_{22}$ values, the interaction is complex, and the conditional probability of the resulting indicator value demonstrates the possibility of beats. There is no mode stabilization.

It can be assumed that under the conditions of two internal subsystems competing for limited common resources, solutions such as solutions of the Lotka-Volter equations may arise. This can give rise to an area of system instability where a small control action can lead to rapid changes in the entire system.

The experiment used a model for two indicators of the innovative component of the assessment of the region's competitiveness. Creating a model with many indicators will lead to more complex solutions and the ability to explore the dependence of regimes on parameters.

Further study of the system involves, firstly, the creation of a model of the control action on the system, and secondly, the creation of a network of interacting indicators. Then, we will introduce indicators of different regions into the interpolation data and study how the dynamics of conditional risk probabilities, represented by indicators for different regions, changes. At the same time, it is precisely the creation of conditions for the possible existence of various regimes that threaten the loss of competitiveness of the regions of the Russian Federation is of particular interest.

The availability of statistical data will make it possible to determine the regimes corresponding to the stabilization of the system, and the occurrence of limit cycles in the phase space of the system. These are the modes and indicators that management structures strive for. In this case, the system perturbed after the measurement stabilizes at specific risk values. However, the emergence of more complex regimes is possible, the study of which is carried out in the phase space of the system.

## 6 Conclusion

The chapter provides a justification for the expediency of using the apparatus of status functions proposed by the authors for mathematical modelling of complex socio-economic objects. Traditional mathematical methods for description, such as probability theory, mathematical statistics, fuzzy models, are not able to describe the specifics of the phenomena under study. These phenomena, on the one hand, have a random character, on the other hand, it is required to formalize the subjective assessment of the person making management decisions.

For this, it is proposed to use status functions representing complex-valued functions of time. At the same time, the real part reflects the objective aspect of measuring the indicators of socio-economic processes, and the imaginary part characterizes the subjective aspect.

The main assumptions that are used in the construction of the proposed quantum-like mathematical models are described. Among them is the representation of the state of the socio-economic system as a set of characteristics at the time of measurement, which are ordered and form a certain vector. Moreover, the basic states of the system form an orthonormal basis, based on which any state of the system is described. In this case, the change in the state of the system is described by a unitary operator, which characterizes the change in the direction of the vector of the state of the system.

Based on status functions, a mathematical model is proposed to describe various transitions of the socio-economic system to a measurable state. At the same time, the possibility is realized based on status functions to create mathematical models similar to quantum ones. The studied states can be of different nature: physical, cognitive, technological, socio-economic.

The final part of the chapter provides a description of the example. It presents the simulation results as well as their analysis. An analysis of the interaction of two hypothetical indicators of the socio-economic system has been carried out. Expressions are obtained for determining the probability of a system transition from one state to another. The initial values of the status functions at the time of modelling are set on the basis of official statistics to assess the competitiveness of the regions of the Russian Federation. An analysis of the creation of conditions for the threat of loss of competitiveness is carried out on the example of the region of the Russian Federation. The data for this are taken from the official statistics of innovation indicators.

**Acknowledgements** This research was partially supported by the Russian Fund of Basic Research (grant No. 22-010-00465).

## References


1. Baymuratova, K.A.: Mathematical modeling of socio-economic processes. Sci. Prog. **2**(8) (2021). https://cyberleninka.ru/article/n/matemathical-modeling-of-socio-economic-processes
2. Davtian, A., Shabalina, O., Sadovnikova, N., Berestneva, O., Parygin, D.: Principles for modeling information flows in open socio-economic systems. In: Society 5.0: Human-Centered Society Challenges and Solutions. Studies in Systems, Decision and Control, vol. 416, pp. 167–173. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-95112-2_14
3. Akberov, K.C., Chernyakov, M.K., Chernyakova, M.M.: Economic-mathematical modeling of the potential of municipalities. Int. J. Prof. Sci. **3** (2017). http://scipro.ru/article/05-03-17
4. Protalinskiy, O., Khanova, A., Shcherbatov, I.: Simulation of power assets management process. In: Dolinina, O., Brovko, A., Pechenkin, V., Lvov A., Zhmud, V., Kreinovich, V. (eds.) Recent Research in Control Engineering and Decision Making. ICIT 2019. Studies in Systems, Decision and Control, vol. 199, pp. 488–501. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12072-6_40
5. Nikulina, T.N., Zhirnova, I.S., Stupina, A.A., Zhirnov, A.A.: Mathematical modeling of economic processes in complex systems (on the example of Krasnoyarsk municipality). J. Phys. Conf. Ser. **1353**, 012118 (2019). https://doi.org/10.1088/1742-6596/1353/1/012118
6. Bondareva, I., Khanova, A., Khanova, Y.: Configuring systems based on petri nets, logic-probabilistic, and simulation models. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Cyber-Physical Systems: Modelling and Intelligent Control. Studies in Systems, Decision and Control, vol. 338, pp. 257–266. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66077-2_21
7. Davtian, A., Shabalina, O., Sadovnikova, N., Berestneva, O., Parygin, D.: Cyber-Social System as a Model of Narrative Management. In: Society 5.0: Cyberspace for Advanced Human-Centered Society, Studies in Systems, Decision and Control, vol. 333, pp. 3–14. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-63563-3_1
8. Kupczynski, M.: Quantum mechanics and modelling of physical reality. Phys. Scr. **93**, 123001 (2018). https://arxiv.org/ftp/arxiv/papers/1804/1804.02288.pdf (or) https://doi.org/10.1088/1402-4896/aae212
9. Kupczynski, M.: EPR paradox, quantum nonlocality and physical reality. J. Phys. Conf. Ser. **701**, 012021 (2016)
10. Allahverdyan, A.E., Balian, R., Nieuwenhuizen, T.M.: A subensemble theory of ideal quantum measurement processes. Ann. Phys. **376**, 324–352 (2017)
11. Haven, E., Khrennikov, A.: Quantum probability and the mathematical modelling of decision-making. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. **374**(2058), 20150105 (2015). https://doi.org/10.1098/rsta.2015.0105
12. Bikhchandani, S., Hirshleifer, J., Riley, J.G.: The Analytics of Uncertainty and Information, p. 191. Cambridge University Press, Cambridge (2013)
13. Belinskii, A.V.: Methodological notes: bell's paradoxes without the introduction of hidden variables. Phys.-Usp. **37**(4), 413–419 (1994). https://doi.org/10.1070/PU1994v037n04ABEH000024
14. Belinskii, A.V.: A generalized Bell's theorem. Zh. Eksp. Teor. Fiz. **105**, 818–827 (1994)
15. Belinskii, A.V.: Objective reality and the paradox of Wigner friends. Opt. Spectrosc. **128**, 1421–1424 (2020)
16. Hardy, L.: Quantum Theory from Five Reasonable Axioms (2001). arXiv:quant-ph/0101012. https://doi.org/10.48550/arXiv.quant-ph/0101012
17. Reyes-Galindo, L.: The Sociology of Theoretical Physics. PhD thesis, Cardiff University School of Social Sciences (2022). https://www.researchgate.net/publication/277106525_The_sociology_of_theoretical_physics
18. Veshneva, I.V., Bolshakov, A.A., Fedorova, A.E.: Organization of engineering education for the development of cyber-physical systems based on the assessment of competences using

status functions. In: Cyber-Physical Systems: Industry 4.0 Challenges. Studies in Systems, Decision and Control, vol. 260, pp. 277–288. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-32648-7_22. ISSN 21984182
19. Veshneva, I., Bolshakov, A., Kulik, A.: Increasing the safety of flights with the use of mathematical model based on status functions. In: Recent Research in Control Engineering and Decision Making. ICIT 2019. Studies in Systems, Decision and Control, vol. 199, pp. 608–621. Springer Nature Switzerland AG (2019). https://doi.org/10.1007/978-3-030-12072-6_49. ISSN 21984182
20. Sokolovskii, D.: Path integral approach to space-time probabilities: a theory without pitfalls but with strict rules. Phys. Rev. D **87**(7), 076001 (2013). https://doi.org/10.1103/PhysRevD.87.076001
21. Feynman, R.P., Hibbs, A.R., Styer, D.F.: Quantum Mechanics and Path Integrals, emended edition. McGraw-Hill, New York, p. 382 (2005). ISBN-13: 978–0–486–47722–0. ISBN-10: 0–486–47722–3
22. Bolshakov, A.A., Veshneva, I.V., Lushin, D.: Mathematical model of integration of cyber-physical systems for solving problems of increasing the competitiveness of the regions of the Russian Federation. In: Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, vol. 333, pp. 129–139. Springer Nature Switzerland AG (2021). https://doi.org/10.1007/978-3-030-63563-3_11. ISSN 2198–4182
23. Bagarello, F.: Quantum Dynamics for Classical Systems: With Applications of the Number Operator, p. 248. Wiley, New York (2012). ISBN 978–1–118–37068–1
24. Bagarello, F.: A quantum-like view to a generalized two player's game. Int. J. Theor. Phys. **54**(10), 3612–3627 (2015)
25. Bagarello, F., Cinà, M., Gargano, F.: Projector Operators in Clustering (2016). arXiv:1605.03093. https://doi.org/10.1002/mma.3963
26. Veshneva, I.V., Chernyshova, G., Bolshakov, A.A.: Regional competitiveness research based on digital models using Kolmogorov-Chapman equations. In: Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, vol. 333, pp. 141–154. Springer Nature Switzerland AG (2021). https://doi.org/10.1007/978-3-030-63563-3_12. ISSN 2198–4182
27. EU Regional Competitiveness Index (RCI) (2021). https://ec.europa.eu/regional_poli-cy/sources/docgener/work/201701_regional_competitiveness2016.pdf

# Development of a Decision Support System for Assessing and Monitoring Regional Competitiveness Risks

**Galina Chernyshova, Ekaterina Piunova, Irina Veshneva, and Gleb Rokakh**

**Abstract** Using of intelligent techniques in analyzing complex socio-economic systems makes it possible to expand the diagnostic and prognostic capabilities of a decision support systems. The implemented modules of the decision support system for assessing and predicting the risks' level of the regions' competitiveness were presented. The functionality of the application provides the use of both fuzzy-logical methods for a comprehensive assessment of the regions competitiveness, and tools for modeling the dynamics of risks based on Chapman–Kolmogorov equations. The developed methodology for applying these models was tested for 78 regions of the Russian Federation. As a result of the benchmarking analysis, the leading regions were identified both in terms of the current level of competitiveness risks and taking into account their dynamics.

**Keywords** Competitiveness assessment · Regions · Risks · Mathematical modeling · Kolmogorov–Chapman equations · Multi-criteria decision-making methods · Fuzzy techniques · Decision support system

## 1 Introduction

The concept of Society 5.0 is largely a strategic direction for long-term development, preservation or conquest of competitive positions for the economy both at the national and regional levels. The advantages of the Society 5.0 are associated with the possibility of creating an organized systemically balanced economy that provides solutions to a number of social problems through high-tech development. The development and implementation of intelligent information systems is important for diagnosing the level of regional development and the process of strategic support

G. Chernyshova (✉) · I. Veshneva · G. Rokakh
Saratov National Research University, 83, Astrakhanskaya Street, N.G. Chernyshevsky Saratov 410012, Russia
e-mail: cherny111@mail.ru

E. Piunova
OOO MetaPrime, St. Petersburg 197022, Russia

 
A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_13

for the competitiveness of regions. The solution of such complex tasks in conditions of incomplete information involves the use of fuzzy techniques, soft calculations, and structural dynamics control in the information processing. At the moment, there is no specialized software for assessing the competitiveness of regions.

The aim of the research is to create an application that provides decision support in assessing the risks of the competitiveness of regions using systemic multivariate analysis and forecasting, taking into account the specifics of the available statistical data for the regions.

At different levels of the economy, the concept of competitiveness continues to develop. The theoretical approach to the analysis of competitiveness at the regional level was initially based on the concept of the four stages of nations' development [1, 2]. The problem of empirical assessment of the regional competitiveness is under discussion in the accepted approach absence [3]. In particular, there are different techniques to identifying a system of factors that determine the complex concept of competitiveness at the regional level [4]. In the conditions of the knowledge economy, there is an expansion of the concept considering not only production costs but also a wide range of qualitative, innovative, institutional factors [5].

A common way to assessing regional competitiveness is based on indices (an indicator of global competitiveness (Global Competitiveness Index [6], European Regional Competitiveness Index [7], Global sustainable Competitiveness index [8]). In domestic practice, a comprehensive regional competitiveness index (AV RCI-2020) is used, it includes seven separate groups of indicators [9]. However, such approaches do not allow building predictive models and do not consider complex relationships between factors.

Selecting factors in studying regional competitiveness is particularly important, extensive set of indicators should be provided. At the same time, in the task of quantitative assessment, the question arises of the availability of their quantitative values. In addition, such factors are in a complex dynamic interaction, it is necessary to consider their mutual influence. This study makes it possible to ensure the accommodation of specified regional competitiveness features using relevant mathematical models implemented as software.

## 2 Models and Methods Implemented in the Application

The functionality of the developed software ensures the application of two models: a model based on multi-criteria decision-making (MCDM) and a model based on the Kolmogorov–Chapman equations.

Multi-criteria decision-making method Fuzzy ELECTRE II allows to use not only quantitative but also qualitative values of the criteria, to implement fuzzy numbers to make accurate and consistent decisions in order to reduce the subjectivity of the assessment [10, 11]. This method and its modifications are not based on classical utility theory; it contains relative estimates of alternatives. Binary relations of dominance are built on the alternatives set. In the ELECTRE method II, indices of consistency and inconsistency are defined, and there are several levels for these indices. This method ranks alternatives based on threshold levels of consistency and inconsistency.

The standard Fuzzy ELECTRE II algorithm uses one metric when evaluating alternatives, such as Euclidean distance or Hamming distance. The choice of metric when ranking alternatives is essential to this MCDM method. The proposed ranking algorithm involves various metrics (Euclidean, Manhattan, Hamming, Hausdorff, Zhuravlev, Tonimoto–Jaccard, Bray–Curtis) [12]. The final rating is based on the results obtained for different metrics.

Methods of system dynamics are used to assess the development of socio-economic objects [13]. The method under consideration, based on the Kolmogorov-Chapman equations, assumes at the initial stage the formation of risk factor set for the competitiveness of the regional socio-economic system. The deterioration of individual indicators corresponding to the identified factors leads to a decrease in the region's competitiveness as a whole. Such a system of indicators is a hierarchical logical-probabilistic model.

For the constructed tree of risk factors, it described a causal graph. A feature of such a graph is logical functions in the nodes of the consequences. Thus, each consequence occurs when a logical operation is performed for the corresponding causes—individual competitiveness risk factors. Graph edges represent possible transitions between system states, which are formed as a result of logical operations on causes. Based on this graph, it formed all combinations of critical events that form fragments of the graph for the implementation of the root event—the loss of the region's competitiveness.

The minimal cut is the set of leaf vertices of the causal graph, united on the basis of binary relations, and forming a subgraph of the causal graph, leading to the implementation of the root event. For each minimal cut, the states of the regional socio-economic system are identified, corresponding to various combinations of critical events that determine the risks of the regions' competitiveness. Then a network structure of possible transitions between these states is built. For the resulting structure of states, a differential equations system is determined [14]. As a result of solving the system of differential equations by numerical methods, it is possible to determine the probabilities of the realization of risks in the nodes of the state graph at given time.

For a minimal cut $s$ of dimension $N$, the set of states $S$ of the system under study contains $S = 2^N - 1$ elements. The system of Kolmogorov–Chapman differential equations can be expressed in matrix form:

$$\frac{dP_k(t)}{dt} = R \cdot \begin{pmatrix} P_0(t) \\ P_1(t) \\ \dots \\ P_S(t) \end{pmatrix}, \tag{1}$$

where $P_k(t), k = 1, \ldots, S$, is the probability that the elements of the system are in state $k$ at time $t$, $R = \{r_{lk}\}$ is the infinitesimal transition matrix of dimension $S \times S$, which is built on the basis of the state graph.

For state $k$, the differential equation has the form:

$$\frac{dP_k(t)}{dt} = \sum_{l=0}^{S} r_{lk} P_l(t) - \sum_{l=0}^{S} r_{kl} P_k(t), \tag{2}$$

where the elements $r_{lk}$ of the matrix R are correspond to the probabilities of transition from state l to state k, $r_{kk} = -\sum_{l,l \neq k} r_{kl}$.

The transition matrix R is formed on the basis of the Poisson parameters for the realization of events corresponding to the elements of the minimal cut: $d_i$ is the risk level, $g_i$ is the control action level for individual indicators, i = 1, …, N. The resulting system of differential equations is solved numerically using the Runge–Kutta method with automatic step adjustment.

The developed application includes the possibilities for users shown in the diagram (Fig. 1).

In the application, users are divided into two groups: administrators and analysts. The second category of users has access rights to work with the program provided by the application administrator.

For the first model, the functionality includes the formation of a set of regions and a set of evaluation criteria, ranking of the selected regions by the modified Fuzzy ELECTRE II method, presenting the result as a ranked list of regions and choosing the leading regions.

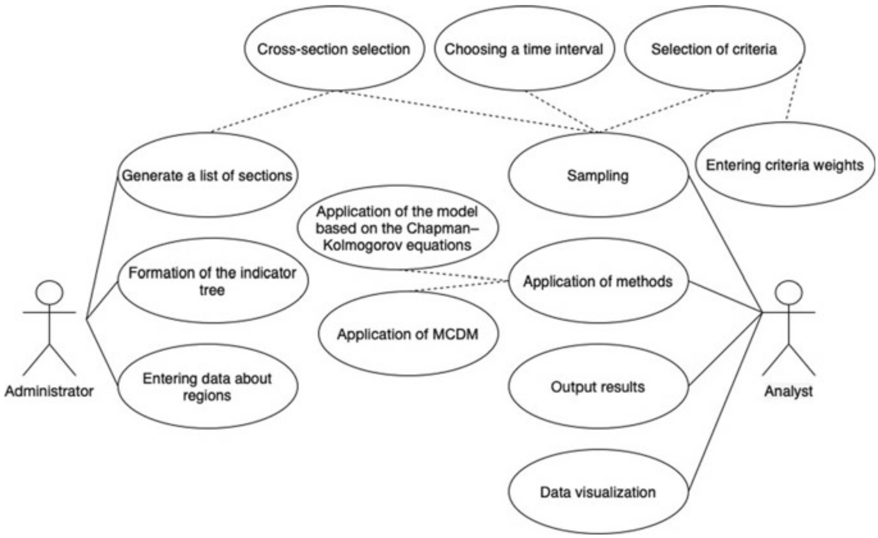



**Fig. 1** Functionality of the application for assessing regional competitiveness

For the second method, the developed application allows you to specify minimal cuts in the graph of cause-and-effect relationships, enter risk levels of control actions for individual indicators, while loading from a text file can be used. The functionality of the application provides visualization of the results. The application allows you to display probability graphs for critical events and their combinations for a given time interval. The calculated data corresponding to the probabilities for various combinations of critical events for the specified time intervals can be saved as a text file.

Web application is client–server, in which the user interacts with the server using a browser. Since the logic of such an Internet service is distributed between the server and the client, the data is stored on the server, and the information itself is exchanged over the network. To run a web application, the user does not need to install any additional software; it is enough to have access to the Internet and a browser on any device. The operation of the application does not depend on the operating system that is installed on the user's computer. Web services are easy to install, maintain, and upgrade the client interface. When changes are made, the application is updated to the latest version the next time the page is loaded.

Mostly, web applications consist of at least three main components: the client and server parts of the web application, as well as a database. The developed application uses the Python programming language extended by a set of libraries (Itertools, Matplotlib, PyPlot, Math, NumPy, SciPy, Tkinter). Python is a high-level scripting programming language that is known for its versatility. Python provides a set of tools with a large number of libraries including mathematical models and methods, various visualization tools.

PostgreSQL was chosen as the database. This database is an object-relational database management system that provides many advantages over other open source databases. A fundamental characteristic of an object-relational database is the support for user-defined objects as well as their behaviors, including data types, functions, operations, domains, and indexes. PostgreSQL is able to create, store and retrieve complex data structures. At the same time, it is a secure and extensible database, and also has an extensive ecosystem of tools available, which allows you to use PostgreSQL in a variety of scenarios.

The classic method for creating web interfaces is to use HTML (Standardized Document Markup Language) using the formal language for describing the appearance of a document, CSS. However, different implementations of HTML, CSS, and other specifications in different browsers cause problems in the development of web applications, and also complicates their subsequent support.

Among the many frameworks, Django was chosen to allow the development of a scalable application. The framework Bootstrap was chosen for front-end development, it is the freeware toolkit that allows you to create a web application. The selected tools allow implementing a scalable, modern and mobile application with an attractive user-friendly interface.

## 3 Methodological Approach to Assessing Competitiveness Risks

For a system analysis of competitiveness risks at the regional level, the following hierarchical structure of factors is used (Fig. 2): regional competitiveness ($E_0$); transformational ($E_1$); transactional ($E_2$); technical ($E_3$); social ($E_4$); natural resource ($E_5$); institutional ($E_6$); informational ($E_7$); innovative ($E_8$) [15]. At the next level, an extended set of indicators is applied: use of fixed assets ($E_9$); development of transport infrastructure ($E_{10}$); income level of the population ($E_{11}$); demographic ($E_{12}$); quality of life ($E_{13}$); mining ($E_{14}$); area of agricultural land ($E_{15}$); endowment with forest resources ($E_{16}$); electricity generation ($E_{17}$); environmental ($E_{18}$); share of unprofitable organizations ($E_{19}$); level of taxation ($E_{20}$); unfair competition ($E_{21}$); use of personal computers ($E_{22}$); use of the Internet in organizations ($E_{23}$); use of electronic document management systems in organizations ($E_{24}$); R & D personnel ($E_{25}$); internal spending on R & D ($E_{26}$); costs of innovative activities of organizations ($E_{27}$); volume of innovative goods ($E_{28}$); cost of fixed assets ($E_{29}$); fixed assets depreciation ($E_{30}$); railway track density ($E_{31}$); density of paved roads ($E_{32}$); average per capita money income of population ($E_{33}$); population with incomes below the subsistence minimum ($E_{34}$); the number of officially unemployed ($E_{35}$); life expectancy ($E_{36}$); coefficient of natural population growth ($E_{37}$); migration growth coefficient ($E_{38}$); improvement of living conditions of the population ($E_{39}$); medical personnel per 10,000 population ($E_{40}$); number of reported crimes ($E_{41}$); emission of pollutants into atmosphere ($E_{42}$); discharge of polluted sewage ($E_{43}$).

For a comprehensive analysis of the current level and dynamics of regional competitiveness, a methodological approach is proposed that allows, based on benchmarking, to identify best practices, create reference estimates of internal indicators, study and adapt the best regional development strategies [16]. This technique includes the implementation of the following steps based on the developed application.

Stage 1. Minimal cut set selection.

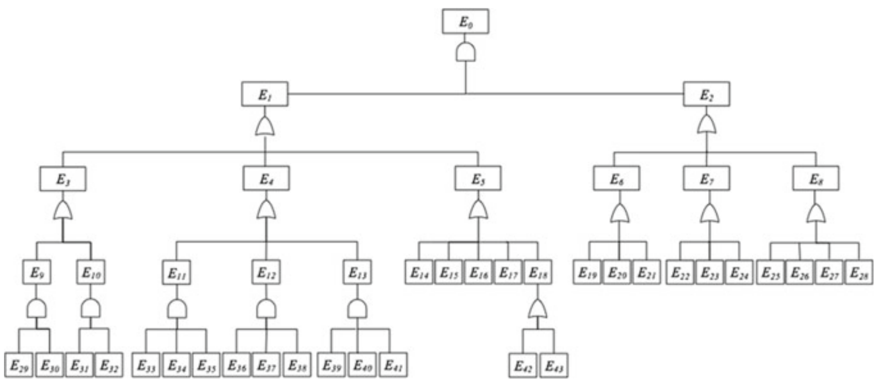


**Fig. 2** The structure of risk indicators of regional competitiveness

Stage 2. Application of the model based on the modified Fuzzy ELECTRE II to assess the current state of regional competitiveness.

Stage 3. Application of the model based on the Kolmogorov–Chapman equations to predict the risks of regional competitiveness.

Stage 4. Formation of the final assessment of the regions based on the distance method.

Let's consider a minimal cut $s$ that includes indicators $E_i^s,\ \ i = 1, \ldots, N$. There is a matrix $E_{ij}^s,\ \ i = 1, \ldots, N,\ \ j = 1, \ldots, M$, values of $N$ indicators included in the minimal cut $s$, for $M$ regions.

For a comprehensive assessment of the current level of regional competitiveness, the value $\overline{E_j^s},\ \ j = 1, \ldots, M$, was used by the modified Fuzzy ELECTRE II method.

The value $P_0(t)$ is used as a predictive assessment $\overline{\overline{E_j^s}}',\ \ j = 1, \ldots, M$, of the competitiveness risks of the region. It is obtained as a result of solving the formed system of the Kolmogorov–Chapman equations, for the region $j$, $j = 1, \ldots, M$, and the user-specified value $t$.

The distance method allows you to take into account the proximity of objects to the reference object. To estimate the distance $\overline{E_j^s}'\left(\overline{\overline{E_j^s}}'\right)$ from region $j$ to the leader region in terms of the indicator $\overline{E_j^s}\left(\overline{\overline{E_j^s}}\right)$, the region with the maximum (minimum) value of the indicator is used, taking into account its multidirectional impact on the final assessment:

$$\left(\overline{E_j^s}'\right) = \max(j = 1, M)\left(\overline{E_j^s}\right) - \left(\overline{E_j^s}\right), \quad j = 1, \ldots, M, \tag{3}$$

$$\overline{\overline{E_j^s}}' = \overline{\overline{E_j^s}} - \min_{j=1,M} \overline{\overline{E_j^s}}, \quad j = 1, \ldots, M \tag{4}$$

The integral assessment of the competitiveness of regions $\widetilde{E_j^s}$ is performed as follows:

$$\left(\widetilde{E_j^s}\right) = a_1\left(\overline{E_j^s}'\right) + a_2\left(\overline{\overline{E_j^s}}'\right), \quad j = 1, \ldots, M, \tag{5}$$

where $a_1, a_2$ are the weights coefficients that determine the significance of indicators $\overline{E_j^s}'$ and $\overline{\overline{E_j^s}}'$.

Stage 5. Risk analysis of regional competitiveness based on benchmarking.

## 4 Application of the Application for the Analysis of Regional Competitiveness of Regions

Consider the phased application of the proposed approach. In the process of testing the developed application for assessing regional competitiveness, the minimal cuts $s_1$ and $s_2$ were considered: $E_{19}$–$E_{20}$–$E_{21}$–$E_{28}$–$E_{30}$, $E_{25}$–$E_{26}$–$E_{27}$–$E_{29}$–$E_{30}$. Minimal cut $E_{19}$–$E_{20}$–$E_{21}$–$E_{28}$–$E_{30}$ includes material and technical risk factors: the share of unprofitable organizations ($E_{19}$); level of taxation ($E_{20}$); unfair competition ($E_{21}$); volume of innovative goods ($E_{28}$); cost of fixed assets ($E_{29}$); degree of depreciation of fixed assets ($E_{30}$). Minimal cut $E_{25}$–$E_{26}$–$E_{27}$–$E_{29}$–$E_{30}$ allows to evaluate such economic risk factors as the number of personnel involved in research and development ($E_{25}$); internal spending on research and development ($E_{26}$); costs of innovative activities of organizations ($E_{27}$); volume of innovative goods ($E_{28}$); cost of fixed assets ($E_{29}$); degree of depreciation of fixed assets ($E_{30}$).

These indicators are open data of regional statistics for 2020 [17, 18]. The study examined 78 regions of the Russian Federation, taking into account the available data on selected sets of indicators. Moscow is not included in the list of regions. According to most of considering indicators, this subject of the Russian Federation is an outlier. In the process of applying the multi-criteria method, discretization of the values of risk indicators of the regions' competitiveness was performed.

To demonstrate the proposed methodology for a comprehensive assessment of regional competitiveness, minimal cuts $s_1$ and $s_2$ and the corresponding models $M_1$, $M_2$, $M_3$, $M_4$, $M_5$, $M_6$ were considered (Table 1). For models based on the Kolmogorov–Chapman equations, the probability values $P_0(t)$ of competitiveness risks for $t = 3$ were calculated. Weight coefficients $a_1 = 0.4$, $a_2 = 0.6$ have been used for the integral assessment of regions' competitiveness.

Table 2 presents the results of modeling regional competitiveness according to $M_1, M_2, M_3, M_4, M_5, M_6$.

The assessment of the competitiveness of regions by the group of indicators of the minimal cut $s_1$ using the $M_1$ model allows you to create a ranked list of regions in terms of the current level of competitiveness in 2020. As leading regions, you can specify: the city of Saint-Petersburg city, Omsk region, Samara region, Moscow

**Table 1** Types of models for assessing the risks of regional competitiveness

| Model number | Minimal cut $s$ | Model |
|---|---|---|
| $M_1$ | $E_{19}$–$E_{20}$–$E_{21}$–$E_{28}$–$E_{30}$ | Model based on Fuzzy ELECTRE II |
| $M_2$ | $E_{19}$–$E_{20}$–$E_{21}$–$E_{28}$–$E_{30}$ | Model based on the Kolmogorov–Chapman equations |
| $M_3$ | $E_{19}$–$E_{20}$–$E_{21}$–$E_{28}$–$E_{30}$ | Integral assessment of regional competitiveness |
| $M_4$ | $E_{25}$–$E_{26}$–$E_{27}$–$E_{29}$–$E_{30}$ | Model based on Fuzzy ELECTRE II |
| $M_5$ | $E_{25}$–$E_{26}$–$E_{27}$–$E_{29}$–$E_{30}$ | Model based on the Kolmogorov–Chapman equations |
| $M_6$ | $E_{25}$–$E_{26}$–$E_{27}$–$E_{29}$–$E_{30}$ | Integral assessment of regional competitiveness |

**Table 2** Assessment of the competitiveness of Russian regions

| Region | $M_1$ | $M_2$ | $M_3$ | $M_4$ | $M_5$ | $M_6$ |
|---|---|---|---|---|---|---|
| | $\overline{E_j^s}$ | $\overline{\overline{E_j^s}}$ | $\widetilde{E_j^s}$ | $\overline{E_j^s}$ | $\overline{\overline{E_j^s}}$ | $\widetilde{E_j^s}$ |
| St. Petersburg city | 0.368 | 0.159 | **0.057** | 0.405 | 0.650 | **0.250** |
| Moscow region | 0.357 | 0.088 | **0.035** | 0.398 | 0.284 | **0.108** |
| Tatarstan | 0.285 | 0.218 | **0.131** | 0.393 | 0.657 | **0.260** |
| Krasnodar region | 0.201 | 0.177 | **0.165** | 0.382 | 0.684 | **0.277** |
| Sverdlovsk region | 0.172 | 0.061 | **0.136** | 0.367 | 0.585 | **0.247** |
| Krasnoyarsk region | 0.123 | 0.101 | **0.181** | 0.395 | 0.570 | **0.224** |
| Rostov region | 0.287 | 0.159 | **0.106** | 0.384 | 0.643 | **0.260** |
| Chelyabinsk region | 0.177 | 0.179 | **0.180** | 0.364 | 0.642 | **0.272** |
| Novosibirsk region | 0.301 | 0.192 | **0.111** | 0.369 | 0.673 | **0.281** |
| Samara Region | 0.359 | 0.063 | **0.024** | 0.352 | 0.340 | **0.158** |
| Bashkortostan | 0.230 | 0.016 | **0.083** | 0.326 | 0.052 | **0.058** |
| Yamalo-Nenets autonomous region | 0.352 | 0.055 | **0.026** | 0.138 | 0.347 | **0.289** |
| Nizhny Novgorod region | 0.125 | 0.148 | **0.199** | 0.371 | 0.644 | **0.268** |
| Tyumen region | 0.202 | 0.138 | **0.149** | 0.339 | 0.479 | **0.222** |
| Irkutsk region | 0.302 | 0.148 | **0.092** | 0.332 | 0.540 | **0.250** |
| Perm region | 0.099 | 0.049 | **0.175** | 0.331 | 0.361 | **0.179** |
| Belgorod region | 0.173 | 0.168 | **0.178** | 0.275 | 0.497 | **0.267** |
| Voronezh region | 0.219 | 0.063 | **0.108** | 0.387 | 0.413 | **0.166** |
| Leningrad region | 0.315 | 0.243 | **0.123** | 0.358 | 0.442 | **0.195** |
| Yakutia | 0.213 | 0.166 | **0.153** | 0.313 | 0.546 | **0.264** |
| Khabarovsk region | 0.245 | 0.111 | **0.112** | 0.318 | 0.639 | **0.298** |
| Stavropol region | 0.296 | 0.254 | **0.139** | 0.241 | 0.730 | **0.380** |
| Kemerovo region | 0.242 | 0.088 | **0.105** | 0.311 | 0.319 | **0.174** |
| Volgograd region | 0.347 | 0.070 | **0.034** | 0.259 | 0.360 | **0.222** |
| Omsk region | 0.362 | 0.148 | **0.057** | 0.314 | 0.808 | **0.368** |
| Primorsky region | 0.237 | 0.110 | **0.116** | 0.327 | 0.663 | **0.302** |
| Saratov region | 0.182 | 0.070 | **0.133** | 0.302 | 0.450 | **0.232** |
| Kaluga region | 0.175 | 0.196 | **0.188** | 0.347 | 0.627 | **0.276** |
| Kaliningrad region | 0.243 | 0.181 | **0.141** | 0.269 | 0.406 | **0.234** |
| Tula region | 0.250 | 0.033 | **0.078** | 0.321 | 0.025 | **0.050** |
| Orenburg region | 0.334 | 0.121 | **0.063** | 0.179 | 0.434 | **0.299** |
| Yaroslavl region | 0.131 | 0.207 | **0.219** | 0.261 | 0.716 | **0.363** |
| Udmurtia | 0.308 | 0.066 | **0.056** | 0.188 | 0.192 | **0.197** |
| Altai region | 0.325 | 0.112 | **0.064** | 0.232 | 0.649 | **0.353** |
| Lipetsk region | 0.164 | 0.122 | **0.165** | 0.211 | 0.489 | **0.302** |

(continued)

**Table 2** (continued)

| Region | $M_1$ | $M_2$ | $M_3$ | $M_4$ | $M_5$ | $M_6$ |
|---|---|---|---|---|---|---|
| | $\overline{E_j^s}$ | $\overline{\overline{E_j^s}}$ | $\widetilde{E_j^s}$ | $\overline{E_j^s}$ | $\overline{\overline{E_j^s}}$ | $\widetilde{E_j^s}$ |
| Sakhalin region | 0.194 | 0.064 | **0.124** | 0.210 | 0.604 | **0.349** |
| Ulyanovsk region | 0.105 | 0.279 | **0.263** | 0.234 | 0.801 | **0.413** |
| Tomsk region | 0.331 | 0.068 | **0.043** | 0.285 | 0.650 | **0.322** |
| Vologda region | 0.337 | 0.162 | **0.077** | 0.183 | 0.624 | **0.373** |
| Vladimir region | 0.335 | 0.179 | **0.085** | 0.316 | 0.663 | **0.309** |
| Murmansk region | 0.195 | 0.110 | **0.142** | 0.267 | 0.444 | **0.251** |
| Penza region | 0.147 | 0.192 | **0.203** | 0.266 | 0.626 | **0.324** |
| Astrakhan region | 0.109 | 0.120 | **0.197** | 0.203 | 0.545 | **0.329** |
| Kursk region | 0.157 | 0.043 | **0.137** | 0.213 | 0.313 | **0.230** |
| Tambov region | 0.192 | 0.140 | **0.155** | 0.173 | 0.598 | **0.369** |
| Komi | 0.285 | 0.046 | **0.062** | 0.246 | 0.318 | **0.213** |
| Ryazan region | 0.166 | 0.085 | **0.149** | 0.190 | 0.698 | **0.398** |
| Arkhangelsk region | 0.126 | 0.192 | **0.216** | 0.222 | 0.599 | **0.340** |
| Dagestan | 0.140 | 0.210 | **0.215** | 0.174 | 0.606 | **0.371** |
| Kamchatka region | 0.228 | 0.250 | **0.178** | 0.253 | 0.666 | **0.348** |
| Kirov region | 0.145 | 0.180 | **0.200** | 0.214 | 0.650 | **0.365** |
| Tver region | 0.275 | 0.290 | **0.165** | 0.239 | 0.966 | **0.476** |
| Amur region | 0.224 | 0.151 | **0.140** | 0.229 | 0.859 | **0.439** |
| Chuvashia | 0.191 | 0.050 | **0.120** | 0.162 | 0.312 | **0.261** |
| Karelia | 0.111 | 0.048 | **0.167** | 0.192 | 0.418 | **0.285** |
| Buryatia | 0.107 | 0.025 | **0.160** | 0.251 | 0.276 | **0.193** |
| Smolensk region | 0.171 | 0.047 | **0.131** | 0.189 | 0.485 | **0.314** |
| Zabaykalsky region | 0.121 | 0.144 | **0.200** | 0.221 | 0.512 | **0.305** |
| Bryansk region | 0.217 | 0.076 | **0.115** | 0.200 | 0.364 | **0.259** |
| Mordovia | 0.128 | 0.123 | **0.187** | 0.145 | 0.820 | **0.474** |
| Magadan region | 0.163 | 0.071 | **0.145** | 0.180 | 0.540 | **0.341** |
| Khakassia | 0.106 | 0.289 | **0.266** | 0.122 | 0.747 | **0.459** |
| Novgorod region | 0.320 | 0.180 | **0.094** | 0.195 | 0.700 | **0.396** |
| Ivanovo region | 0.093 | 0.119 | **0.206** | 0.159 | 0.532 | **0.350** |
| Mari El | 0.120 | 0.284 | **0.256** | 0.102 | 0.746 | **0.470** |
| Kostroma region | 0.235 | 0.205 | **0.155** | 0.121 | 0.678 | **0.432** |
| Oryol region | 0.297 | 0.244 | **0.134** | 0.144 | 0.825 | **0.477** |
| North Ossetia | 0.283 | 0.145 | **0.103** | 0.141 | 0.667 | **0.415** |
| Kabardino-Balkaria | 0.289 | 0.124 | **0.091** | 0.154 | 0.815 | **0.467** |

(continued)

**Table 2** (continued)

| Region | $M_1$ | $M_2$ | $M_3$ | $M_4$ | $M_5$ | $M_6$ |
|---|---|---|---|---|---|---|
| | $\overline{E_j^s}$ | $\overline{\overline{E_j^s}}$ | $\widetilde{E_j^s}$ | $\overline{E_j^s}$ | $\overline{\overline{E_j^s}}$ | $\widetilde{E_j^s}$ |
| Yamalo-Nenets autonomous region | 0.158 | 0.156 | **0.182** | 0.109 | 0.438 | **0.343** |
| Pskov region | 0.349 | 0.041 | **0.022** | 0.117 | 0.406 | **0.325** |
| Adygea | 0.208 | 0.090 | **0.126** | 0.116 | 0.919 | **0.531** |
| Kurgan region | 0.132 | 0.180 | **0.207** | 0.133 | 0.665 | **0.419** |
| Karachaev-Circassia | 0.292 | 0.122 | **0.088** | 0.135 | 0.378 | **0.303** |
| Altai | 0.259 | 0.133 | **0.112** | 0.132 | 0.479 | **0.346** |
| Tyva | 0.185 | 0.167 | **0.170** | 0.139 | 0.872 | **0.498** |
| Ingushetia | 0.110 | 0.110 | **0.193** | 0.120 | 0.571 | **0.390** |
| Kalmykia | 0.211 | 0.174 | **0.158** | 0.118 | 0.689 | **0.438** |

Bold indicates the final results of the integrated competitiveness risk assessment using both fuzzy methodologies and Kolmogorov-Chapman equations

region, Volgograd region, Yamalo-Nenets autonomous region, Pskov region, Vologda region, Vladimir region, Orenburg region.

The application of the $M_2$ model made it possible to single out the regions with the lowest risks relative to the group of indicators included in the $s_1$. The first ten regions include: Bashkortostan, Buryatia, Tula region, Pskov region, Kursk region, the Komi, Smolensk region, Karelia, Perm region, the Chuvashia.

In accordance with the final assessment based on the $M_3$ model, the following leading regions were identified for this category: Pskov region, Samara region, Yamalo-Nenets autonomous region, Volgograd region, Moscow region, Tomsk region, the Udmurtia, Omsk region, St. Petersburg city, Komi. The use of the estimates $\overline{E_j^s}'$ and $\overline{\overline{E_j^s}}'$, $j = 1, \ldots, M$ obtained as a result of using the $M_1$ and $M_2$ models, which determine the distance to the leading region, and the integral indicator $\widetilde{E_j^s}$, $j = 1, \ldots, M$, makes it possible to identify the features of regional development (Fig. 3). For example, the city of Saint-Petersburg city and the Omsk region lag far behind the leading regions in terms of predictive values of competitiveness risks. This led to the deterioration in the final assessment of these regions.

Let consider the minimal cut set $s_2$. The assessment of the competitiveness of regions by indicators of $s_2$ using the $M_4$ model allows you to create a ranked list of regions by the level of competitiveness in 2020. In this case, the following can be indicated as leading regions: the city of Saint-Petersburg, Moscow region, Krasnoyarsk region, Tatarstan, Voronezh region, Rostov region, Krasnodar region, Nizhny Novgorod region, Novosibirsk region, Sverdlovsk region.

Using the $M_5$ model, you can highlight the regions with the lowest risks relative to the group of indicators included in the minimal cut set $s_2$. The leading regions are: St. Petersburg, Moscow region, Krasnoyarsk region, Tatarstan, Voronezh region, Rostov

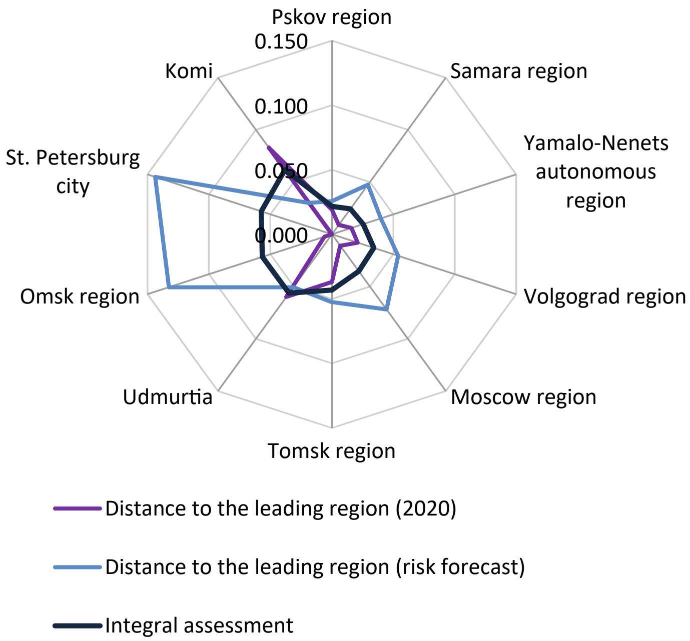



**Fig. 3** Distance to the leading region in accordance with the indicators of the minimal cut $E_{19}$–$E_{20}$–$E_{21}$–$E_{28}$–$E_{30}$

region, Krasnodar region, Nizhny Novgorod region, Novosibirsk region, Sverdlovsk region.

The final assessment using the $M_6$ model for the minimal cut set $s_2$ allows us to identify such leading regions as: Tula region, the Republic of Bashkortostan, Moscow region, Samara region, Voronezh region, Kemerovo region, Perm region, Buryatia, Leningrad region, the Udmurtia.

For benchmarking analysis, we present the distances from regions to reference samples in accordance with the models $M_5$ and $M_6$ (Fig. 4). For minimal cut indicators $s_2$, the lag behind the leading regions in terms of predictive risk levels is much smaller. It should be noted that in terms of the level of competitiveness in 2020 there is a slight lag in such relatively developed regions as the Leningrad region, Perm territory, Voronezh region.

As a result, it is possible to compare the estimates of the current level of competitiveness of the regions for the period 2020 and the forecast risk assessments with the position of the leading regions. It helps to identify the key problems of regional development.

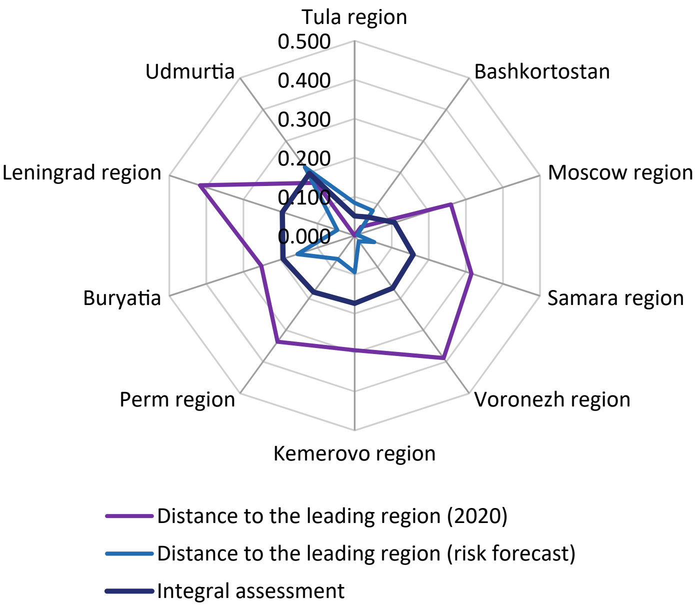



**Fig. 4** Distance to the leading region in accordance with the indicators of the minimal cut $E_{25}$–$E_{26}$–$E_{27}$–$E_{29}$–$E_{30}$

## 5 Conclusion

The approach implemented in the application involves an assessment according to certain sets of regional competitiveness indicators. Multi-criteria decision-making method Fuzzy ELECTRE II provides an estimation of the current state of the regional socio-economic system in a certain period. Methods for measuring risks dynamics based on the Kolmogorov–Chapman equations predict the level of regional competitiveness.

The proposed methodology based on the developed application makes it possible to study the experience of the regions, form benchmarks for comparing the performance of the regions and introduce best practices in the process of developing strategic directions for the development of the regions of the Russian Federation. A further direction for improving the decision support system for assessing the risks of regional competitiveness is the use of hybrid data mining models [19] for predictive assessment of competitiveness factors.

**Funding** The research was funded by RFBR, grant number 20-010-00465.

## References


1. Porter, M.E.: Competitive Advantage: Creating and Sustaining Superior Performance. Free Press, Riverside (1998)
2. Porter, M.E.: The economic performance of regions. Reg. Stud. **37**(6&7), 549–678 (2003)
3. 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)
4. Martin, R., Kitson, M., Tyler, P.: Regional Competitiveness, 176 p. Routledge (2006)
5. Sadki, T., Nmili, M., Jalal, C.: Regional competitiveness: a critical theoretical review and managerial and political implications. Revue Alternatives Manageriales et Economiques **2**(4), 194–216 (2020)
6. Global Competitiveness Report Special Edition 2020. https://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2020.pdf. Last accessed 28 March 2022
7. Regional Competitiveness Index. https://cohesiondata.ec.europa.eu/stories/s/Regional-Competitiveness-Index-2019/363v-4uq6. Last accessed 28 March 2022
8. Global Sustainable Competitiveness Index. https://solability.com/global-sustainable-competitiveness-index/the-global-sustainable-competitiveness-index-2021. Last accessed 28 March 2022
9. AV RCI-2020. http://lc-av.ru/wp-content/uploads/2020/05/AV-RCI-2020-alfa-200219.pdf. Last accessed 28 March 2022
10. Roy, B., Bertier, P.: La methode ELECTRE II: une method de classement en presence de critteres multiples. SEMA (Metra International), Direction Scientifique, Note de Travail, Paris, No. 142 (1971)
11. Sevkli, M.: An application of fuzzy ELECTRE method for supplier selection. Int. J. Prod. Res. **48**(12), 3393–3405 (2010)
12. Chernyshova, G., Piunova, E., Veshneva, I.: Analysis of Russian regions innovative competitiveness dynamics using fuzzy methods. In: Society 5.0: Human-Centered Society Challenges and Solutions. Studies in Systems, Decision and Control, vol. 416, pp. 115–128. Springer, Cham (2022)
13. Rezchikov, A.F., Kushnikov, V.A., Jandybaeva, N.V.: Model for assessing the state of national security in Russia based on the theory of system dynamics. Appl. Inf. **12**(2(68)), 106–117 (2017)
14. Veshneva, I., Chernyshova, G., Bolshakov, A.: Regional competitiveness research based on digital models using Kolmogorov–Chapman equations. Studies in Systems, Decision and Control, vol. 333, pp. 141–154 (2021)
15. Veshneva, I., Chernyshova, G.: The scenario modeling of regional competitiveness risks based on the Chapman−Kolmogorov equations. J. Phys. Conf. Ser. (JPCS) **1784**(1), 012008 (2021)
16. Atkinson, E.A., Bunker, R.D., Kaplan, R.S., Young, M.S.: Management Accounting, p. 874. Williams, Moscow (2005)
17. Regions of Russia. Socio-economic indicators, https://gks.ru/bgd/regl/b21_14p/Main.htm. Last accessed 28 March 2022
18. Operational statistics. http://bi.gks.ru/biportal/. Last accessed 28 March 2022
19. Chernyshova, G., Rasskazkin, N.: Software implementation of ensemble models for the analysis of regional socio-economic development indicators. Bull. Saratov Univ. New Series. Series: Math. Mech. Inf. **22**(1), 130–137 (2022)

# Mental Analysis of Stackelberg Leadership in Nonlinear Oligopoly Model for Cyber-Physical Systems of Oil Market

**Mikhail Geraskin**

**Abstract** A problem of the Stackelberg leadership analysis in an oligopoly model with nonlinear functions of market demand and firms' costs is considered. On the basis of trends in the dynamics of the oil market, the Cournot and Stackelberg equilibria are investigated in the case of one or several leaders. Because the oil market connects consumers, services supply equipment and management of the firms, the market is related to cyber-physical systems. In this case, the decision-making mechanism of the management is based on the analysis of mutual assumptions of the firms about the possible strategies of competitors, taking into account the technical characteristics of the equipment. We prove the following results. First, despite government regulation, the Russian oil market is characterized by a declining non-linear demand curve. Second, the modeling demonstrates that the Stackelberg equilibrium in the struggle for leadership game is closest to the actual market indicators, which allows us to conclude about the asymmetry of firms' awareness. Third, in the equilibria with the second level leaders, two attractors are found, i.e. the nonlinear model provides bifurcations.

**Keywords** Oligopoly · Nonlinear demand and cost functions · Cournot-Nash equilibrium · Stackelberg equilibrium · Oil market

## 1 Introduction

A linear model for choosing the optimal actions of the oligopoly market agents, in which the inverse demand function and the cost functions of agents are linear, widely used [1–5] to analyze equilibria in the game of oligopolists, because this model allows us a simple analytical solution in the form of the Cournot-Nash equilibrium [6, 7]. A nonlinear model for choosing optimal actions, in which the inverse demand function is linear, and the agents' cost functions are nonlinear, was used less often [8–11],

M. Geraskin (✉)
Samara National Research University, 34 Moscovskoe Shosse, Samara 443086, Russia
e-mail: innovation@ssau.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_14

because it does not allow us finding an analytical solution. In this case, special difficulties arise when calculating the equilibrium with the Stackelberg leadership [12].

As a consequence of the computational complexity of modeling, it is even more rarely that researchers focus on analyzing an oligopoly model with a nonlinear demand function. Meanwhile, actual markets often offer situations with nonlinear functions of firms' demand and costs. The recent studies shown interesting consequences of the nonlinearity.

In the case of a hyperbolic demand function, a bifurcation of the one fixed Cournot point was established [13], which is the resonant case of the Neimark-Saker bifurcation [14, 15], and when the fixed Cournot point is unstable, the stable cycles of any period can exist. An analysis of the hyperbolic demand function for the case of a duopoly [16, 17] on the phase portrait of firms' output demonstrated periodic cycles and chaotic attractors between which bifurcations arise. Additionally, in the case of an exponential demand function, the studies of the Australian electricity market proved a high degree of the price volatility [18]. In the case of a trigonometric demand function, the complex dynamics with two types of bifurcation is possible: the cascade of the double bifurcation and chaos and the Neimark-Saker bifurcation and, as a consequence, chaos [19]. The power fractional rational demand functions can lead to destabilization of the Nash equilibrium and bifurcations, which cause qualitative changes in the structure of attractors and their ranges [20]. An analysis of the hyperbolic demand function showed that there are many complex dynamic scenarios, and chaotic trajectories, as well as multistability, when various attractors coexist [21]. Interestingly, in the case of firms' demand functions linearly inhomogeneous in terms of outputs [22], the bifurcations of the equilibrium points were demonstrated, variations in model parameters lead to chaos, and loss of stability can be caused by period doubling bifurcations.

Our research has the following objectives. First, we formulate the situation on the oil market as a game model of an oligopoly. Second, we calculate the Cournot equilibrium in this game, as well as the Stackelberg equilibrium of the first and second levels based on formulas [23]. Third, we analyze the consequences of nonlinearity in terms of the possibility of multiple attractors.

As a rule, vertically integrated corporations are formed in the commodity markets. These companies form networks that consist of the mining equipment and the administrative staff (i.e., management). Consequently, these human–machine networks belong to the cyber-physical systems [24–26]. In this case, we investigate the cyber-physical system in terms of the influence of the mental management processes on such technical characteristics as the fuel transfer volumes.

## 2 Formulation of Problem

The model for choosing the optimal market strategies (sales volumes) of firms in the oligopoly market includes an objective function (firm's profit)

$$\Pi_i = P(Q)Q_i - C_i(Q_i), \quad Q_i \geq 0, \quad i = 1, \ldots, I, \tag{1}$$

the inverse demand function

$$P(Q) = AQ^{\alpha}, \; A > 0, \; \alpha < 0, \; |\alpha| < 1, \; Q = \sum_{i=1}^{I} Q_i, \tag{2}$$

and the power functions of agent costs

$$C_i(Q_i) = C_{Fi} + B_i Q_i^{\beta_i}, \; C_{Fi} > 0, \; B_i > 0, \; 0 < \beta_i < 2, \quad i = 1, \ldots, I, \tag{3}$$

where $Q_i$ and $\Pi_i$ are the sales volume and the profit of the *i*-th agent; $I$ is the number of market agents; $P(Q)$ is the inverse function of demand; $A$, $\alpha$ are the coefficients of the demand function; $B_i$, $\beta_i$, $C_{Fi}$ are the coefficients of the cost functions. The parameters of the cost functions have the following meaning: $C_{Fi}$ is the constant cost, $B_i$ is the marginal cost at $\beta_i = 1$, $\beta_i$ characterizes the return to scale effect, which can be positive ($0 < \beta_i < 1$) or negative ($1 < \beta_i < 2$).

In model (1)–(3), the Nash equilibrium is calculated by solving a system of necessary optimality conditions for a given vector of conjectural variations:

$$\frac{\partial \Pi_i \left( Q_i, Q'_{iQ_j} \right)}{\partial Q_i} = 0, \quad i, j = 1, \ldots, I, \tag{4}$$

where $Q'_{iQ_j}$ is the conjectural variation, i.e. the expected change in the sales volume of the *j*-th agent in response to a unit increase in the sales of the *i*-th agent.

In [23], the following formulas of system (4) were proved in various game situations.

1. The case of Cournot equilibrium, which corresponds to the unawareness of all agents (i.e. the followers) about the actions of the environment, and gives the zero conjectural variations $Q'_{iQ_j} = 0$:

$$AQ^{\alpha}\left[1 + \alpha \frac{Q_i}{Q}\right] - B_i \beta_i Q_i^{\beta_i - 1} = 0, \quad i = 1, \ldots, I \tag{5}$$

2. The case of equilibrium with the Stackelberg leadership. In this case, some agent (i.e. the leader) with a number $k$ believes that he is informed about the actions of the environment, i.e. there is an asymmetry of presumptive awareness in the market. It is also possible that there are several leaders who are informed that the rest of the agents are the followers. In this case, the mutual reaction of the leaders is not assumed, that is, in relation to each other, the leaders are the followers. As a result, system (4) for thee followers (whom we enumerate as $i = 1, \ldots, I_0$) has the form (5), and for the Stackelberg leaders (which we enumerate as $k = 1, \ldots, I_1$), Eq. (4) have the following form:

$$AQ^{\alpha}\left[1+\alpha\frac{Q_k}{Q}\left(1+\sum_{i=1,i\neq k}^{I}\frac{F_i}{u_i-F_i}\right)\right]-B_k\beta_k Q_k^{\beta_k-1}=0, k=1,\ldots,I_1, \tag{6}$$

where
$u_i = u_i(Q_i) = B_i\beta_i(\beta_i-1)Q_i^{\beta_i-2}$, $F_i = F_i(Q,Q_i) = A\alpha Q^{\alpha-1}\left(1+(\alpha-1)\frac{Q_i}{Q}\right)$, $I_1$ is a number of the first-level leaders.

3. The case of equilibrium with the two-level Stackelberg leadership. In this case, from the set of agents-leaders, the agents-leaders of the second level are distinguished (let us denote a number of these agents by $I_2$), informed that the remaining $I_1 = I - I_2 - I_0$ leaders choose actions according to reactions (6). We index the leaders of the second level by the symbol "$m$", where $m = 1, \ldots, I_2$. As a result, system (4) for the Cournot agents has the form (5), for the first level Stackelberg leaders has the form (6), and for the second level leaders, Eq. (4) have the following form:

$$AQ^{\alpha}\left[1+\alpha\frac{Q_m}{Q}\left(1-\sum_{k=1,k\neq m}^{I}\frac{F_k+A\alpha Q^{\alpha-2}Q_k\left[\Sigma_1+Q\frac{F^{/}_{mQ_m}u_m-W_mF_m}{(u_m-F_m)^2}\right]}{F_k[1+\Sigma_1]-A\alpha Q^{\alpha-1}[1+Q_k\Sigma_2]-W_k}\right)\right]$$
$$-B_m\beta_m Q_m^{\beta_m-1}=0,\quad m=1,\ldots,I_2, \tag{7}$$

where
$W_m = W_m(Q_m) = u'_{mQ_m} = B_m\beta_m(\beta_m-1)(\beta_m-2)Q_m^{\beta_m-3}$, $F^{/}_{mQ_m} = F^{/}_{mQ_m}(Q,Q_m) = A\alpha(\alpha-1)Q^{\alpha-2}\left(2+(\alpha-2)\frac{Q_m}{Q}\right)$, $\Sigma_1 = \Sigma_1\left(F_j,u_j\right) = \sum_{j=1,j\neq k}^{I}\frac{F_j}{u_j-F_j}$, $F^{/}_{iQ_k} = F^{/}_{iQ_k}(Q,Q_i) = A\alpha(\alpha-1)Q^{\alpha-2}\left(1+(\alpha-2)\frac{Q_i}{Q}\right)$, $\Sigma_2 = \Sigma_2\left(F_j,u_j\right) = \sum_{j=1,j\neq k}^{I}\frac{F^{/}_{jQ_k}u_j}{\left[u_j-F_j\right]^2}$.

We consider the problem of calculating the coefficients of the demand function (2), agents' cost functions (3) for the oil market, and on the basis of these data, we calculate equilibria (5)–(7) in cases 1–3 for various combinations of leaders and followers.

## 3 Results and Discussion

In the Russian oil market, the oil company Tatneft holds the largest market share, because other companies supply 2–5% of total oil sales to the domestic market, while Tatneft sells 35–50% in the domestic market. In 2019, the largest companies Rosneft, Tatneft and Gazpromneft held market shares of 24.5%, 35.1% and 19.6%, respectively, which together amount to 79.2%. Consequently, the characteristics of these companies adequately express the trends and factors of the state of the Russian oil market.

On the basis of the statistics from the largest firms in the Russian oil market (Rosneft, Tatneft and Gazpromneft), we study the volume of oil sales in the domestic market, prices and costs in 2009–2019 by quarters (44 periods). We calculate the weighted average price based on the firms' revenues. Using the least squares method, we define the regression models for demand function (2) and agents' cost functions (3) in the following form:

$$P(Q) = \begin{cases} 9236\,Q^{-0.91}, & t = 2009, \ldots, 2012, \\ 264196770\,Q^{-1.96}, & t = 2013, \ldots, 2019, \end{cases} \quad (8)$$

$$C_1(Q_1) = 1.61 Q_1^{1.13},\ C_2(Q_2) = 2.65 Q_2^{1.15},\ C_3(Q_3) = 1.04 Q_1^{1.15}, \quad (9)$$

where $t$ is the year, and the following indexation of agents is introduced: 1—Rosneft, 2—Tatneft, 3—Gazpromneft. The statistical characteristics of the regression models (8), (9) are high (Tables 1 and 2): the coefficients of determination of the demand functions are 0.66–0.71, for the cost functions these are 0.95, with the exception of the Tatneft function, for which this coefficient is 0.65; the calculated value of the Fisher criterion is at least 4 times higher than the table value at a significance level of 0.05. This confirms the adequacy and reliability of the developed models.

An analysis of the time series of weighted average prices and volumes of oil sales in the Russian market (Fig. 1) demonstrate that, despite the state regulation of the oil market, there are diminishing price trends. We identify two trends, the first of which describes the dynamics of oil prices in 2009–2012, and the second—in 2013–2019. Interestingly that for the latter time interval the slope of the inverse demand curve is higher, i.e. demand became less elastic (the price elasticity coefficient was 0.51) than in 2009–2012 (the price elasticity of demand was 1.01). In addition, the demand curve of the second interval passes above and to the right of the curve of the

**Table 1** Statistical estimates of regression models of demand functions

| Period | $R^2$ | Fisher criterion | |
|---|---|---|---|
| | | Table value ($\alpha = 0.05$) | Actual |
| 2009–2012 | 0.71 | 4.45 | 19.59 |
| 2013–2019 | 0.66 | 3.74 | 27.3 |

**Table 2** Statistical estimates of regression models of cost functions

| Firm | $R^2$ | Fisher criterion | |
|---|---|---|---|
| | | Table value ($\alpha = 0.05$) | Actual |
| Rosneft | 0.95 | 3.23 | 774.11 |
| Tatneft | 0.65 | 3.23 | 76.55 |
| Gazpromneft | 0.96 | 3.23 | 1 002.26 |

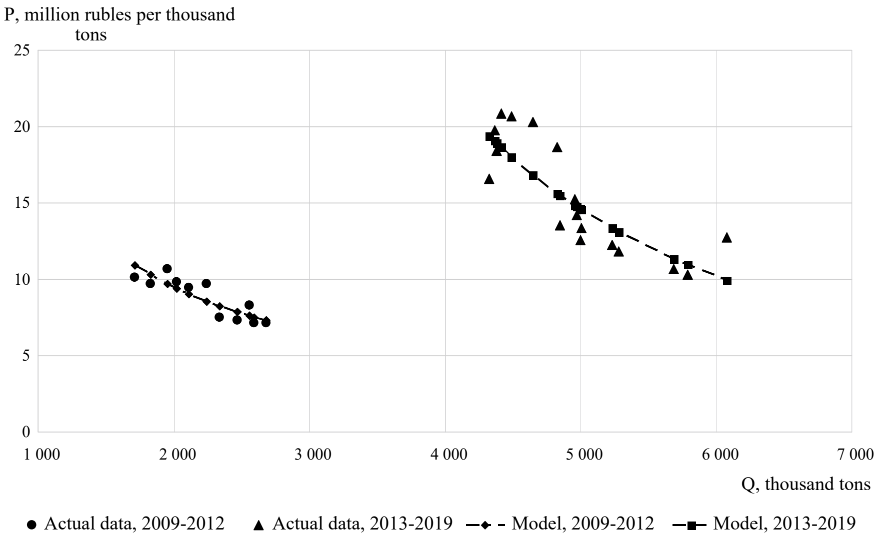



**Fig. 1** Dependences of weighted average prices on oil sales and regression models of demand functions

first interval, i.e. in 2013–2019 prices and sales volumes in the market were higher than in 2009–2012. Consequently, the market has expanded. When calculating the equilibria, we will use a regression model corresponding to the later interval.

The approximations of the agents' cost functions (9), presented in Figs. 2, 3 and 4 are increasing power functions with decreasing returns to scale ($\beta > 1$). A comparative analysis of the trends in agents' costs (Fig. 5) shows that in the range $Q \in [500, +\infty)$ the Tatneft's cost curve falls below the Gazpromneft's cost curve, which, in turn, is located below the Rosneft's curve. Consequently, Tatneft has a relative cost advantage, followed by Gazpromneft, while Rosneft loses to the first two firms in terms of costs.

The calculated Cournot equilibrium (Table 3) has a significant deviation from the actual market indicators, because the total sales volume in the equilibrium is 31% less than the actual volume in 2019, and the equilibrium price exceeds the actual prices of firms by 58–121%. Therefore, firms do not follow the classic follower strategies, i.e. there is an asymmetry of awareness. The Cournot equilibrium analysis demonstrates that the distribution of sales volumes corresponds to the relative cost advantage, i.e. model predictions are consistent with logic. In the actual distribution of the market, these firms are stratified by profit also in accordance with their relative cost advantage, although Rosneft is the second in market share and Gazpromneft is the third. This means that Cournot's model demonstrates a partial adequacy of reality, in part of the stratification of agents in terms of the maximum profit.

In the Stackelberg leadership struggle game (Table 4), if all three agents are the first-level leaders, the distribution of sales is closest to the actual distribution of the market. In this case, the deviation of the calculated value of the total sales volume

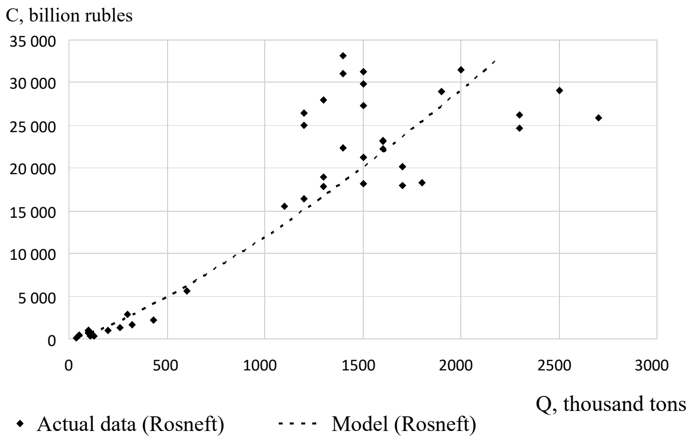



**Fig. 2** Dependences of Rosneft's costs on volume of oil sales and regression model

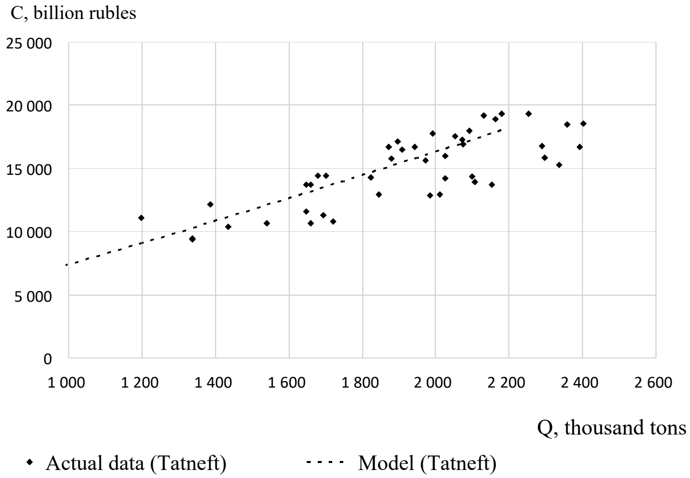



**Fig. 3** Dependences of Tatneft's costs on volume of oil sales and regression model

from the actual value in 2019 is 18%, the deviation in the equilibrium prices belongs to the range of 13–58%, the deviations in the revenues are 1%, 9% and 60% for the first, the second and the third agents, respectively. Consequently, the asymmetry of agents' awareness is expressed in the fact that the behavior of firms follows the principle of perceiving counterparties as the followers.

The most interesting results we obtain in the case with the second level leaders (Tables 5 and 6). We describe a game in which Tatneft is the second level leader, and Rosneft and Gazpromneft are the first level leaders, because this situation corresponds to the actual market. In this game, we search for two attractors, in each the second

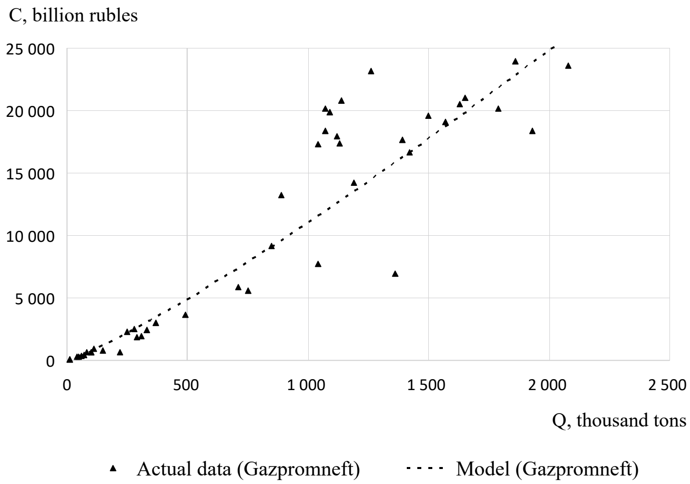



**Fig. 4** Dependences of Gazpromneft's costs on volume of oil sales and regression model

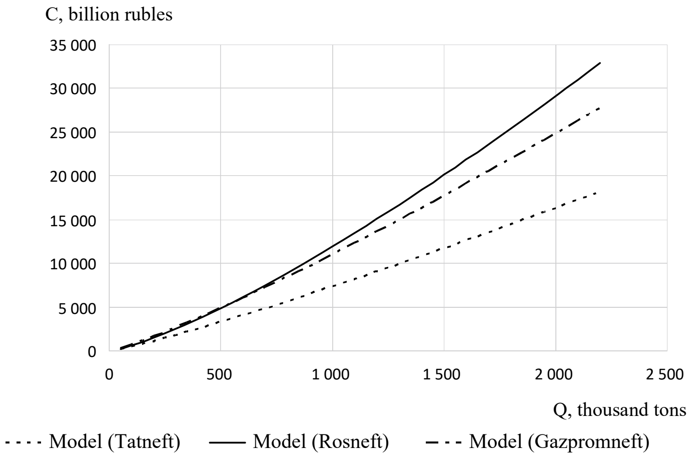



**Fig. 5** Comparative analysis of agents' cost trends

level leader (Tatneft) has an advantage in terms of sales and receives the greatest maximum profit in comparison with competitors. In the second attractor, the total sales of agents are lower, and the calculated equilibrium price is higher, as a result of which the profit of the second level leader is higher than in the first one. Consequently, in this case there are bifurcations of the trajectory of the groping for equilibrium. In addition, in the second attractor, the deviation of the calculated value of the total sales volume from the actual value in 2019 is 20%, the deviation in the equilibrium prices belongs to the range of 48–63%, the deviations in revenues are 20%, 1% and 69% for the first, second and the third agents, respectively. Therefore, the second

**Table 3** Actual and optimal indicators in the Cournot equilibrium

| Indicators | Agent | $Q_i$, thousand tons | Q, thousand tons | P, RUB mln per thousand tons | R, RUB mln | C, RUB mln | Π, RUB mln |
|---|---|---|---|---|---|---|---|
| Optimal | Rosneft | 933 | 3185 | 35 | 33,024 | 10,889 | 22,135 |
| | Tatneft | 1223 | | | 43,279 | 9283 | 33,996 |
| | Gazpromneft | 1029 | | | 36,393 | 11,378 | 25,016 |
| Actual | Rosneft | 1500 | 4629 | 22 | 33,600 | 31,213 | 2387 |
| | Tatneft | 2089 | | 16 | 33,468 | 16,780 | 16,688 |
| | Gazpromneft | 1040 | | 19 | 19,885 | 17,281 | 2603 |
| Deviation | Rosneft | −38% | −31% | 58% | −2% | −65% | 827% |
| | Tatneft | −41% | | 121% | 29% | −45% | 104% |
| | Gazpromneft | −1% | | 85% | 83% | −34% | 861% |

**Table 4** Actual and optimal indicators in the Cournot equilibrium Actual and optimal indicators in the equilibrium of the Stackelberg leadership struggle (three leaders of the first level)

| Indicators | Agent | $Q_i$, thousand tons | Q, thousand tons | P, RUB mln per thousand tons | R, RUB mln | C, RUB mln | Π, RUB mln |
|---|---|---|---|---|---|---|---|
| Optimal | Rosneft | 1319 | 3781 | 25 | 33,329 | 17,009 | 16,323 |
| | Tatneft | 1202 | | | 30,366 | 9096 | 21,278 |
| | Gazpromneft | 1260 | | | 31,858 | 14,441 | 17,419 |
| Actual | Rosneft | 1500 | 4629 | 22 | 33,600 | 31,213 | 2387 |
| | Tatneft | 2089 | | 16 | 33,468 | 16,780 | 16,688 |
| | Gazpromneft | 1040 | | 19 | 19,885 | 17,281 | 2603 |
| Deviation | Rosneft | −12% | −18% | 13% | −1% | −46% | 584% |
| | Tatneft | −42% | | 58% | −9% | −46% | 28% |
| | Gazpromneft | 21% | | 32% | 60% | −16% | 569% |

attractor of the game with two first level leaders and a leader of the second level also corresponds quite closely to the real market, like the aforementioned leadership struggle game with the first level leaders.

**Table 5** Actual and optimal indicators in equilibrium of game with two first level Stackelberg leaders and one leader of the second level (Tatneft), attractor 1

| Indicators | Agent | $Q_i$, thousand tons | Q, thousand tons | P, RUB mln per thousand tons | R, RUB mln | C, RUB mln | Π, RUB mln |
|---|---|---|---|---|---|---|---|
| Optimal | Rosneft | 2315 | 6637 | 8 | 19,398 | 35,148 | −15,750 |
| | Tatneft | 2970 | | | 24,888 | 18,705 | 6183 |
| | Gazpromneft | 1352 | | | 11,327 | 15,673 | −4346 |
| Actual | Rosneft | 1500 | 4629 | 22 | 33,600 | 31,213 | 2387 |
| | Tatneft | 2089 | | 16 | 33,468 | 16,780 | 16,688 |
| | Gazpromneft | 1040 | | 19 | 19,885 | 17,281 | 2603 |
| Deviation | Rosneft | 54% | 43% | −63% | −42% | 13% | −760% |
| | Tatneft | 42% | | −48% | −26% | 11% | −63% |
| | Gazpromneft | 30% | | −56% | −43% | −9% | −267% |

**Table 6** Actual and optimal indicators in equilibrium of game with two first level Stackelberg leaders and one leader of the second level (Tatneft), attractor 2

| Indicators | Agent | $Q_i$, thousand tons | Q, thousand tons | P, RUB mln per thousand tons | R, RUB mln | C, RUB mln | Π, RUB mln |
|---|---|---|---|---|---|---|---|
| Optimal | Rosneft | 2238 | 5537 | 12 | 26,760 | 33,642 | −6883 |
| | Tatneft | 2782 | | | 33,268 | 16,844 | 16,424 |
| | Gazpromneft | 517 | | | 6181 | 5079 | 1102 |
| Actual | Rosneft | 1500 | 4629 | 22 | 33,600 | 31,213 | 2387 |
| | Tatneft | 2089 | | 16 | 33,468 | 16,780 | 16,688 |
| | Gazpromneft | 1040 | | 19 | 19,885 | 17,281 | 2603 |
| Deviation | Rosneft | 49% | 20% | −47% | −20% | 8% | −388% |
| | Tatneft | 33% | | −25% | −1% | 0% | −2% |
| | Gazpromneft | −50% | | −37% | −69% | −71% | −58% |

## 4 Conclusion

We investigate the game models of the oligopoly with nonlinear functions of market demand and firm costs. We consider the following models of agents' behavior: the followers, i.e. the agents correspond to the Cournot's hypothesis, and the Stackelberg leaders of the first and second levels. The statistical analysis of the Russian oil market proves a decreasing non-linear demand curve. Consequently, it turned out to be possible to model the states of this market in the form of an aggregative game with nonlinear utility functions of agents.

The main difference between the equilibria calculated in this study in the case of a nonlinear demand curve and equilibria in the case of a linear dependence of demand [23] is manifested in the non-uniqueness of the game decisions under the conditions of the second level leaders. Because there are at least two attractors in the game, the possibility of the existence of other attractors is not excluded, when passing between them the bifurcations arise. Thus, modeling based on the actual market statistics confirmed the theoretical conclusions, which are derived earlier on the basis of model test parameters and described in the literature review.

## References


1. Karmarkar, U.S., Rajaram, K.: Aggregate production planning for process industries under oligopolistic competition. Eur. J. Oper. Res. **223**(3), 680–689 (2012)
2. Ledvina, A., Sigar, R.: Oligopoly games under asymmetric costs and an application to energy production. Math. Financ. Econ. **6**(4), 261–293 (2012)
3. Currarini, S., Marini, M.A.: sequential play and cartel stability in Cournot oligopoly. Appl. Math. Sci. **7**(1–4), 197–200 (2013)
4. Vasin, A.: Game-theoretic study of electricity market mechanisms. Procedia Comput. Sci. **31**, 124–132 (2014)
5. Sun, F., Liu, B., Hou, F., Gui, L., Chen, J.: Cournot equilibrium in the mobile virtual network operator oriented oligopoly offloading market. In: IEEE International Conference of Communications, ICC 2016, p. 7511340 (2016)
6. Nash, J.: Non-cooperative games. Ann. Math. **54**, 286–295 (1951)
7. Cournot, A.A.: Researches into the Mathematical Principles of the Theory of Wealth. Hafner, London (1960, Original 1838)
8. Naimzada, A.K., Sbragia, L.: Oligopoly games with nonlinear demand and cost functions: two boundedly rational adjustment processes. Chaos, Solitons Fractals **29**(3), 707–722 (2006)
9. Askar, S., Alnowibet, K.: Nonlinear oligopolistic game with isoelastic demand function: rationality and local monopolistic approximation. Chaos, Solitons Fractals **84**, 15–22 (2016)
10. Naimzada, A., Tramontana, F.: Two different routes to complex dynamics in an heterogeneous triopoly game. J. Differ. Equations Appl. **21**(7), 553–563 (2015)
11. Cavalli, F., Naimzada, A., Tramontana, F.: Nonlinear dynamics and global analysis of a geterogeneous Cournot duopoly with a local monopolistic approach versus a gradient rule with endogenous reactivity. Commun. Nonlinear Sci. Numer. Simul. **23**(1–3), 245–262 (2015)
12. Stackelberg, H.: Market Structure and Equilibrium, 1st edn. Translation into English, Bazin, Urch & Hill, Springer (2011, Original 1934)
13. Tramontana, F., Gardini, L., Puu, T.: Global bifurcations in a piecewise-smooth Cournot duopoly game. Chaos, Solitons Fractals **43**(1–12), 15–24 (2010)
14. Neimark, Ju.I.: On some cases of periodic motions depending on parameters. Dokl. Akad. Nauk SSSR **129**, 736–739 (1959)
15. Sacker, R.: On invariant surfaces and bifurcation of periodic solutions of ordinary differential equations. Report IMM-NYU №333. New York University (1964)
16. Askar, S.S., Alshamrani, A.M., Alnowibet, K.: Dynamic Cournot duopoly games with nonlinear demand function. Appl. Math. Comput. **259**(20834), 427–437 (2015)
17. Askar, S.S., Al-khedhairi, A.: On complex dynamic investigations of a piecewise smooth nonlinear duopoly game. Chaos, Solitons Fractals **139**, 110001 (2020)
18. Masoumzadeh, A., Nekouei, E., Alpcan, T.: Impact of a coal power plant closure on a multi-region wholesale electricity market. In: 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe, ISGT-Europe 2017—Proceedings 2018-January, pp. 1–6 (2018)

19. El-Wakeel, M.F., Askar, S.S., Alkhunayfir, N.A.: Toward investigation of the complex behavior of a monopoly game. J. Comput. Theor. Nanosci. **13**(11), 8552–8559 (2016)
20. Agliari, A., Naimzada, A.K., Pecora, N.: Nonlinear dynamics of a Cournot duopoly game with differentiated products. Appl. Math. Comput. **281**, 1–15 (2016)
21. Baiardi, L.C., Naimzada, A.: Imitative and best response behaviors in a nonlinear Cournotian setting. Chaos **28**(5), 055913 (2018)
22. Peng, Y., Lu, Q., Xiao, Y., Wu, X.: Complex dynamics analysis for a remanufacturing duopoly model with nonlinear cost. Physica A **514**, 658–700 (2019)
23. Geraskin, M.I., Chkhartishvili, A.G.: Modeling of oligopoly market structures with nonlinear functions of demand and costs of agents. Problemy Upravleniya **6**, 10–22 (2015)
24. Basinya, E., Rudkovskiy, A.: Automatic traffic control system for SOHO computer networks. Stud. Syst. Decis. Control **199**, 743–754 (2019)
25. Mammadov, J.F., Abdullaev, K.S., Agaev, U.H., Aliev, I.R., Huseynova, G.G.: Searching and selection of a flexible manufacturing system by means of frame model. In: Cyber-Physical Systems: Modelling and Intelligent Control. Studies in Systems, Decision and Control, pp. 139–152. Springer (2021)
26. Bolshakov, A., Veshneva, I., Lushin, D.: Mathematical model of integration of cyber-physical systems for solving problems of increasing the competitiveness of the regions of the Russian Federation. In: Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, pp. 129–139. Springer (2021)

# Society 5.0: Industrial Cyber-Solutions

# Planning the Optimal Railway Transport Load of a Mining Company

**Evgeny Eletin, Galina Borovkova, and Alexander Galkin**

**Abstract** The problem of planning of optimal loading of freight trains on a limited railway network—the network of a mining company—is considered in this chapter. The solution algorithm represents a combination of private scheduling problems solutions for a single locomotive, taking into account the constraints on the possibility of their joint movement. The chapter presents the problem statement, including the presence of several locomotives of different types, the solution algorithm and its application to the special case.

**Keywords** Integer linear programming · Railway planning problems · Theory of schedules

## 1 Introduction

The functioning of railway transport for mining companies directly affects the volume and continuity of mineral resource deliveries. The complexity of the task of planning the optimization of railway transport load depends on the size of the railway network, its closure, directions of cargo transportation, availability of a large fleet of railway vehicles, the need to transport different types of materials, etc. [1–5]. Although management engineering is now a mature discipline, it remains an area in which much research activity is conducted [6]. Optimization of rail transport load makes it possible to reduce the costs of freight transportation, create capacity reserves, and deliver goods within the specified time, which improves the efficiency of the railway network operation [7].

A number of automated systems have been created and are being improved for efficient transportation management in the industry [8–11]. The capabilities of the

E. Eletin (✉)
Corporate Solutions Center, 20 Sovetskaya, Lipetsk 398001, Russia
e-mail: eletin_ev@cscentr.com

G. Borovkova · A. Galkin
Lipetsk State Technical University, 30 Moskovskaya, Lipetsk 398055, Russia
e-mail: galkin_av@stu.lipetsk.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_15

network of information and computer centers are being increased [12, 13]. All this is aimed at making rational decisions on the use of the railway transport fleet [14]. The improvement of railway operation technology, and, mainly, the management of wagon fleet operation, is an important task for improving the efficiency of railway operation [15–17]. Therefore, the tasks of railway planning and optimization remain relevant and require development of new approaches to their solution [18–20].

The chapter presents two problem statements, the first involves several locomotives of the same type, the second involves locomotives of different types, as well as an algorithm for solving such problems and its application to the special case.

## 2 Problem Statement with One Type of Locomotives

In [16, 18] the formalization and solution of the problem of scheduling traffic for a single locomotive was considered. A formalization considering several locomotives with the same characteristics, transporting cargo between few stations is given below.

Let us consider the following notations. Let $h$ be the number of locomotive,

$h = 1,\ldots,m$, where $m$ is the total number of locomotives. Then we will say that locomotive $h$ is in the state $S^h\left(s^h, t^h, k^h\right)$, if at the moment of time $t^h \in T$ it is located at station $s^h$, and by this moment it has delivered $k^h = \left\{k^h_{ij}, i \neq j, i, j = 1, ..., n\right\}$ orders, where $n$ is the total number of stations, and $k^h_{ij}$ is the number of orders transported from $i$-th station to $j$-th.

There are also orders that determine the amount of cargo to be transported from the $i$-th station of departure to the $j$-th station of destination. The orders appear at certain moments of time and cannot be served earlier. It is necessary to make a schedule of locomotive movement that ensures fulfillment of all orders. The schedule is specified by the state of the locomotives at each moment of time.

The total time of completion of all orders by all locomotives in the last state (the moment of completion of all orders) is used as the criterion of optimality of scheduling:

$$F = \sum_{h=1}^{m} t^h. \tag{1}$$

It should be noted that in this section we consider the solution of the problem with several locomotives of the same type. In this case, the state ceases to be a vector [18] and becomes a matrix, where the $h$-th row is the state of the $h$-th locomotive. Due to the fact that locomotives of the same type are considered, we will consider states that include the same set of rows to be identical, as in Fig. 1, and we will glue them together as well. Here the states have the format $(s, t, k_{12}, k_{13}, k_{23})$.

In [16], a vertex gluing algorithm for solving the problem of forming trains and scheduling their movement for one locomotive was presented. This chapter presents

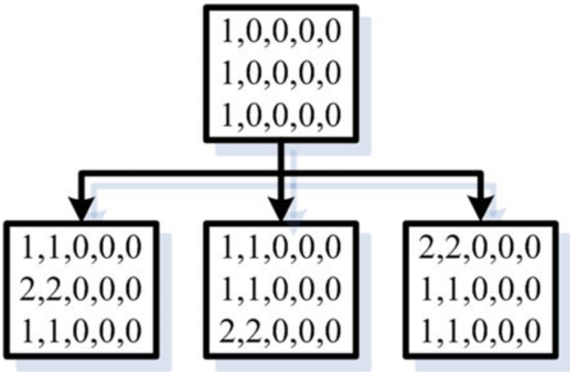



**Fig. 1** Example of identical states

the development of this algorithm for the case with several locomotives of the same type, as well as with several locomotives of different types. The algorithms themselves and examples of their use are presented below.

## 3 Algorithm for Locomotives of One Type

The tree of states is constructed by enumerating different locomotive motion variants. At each moment of time, taking into account the constraints [17], the locomotive can be in one of the following states.

1 Waiting is a condition when a locomotive is waiting for a new order at one station, several consecutive waits are allowed.
2 Idling is a condition when the locomotive is running empty, more than one consecutive idling is unacceptable.
3 Delivery is the state when the locomotive delivers the order from the point of departure to the destination.
4 Vacant is a condition when the locomotive is not functioning.

To reduce the dimensionality of the state tree the following algorithm of gluing its vertices is used.

Step 1. We add a new state to the tree and fix it as a matrix $S_j$, each row of which represents the state of an individual locomotive.

Step 2. We look for such a state among the constructed states, in which the set of rows coincides with the set of matrix rows $S_j$, we compare them line by line. That is, for any $k$-th row in matrix $S_j$ there is a similar $h$-th row in matrix $S_i$, $S_i^h = S_j^k$. If there is no such condition, go to step 1. Otherwise, go to step 3.

Then we find the parent state $S_{i-1}$, preceding the state $S_i$, and the state $S_{j-1}$, preceding $S_j$. Compare them by counting their values of the states $c_{i-1}$ and $c_{j-1}$ using the formula:

$$c_z = \sum_{h=1}^{m} \sum_{l,q=1}^{n} k_{z_{lq}}^{h} \tag{2}$$

where $z$ is the state number, equal to $(i-1)$ or $(j-1)$, $l$ is the departure station, $q$ is the destination station.

Then two options are possible.

Step 3a. If $c_{i-1} \geq c_{j-1}$, then stop building the branch containing the state $S_j$ and delete it. Move on to step 1.

Step 3b. If $c_{i-1} < c_{j-1}$, then delete the branch containing the state $S_i$, up to and including $S_i$, and transfer the states of the following levels to the state $S_j$. . Move to the construction of the next branch (step 1).

Formula (2) is the sum of all executed orders for a given state. We will consider more profitable the schedule in which more orders have been executed at a given moment of time.

## 4 Example with Several Locomotives of One Type

Let us consider an example of solving the problem of scheduling with three locomotives for three stations arranged in a circle. Let us use the notations introduced above. The layout of the vertices is shown in Fig. 2.

Let's set stations 1 and 2 as departure stations, 2 and 3 as destination stations, routes 12 (from station 1 to station 2), 13, 23. Orders arrive at departure station at time 1 and 3, the size of one order is 1 unit, the capacity of each locomotive is 2 units. The task is to find the schedule in which all orders will be executed in the shortest time.

Let us build a tree of states for this statement, using the algorithm of vertex gluing. The result is shown in Fig. 3. At the bottom after each schedule the value of the optimality criterion according to the formula (1) is calculated and the number of delivered orders for each route is given. As can be seen from Fig. 3, 16 possible schedules were obtained, the lowest value of the optimality criterion is 15, this result was obtained by waiting for orders until the locomotives are full.

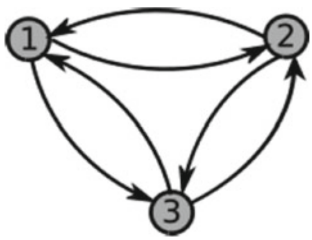



**Fig. 2** Scheme of vertex arrangement

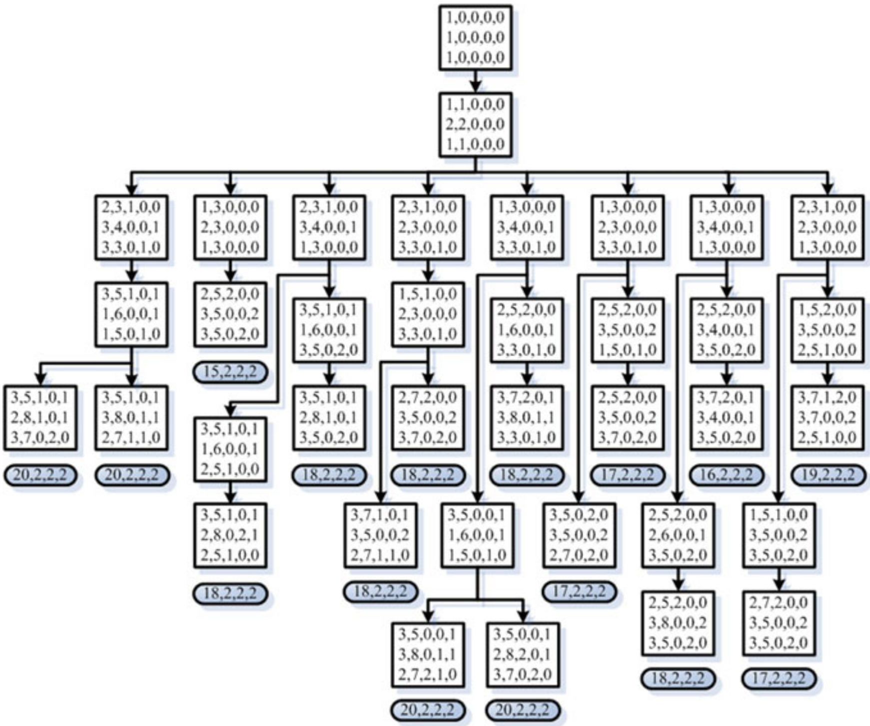


Fig. 3 State tree for three locomotives

## 5 Problem Statement with Different Types of Locomotives

Comparing to the previous formulation, when the problem contains locomotives with different capacities, the dimensionality of this problem increases and, as a consequence, the complexity increases too.

Let us consider the following notations. Let $h_l$ be the number of locomotives of type $l$, $l = 1,\ldots,L$, $L$ is number of locomotive types, $\sum_{l=1}^{L} h_l = m$ where $m$ is the total number of locomotives. Then let's consider that the locomotive $h_{l_g}$ is in the state $S^{h_{l_g}}(s^{h_{l_g}}, t^{h_{l_g}}, k_{ij}^{h_{l_g}})$, $h_{l_g} = 1, \ldots, h_l$, $g$ is the number of locomotive of a given type if at the moment of time $t^{h_{l_g}} \in T$ it is located at station $s^{h_{l_g}}$, and by this moment it has delivered $k_{ij}^{h_{l_g}}$ orders from $i$-station to $j$-station, $i \neq j$, $i$, $j = 1,\ldots, n$, where $n$ is the total number of stations.

There are also orders that determine the amount of cargo to be transported from the $i$-th station of departure to the $j$-th station of destination. The orders appear at certain moments of time and cannot be served earlier. It is necessary to make a schedule of locomotive movement that ensures fulfillment of all orders. The schedule is specified by the state of the locomotives at each moment of time.

The minimum time of completion of all orders by all locomotives, the minimum time among the maximum ones of completion of all orders in the last state (the moment of completion of all orders), is used as the criterion of optimality:

$$F = \min(\max_i t_{end}^i) \quad (3)$$

where $t_{end}^i$ is completion time of the $i$-th movement, $i = 1,\ldots, m$.

For this formulation, it should be noted that here the state of the transport system is described by a three-dimensional array, where each matrix is the state of all locomotives of the corresponding type. In addition, the algorithm for constructing the state tree is similar to the previous statement, differing in a large dimension due to the constraints given in [16]. The following modification of the vertex gluing algorithm is proposed for the developed optimal schedule search algorithm.

## 6 Algorithm for Different Types of Locomotives

Step 1. We add a new state to the tree and fix it as current $S^{h_{lg}}(s^{h_{lg}}, t^{h_{lg}}, k_{ij}^{h_{lg}})^{(cur)}$. It is a three-dimensional array, each two-dimensional slice of which is the state of all locomotives of the corresponding type $l$, i.e. a line in it describes the state of one locomotive.

Step 2. Find among the constructed states $S^{h_{lg}}(s^{h_{lg}}, t^{h_{lg}}, k_{ij}^{h_{lg}})^{(alt)}$, for which the condition is fulfilled: for any locomotive $g, d = 1, \ldots, h_l,\ l = 1,\ldots, L$

$$S^{h_{lg}} = s^{h_{ld}}, k_{ij}^{h_{lg}(cur)} \geq k_{ij}^{h_{ld}(alt)}.$$

If no such state is found, go to step 1. Otherwise, fix it as an alternative and go to step 3.

Step 3. Find the running times of locomotives in the current state and in the alternative state and choose the least of them. If $\max_i t^{i(cur)} > \max_i t^{i(alt)}$, then we delete the current route and proceed to construct the next branch (step 1). If $\max_i t^{i(cur)} = \max_i t^{i(alt)}$, then we compare the total running time of the locomotives: $\sum_{i=1}^{m} t^i$. We leave the route with less running time. If $\max_i t^{i(cur)} < \max_i t^{i(alt)}$, the route containing the state $(alt)$ is deleted and the current route is built (step 1).

Thus, we will consider more profitable the schedule in which not less orders are completed in less time. For cases where the same number of orders are completed in equal time, the route that has already been fully constructed will be given priority.

## 7 Example with Several Locomotives of Different Types

As an example, consider the railroad network from the example above, the station layout is shown in Fig. 1.

Let us suppose two types of locomotives with different capacities: one locomotive with capacity 1 unit and two locomotives with capacity 2 units.

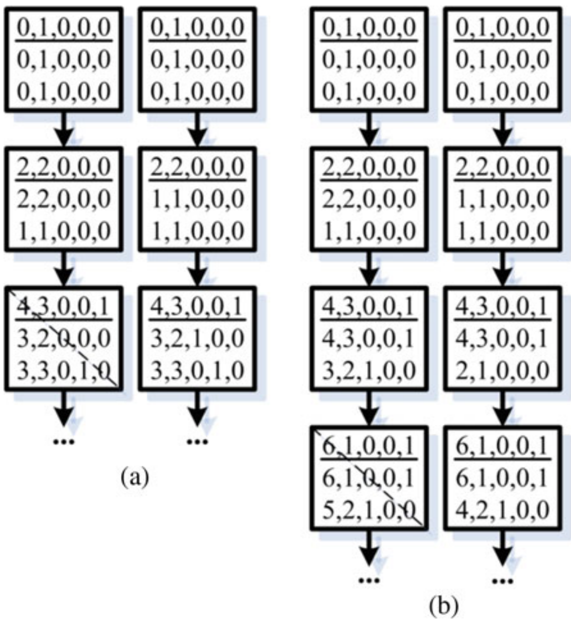
Let's set departure stations 1 and 2, destination stations 2 and 3, routes 12, 13, 23 as in example above. Orders arrive to departure station at time 1 and 3, the size of one order is 1 unit. It is required to find a schedule in which all orders will be executed in the shortest time.

The state tree for this problem will have a large number of branches, so below are examples of gluing nodes and examples of ready-made routes.

Thus, Fig. 4 shows examples of vertex gluing. The states have the format $(t, s, k_{12}, k_{13}, k_{23})$. The first line is the state of locomotive type 1, the second and third are the states of two locomotives of type 2. Under 4a is the case of gluing when different number of orders are transported in equal time, and under 4b is the case when the same state is reached in different time.

Figure 5 shows a fragment of the final schedule tree including the most optimal schedule. The value of optimality criterion, calculated by formula (3), and the number of delivered orders for each route are given at the bottom after each schedule. The smallest value of optimality criterion is 7, it was reached in several schedules, and we choose the one with less total time of locomotives operation. The first schedule has total running time equal to 16, and the second schedule has 12, in Fig. 5 it is

**Fig. 4** Examples of vertex gluing



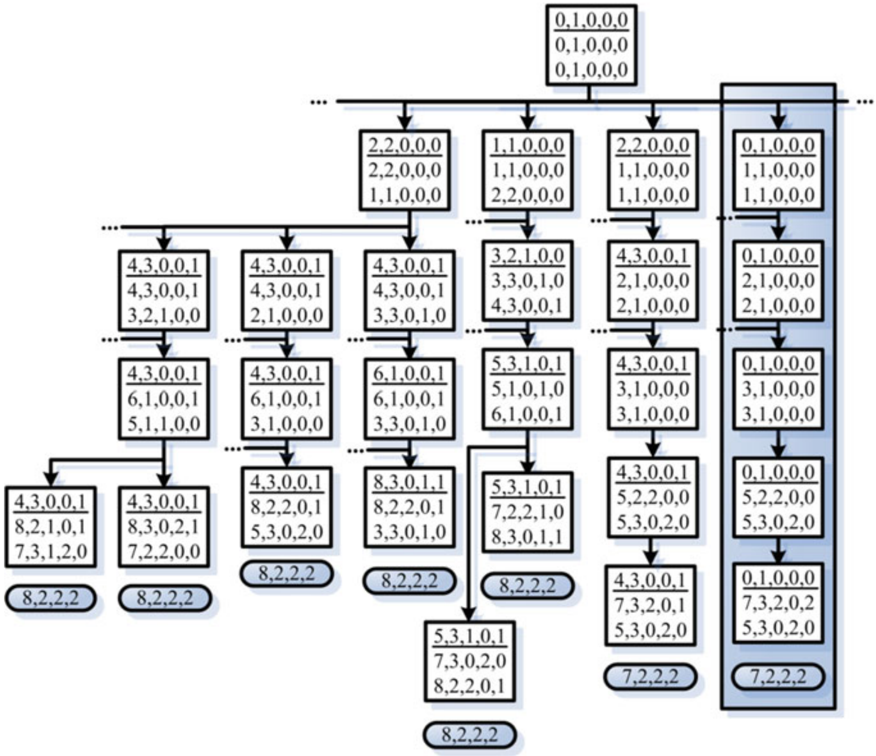



**Fig. 5** Fragment of the schedule tree

highlighted by a rectangle. In the process of construction, 24 pairs of routes are glued together.

## 8 Conclusion

Thus, the chapter presents two algorithms for planning the optimal loading of railway transport—for single-type locomotives and for locomotives of different types. The state of the system is described in the first case by a matrix, each row of which corresponds to a separate locomotive, and in the second case by a three-dimensional array, where matrices describe the states of the locomotives of corresponding types. The schedule represents the state of each locomotive at each moment of time. The algorithms for finding the optimal schedule are given. The solution is represented as a hierarchical structure corresponding to all possible state variants for each moments of time. The gluing of identical states according to optimality criteria reduces the dimensionality of the problem. We present the solution of a particular case of the described problem, namely, scheduling three locomotives between three marshalling stations which are located in the form of a circle. In the future, it is planned to use these algorithms as a basis for intelligent railway scheduling system.

## References

1. Yıldırım, M.S.: A management system for autonomous shuttle Freight Train Service in shared railway corridors. Int. J. Civ. Eng. **20**(2022), 273–290 (2022)
2. Singhania, V., Marinov, M.: An event-based simulation model for analysing the utilization levels of a railway line in urban area. Promet **29**(2017), 521–528 (2017)
3. Moeinaddini, A., Shafahi, Y., Hasany, R.M.: A simulation model for train movements in the rail network. Proc. Inst. Civ. Eng-Transp. **172**(2019), 152–163 (2019)
4. Bettinelli, A., Santini, A., Vigo, D.: A real-time conflict solution algorithm for the train rescheduling problem. Transp. Res. Part B Methodol. **106**(2017), 237–265 (2017)
5. Zhu, T., Mera, J.M., Suarez, B., Maroto, J.: An agent-based support system for railway station dispatching. Expert. Syst. Appl. **61**, 39–52 (2016)
6. Encyclopedia of Systems and Control 2nd ed. Edition. by John Baillieul (Editor), Tariq Samad (Editor). ISBN-13: 978-3030441852. ISBN-10: 3030441857 (2021)
7. Matyukhin V.G., Shabunin A.B.: Intelligent control systems on railway transport // Intelligent control systems on railway transport. (ISUZhT-2014): proc. M.: JSC NIIAS, 4–7 (2014)
8. Krmac, E., Djordjević, B.: An evaluation of train control information systems for sustainable railway using the analytic hierarchy process (AHP) model. Eur. Transp. Res. Rev. **9**, 35 (2017). https://doi.org/10.1007/s12544-017-0253-9
9. Shubinsky, I.B., Zamyshlaev, A.M.: Resource management of railway transport facilities. In: Technical Asset Management for Railway Transport. International Series in Operations Research & Management Science, vol. 322, pp. 113–140. Springer, Cham (2022).
10. Morozov V.N.: To enhance the efficiency and coordination of rail transport. Loading in ports. Zheleznodorozhn. Transp. **2** (2015)
11. Pekhterev, F.S., Shestakov, P.A.: Prospects for the development of freight traffic. Zheleznodorozhn. Transp. **10**, 2014 (2014)
12. Geiger M.J., Huber S., Langton S., Leschik M., Lindorf C., Tüshaus U.: Multi-attribute assignment of trains to departures in rolling stock management. Ann. Operat. Res. 271(2):1131–1163
13. Filina, V.N.: Development vectors of railway transport. Stud. Russ. Econ. Dev. **27**(2016), 400–411 (2016)
14. Shubinsky I.B., Zamyshlaev A.M.: Problems of managing technical assets in railway domain. In: Technical Asset Management for Railway Transport. International Series in Operations Research & Management Science, vol. 322, pp. 3–14. Springer, Cham (2022)
15. Tselsova A.Y., Khobotov E.N.: Development of the methodology of formation of routes and schedules of freight trains on the railway network // Intelligent control systems on the railway transport. (ISUZhT-2014): proc. The third scientific-technical conf. Moscow: JSC NIIAS, 11–14 (2014)
16. Eletin, E.V., Borovkova, G.S., Galkin, A.V.: Formalising a rail planning task for a mining company. Appl. Math. Cont. Sci. **2022**(1), 37–51 (2022). https://doi.org/10.15593/2499-9873/2022.1.02
17. Lazarev A.A., Musatova E.G., Kvaratskhelia A.G., Gafarov E.R.: Theory of Schedules. The Tasks of Control over Transport Systems, 160 p. Faculty of Physics of Lomonosov Moscow State University, Moscow (2012)
18. Eletin E.V., Borovkova G.S., Galkin A.V.: Solving the Task of Forming Trains and Their Schedule for Four Stations Using the Algorithm of Vertex Gluing // 3rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), Lipetsk, 10–12 Nov. 2021, pp. 1013 – 1016 (2021).

19. Chakraborty, S., Mengersen, K., Fidge, C.: A Bayesian Network-based customer satisfaction model: a tool for management decisions in railway transport. Decis. Anal. **3**, 4 (2016). https://doi.org/10.1186/s40165-016-0021-2
20. Zhaotong D., Wei L., Sheng L., Zheng C., Songqi W.: The research of two-train tracking system dynamic test based on train-train communication. In: Liang, J., Jia, L., Qin, Y., Liu, Z., Diao, L., An, M. (eds) Proceedings of the 5th International Conference on Electrical Engineering and Information Technologies for Rail Transportation (EITRT) 2021. EITRT 2021. Lecture Notes in Electrical Engineering, vol. 867. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-9909-2_17

# Intelligent System for Railway Joint Diagnostics

**N. Efimushkin, N. Efimushkina, and S. Orlov**

**Abstract** Improving rail transportation reliability is ensured by constant monitoring the rail track state. The chapter proposes a deep convolutional neural network for recognizing and classifying defects in rail joints based on rail images. Modern track-measuring cars are equipped with video cameras to receive rail track frames. However, operators' real-time processing of video data is impossible due to the extensive information volume and the high speed of the track measuring car. Rail joints are the elements most exposed to external influences, which causes a failure in their assembly. The structure of the diagnostic complex for monitoring rail tracks is described. An array of rail joint images is formed in the rail's video recording during the passage of the track measuring car. The formation of a set of classes of rail joints, including defective states and images without rails, is considered. This set includes regular rail joints with connectors, insulating joints, and welded joints. Additional classes are identified, corresponding to various anomalous configurations of rail joints in the images. A modified structure of a pre-trained deep convolutional network is proposed to classify rail joints and detect defective states. The actual images of rail joints were supplemented with artificial images obtained by affine transformations when preparing training samples. The process of training and testing the classifier based on the developed convolutional network is described. Transfer Learning is used to train the neural network. As a result of the experiments, rail joint classification accuracy and defect detection were at least 96%. It is proposed to use the method of accented visualization with augmented reality combined with an artificial neural network to train operators of track measuring cars to detect defects in rail fastenings and joints on video frames of a rail track.

**Keywords** Technical diagnostics · Cyber-physical system · Artificial neural network · Machine learning · Rail track · Splice-joints

N. Efimushkin (✉) · N. Efimushkina · S. Orlov
Samara State Technical University, 244 Molodogvardeyskaya Str., Samara 443100, Russia
e-mail: iefimushkin1999@mail.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_16

## 1 Introduction

Modern high-speed railways are cyber-physical systems consisting of interacting subsystems: rail track, rolling stock, energy system, information-measuring, and control systems [1, 2]. The key criteria for the operation of railways are reliability and profitability [3]. At the same time, the most significant problems arise in the field of intensive energy interaction of wheelsets of railway vehicles and the rail track.

The rail joint is the junction of two adjoining ends of the rails with a fastening. The rail joint must ensure the transition of the wheel from rail to rail with a minimum track pounding. The rail joint, the weakest part of the track, is constantly studied and improved. Defects can occur in rail joints during operation, leading to undesirable and sometimes catastrophic consequences.

Vibration-based diagnostics and statistical signal processing [4–7] and visual observation [8–11] are the most effective methods for determining the technical conditions of track elements.

In addition, the rail joint is involved in the formation and operation of the electrical rail circuit. The primary purpose of the track circuit is to detect the presence of rolling stock on the section of the track that forms the track circuit. Rail chains are the main element of automatic blocking used for various purposes in centralized traffic control devices, automatic crossing gates, and mechanized gravity marshaling yards [12]. The quality of the electrical rail circuit depends on the joint state and the value of its electrical resistance. In this regard, it is necessary to control the integrity of the rail joint structure.

There is a problem with the operational processing and analysis of track monitoring data. Visual control involves filming at high-speed hundreds of kilometers of rail for one passage of the track measuring car. The result is tens of thousands of images that operators cannot process in real-time. The solution to this problem is proposed using intelligent methods for processing big data arrays [13, 14]. In this regard, the control and diagnostics of the rail joint state using artificial neural networks is a focal problem.

## 2 Classification of Rail Joint Technical States

The diagnostic complex as part of a track measuring car was described in [14]. The complex includes four video cameras VC1–VC4 for continuous recording of the visual image of rails, an image preprocessing unit, and a decision-support system DSS on the technical state of the rail track (Fig. 1). Video recording data and image processing results are stored in the database DB.

The rail fastener classification and the identification of their defects were performed using a convolutional neural network and analyzed by the operator at the workstation WS. This diagnostic complex has already been used to control only rail fasteners.

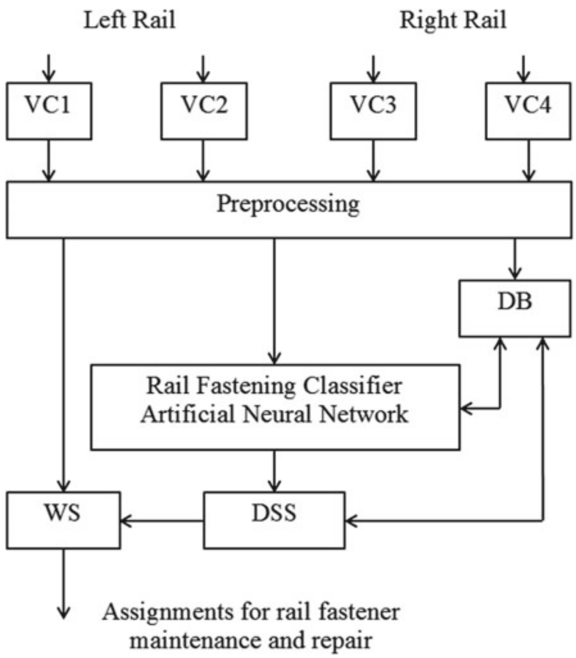



**Fig. 1** Diagnostic complex for track measuring car

Our chapter proposes modernizing the diagnostic complex by building a deep neural network to detect rail joint defects.

Three main types of rail joints are considered [15]:

- fish plate rail joint;
- insulating joint;
- welded joint.

The first type corresponds to the rail splice with the fishplates and fish bolts. The rail junction gap in such a joint can be zero. The fishplate is usually fastened with four or six bolts. There are two variants of the fished joint: without a connector and with a connector. Figure 2a, b show examples of the fished rail joints. The image of the insulating joint is shown in Fig. 2c, and the image of the welded joint is in Fig. 2d.

In this chapter, we consider two main types of defects that can occur in a typical joint with a connector:

- the connector is cut off at one end;
- the connector is completely cut off.

Examples of detectable defects are shown in Fig. 3a, b, respectively. In Fig. 3a, one of the ends of the connector is not fixed and hangs in the air. Traces of the connector are visible on the rail. In Fig. 3b, any connector elements are absent. In this case, at least one rail must have the connector cut end traces. It should be noted

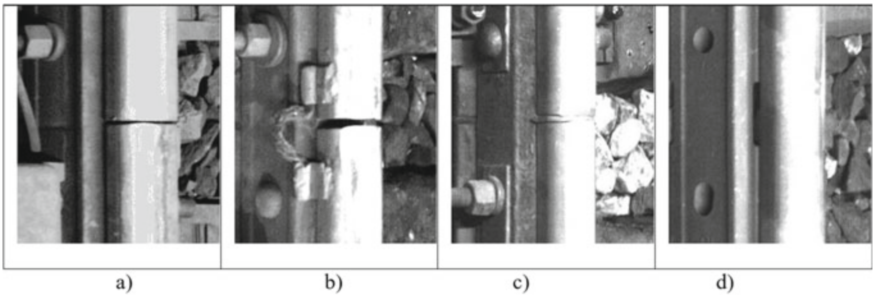


Fig. 2 Rail joint types: **a** fish plate rail joint without connector, **b** fish plate rail joint with connector, **c** insulating joint, **d** welded joint

that the presence of such traces on an insulating or welded joint is not a defect since connectors are not used in these joints.

For a complete description of the possible classes of images of rail joints, it is proposed to introduce two additional types of images: (a) **"fake_rail"**—rail without a joint, and (b) **"fake"**—an image without a rail. Examples of such situations are shown in Fig. 4. In the first case, a section of the road was fixed where there was no joint on the rail (Fig. 4a).

Such images can be obtained as the video is recorded continuously, and the distances between the rails may vary. The second option is related to the probability of hitting a rail track section with a turn in the frame. In this case, there may be no joint on the rails, or several rails in the frame may appear at once.

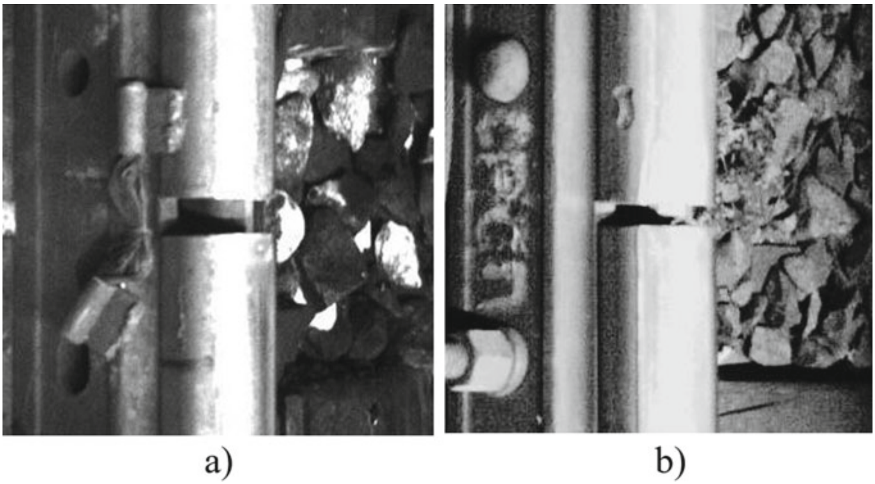


Fig. 3 Defective joints: **a** a fish plate rail joint with a connector cut off at one end, **b** a fish plate rail joint with an utterly cut connector

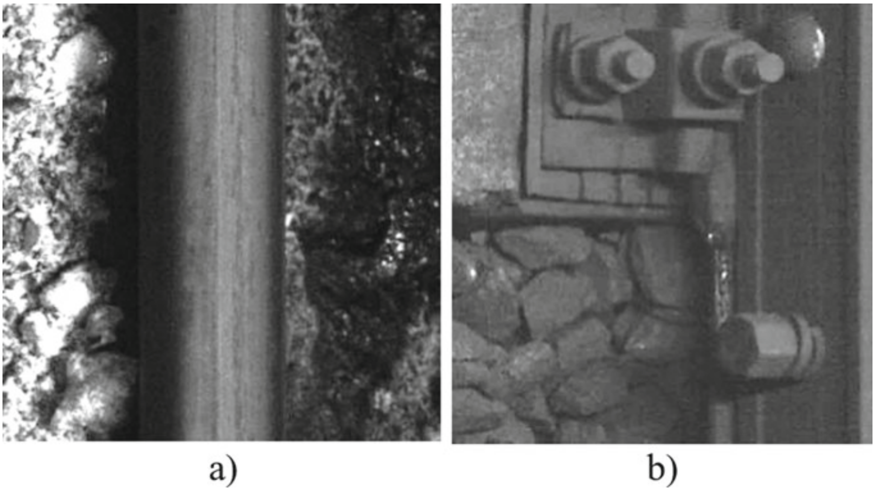



**Fig. 4** Additional image classes: **a** section of the rail without a joint, **b** image without a rail

The rail may be missing from the image in case of failures in the video camera. The camera may be initially incorrectly configured or knocked down during the movement of the track measuring car. Such frames often include fishplates, sleepers, fasteners, stones, and other objects (Fig. 4b).

Eight classes are proposed to identify the rail joint's type and technical state. Each of these classes describes specific characteristics of the rail junction. The image class names and their correspondence to the types and states of the joints are given in Table 1. The classes **"fake_rail"** and **"fake"** contain additional images obtained during video shooting. They have been described previously and are shown respectively in Fig. 4.

**Table 1** Image classes and types of rail joint states

| Class number | Class designation | Type and state of rail joint | Number of images in the class |
|---|---|---|---|
| 1 | conn | Fish plate rail joint with the connector | 7010 |
| 2 | no_conn | Fish plate rail joint without the connector | 5010 |
| 3 | iso | Insulating joint | 5024 |
| 4 | weld | Welded joint | 5008 |
| 5 | cut | The connector is torn off at one end | 1446 |
| 6 | missing | The connector is completely torn off | 5024 |

Video footage of the railway sections is recorded in special files and placed in the archive. Each video camera corresponds to one file, the name of which contains the identification numbers of the car, camera, section of the passage, the date and time of the shooting.

Thzn frames of black-and-white joint images with 300 × 300 pixels are formed. Frames contain the same information in their names as files. When a defect is detected, this allows determining where the problem with the rail joint occurred accurately. Frames are distributed into directories with class names describing the state of the joint and the defect. The database stores many images (up to several million frames.

## 3 The Structure of the VGG16 Convolutional Neural Network for Classifying the Rail Joint States

The rail joint classification system uses 300 × 300 pixel grayscale images. Each dataset image is a three-dimensional array, with the array length corresponding to the height, the number of elements in the second dimension to the width, and the number of elements in the third dimension to the quantity of the video recording channels. The image is reduced by 2 to obtain a frame with 150 × 150 dimensions. Each pixel array element in the frame is divisible by 255, which allows the pixel brightness values to be reduced to a range from 0 to 1.

It is advisable to use the Python programming language to develop a system for classifying the states of rail joints. The programming code was designed in the PyCharm development environment.

The most common tools for designing and training neural networks are Tensorflow and Keras libraries [16]. Tensorflow implements the concepts of the machine and deep learning, allows finding and classifying images, makes it possible to fine-tune any parameters of the network training process, and uses existing characteristics (loss functions available in the library, optimizers, and activation functions). Keras is a high-level Python language library that runs on top of TensorFlow. Keras helps solve image classification problems using convenient tools for setting up network layers, training parameters and characteristics, presenting structures and results.

A network based on pre-trained convolutional network VGG16 was designed to solve the states of rail joint classifying [17]. The network was trained on many images from the ImageNet website: animals and plants, houses, cars, and other objects.

The architecture of the neural network used to classify rail joints is shown in Fig. 5. The network model contains layers: InputLayer, Conv2D (convolution layer), MaxPolling2D (maximum pooling layer), BatchNormalization, Flatten, Dense (fully connected), and Activation. Arrows in the diagram indicate the order of the layers. Each block in Fig. 4 contains information about the name and type of the layer. The upper right field of the block contains the input array dimensions, the lower right field—the output array dimensions.

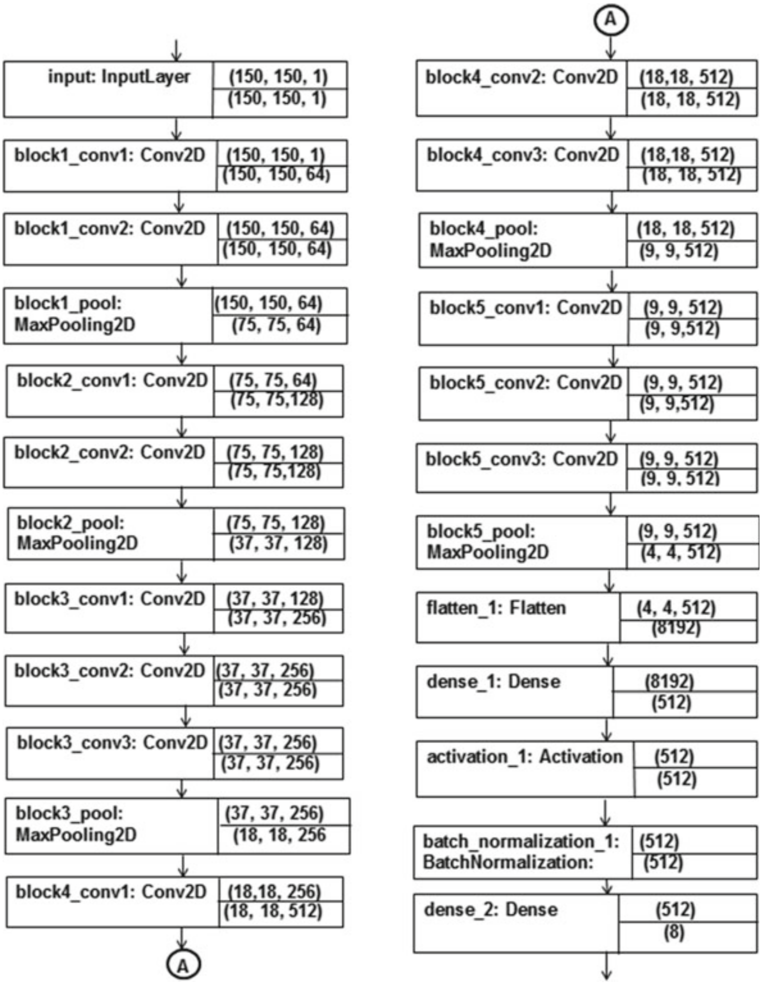



**Fig. 5** The structure of the rail joint classifier based on the neural network VGG16

The input layer of the neural network VGG16 receives arrays of 150 × 150 × 3 pixels. Since grayscale frames of 300 × 300 are used, changing the images and the network's input layer is required. The input layer dimension is changed, and the following weights are adjusted. The architecture of the VGG16 convolution section remains unchanged. The modified weights are loaded into the second layer of the network, and the original weights are loaded into the remaining layers.

The classifier in the original VGG16 network has been completely replaced with a classifier of our design (the last four blocks in Fig. 5).

This is due to the training method used, which is described below. The classifier contains two fully connected layers: activation and normalization layers. The initial weights in the pre-trained VGG16 section of the convolutional network are not saved since the classifier of our design was prepared for the new problem being solved, and its weights must be obtained by retraining.

Each image with some probability belongs to a specific class. The output layer of the neural network contains 8 neurons, with each of the neurons corresponding to a specific class in Table 1. The output neuron signal is equal to the probability that an image belongs to a given class after passing through the neural network.

As part of developing a deep neural network, one of the varieties of gradient descent, Adam (Adaptive Moment Estimation) [18], was used. This method, in particular, calculates the exponential moving average of the gradient and the quadratic gradient. The weight update rule for the Adam method is determined based on the gradient estimates and the second moments of the gradients, which are given by the expressions:

$$m_w^{(t+1)} \leftarrow \beta_1 m_w^{(t)} + (1-\beta_1)\nabla_w L^{(t)},\ \ v_w^{(t+1)} \leftarrow \beta_2 v_w^{(t)} + (1-\beta_2)(\nabla_w L^{(t)})^2,$$

where $w$ are parameters; $L^{(t)}$ is the loss function; $t$ is the index of the current iteration, $\beta_1$ and $\beta_2$ are the forgetting coefficients for the gradients and the second moments of the gradients, respectively.

The calculated moments are corrected according to the following formulas:

$$\hat{m}_w = m_w^{(t+1)} / \left(1 - \beta_1^{(t+1)}\right),\ \ \hat{v}_w = v_w^{(t+1)} / \left(1 - \beta_2^{(t+1)}\right).$$

The weights are recalculated according to the formula:

$$w^{(t+1)} \leftarrow w^{(t)} - \eta \hat{m}_w / \left(\sqrt{\hat{v}_w} + \delta\right),$$

where $\eta$ is the step size or learning rate; $\delta$ is small to prevent division by zero.

## 4 Forming Train Dataset and Training a Neural Network

It was decided to use the supervised learning mode to train the developed neural network. For this, 90% of the images should be placed in the training dataset and 10% in the test dataset. The program code sequentially reads and prepares for training images from the input array. The first 90% of the processed images belong to the training set and the rest to the test set.

Since there is a small number of source images (less than 10,000 for the most frequently occurring class), data augmentation is the best to increase the feature for the training of the learning models. The following affine transformations create new frames:

- vertical rotation of the original frame (rotation by 180 degrees),

- decrease and increase the brightness of the frame (maximum by 70%),
- rotate 2 degrees in a random direction relative to the center of the frame,
- image zoom (maximum by 10%).

Such augmented images are generated only for the training sample.

Loss estimation is performed using categorical cross-entropy. A value of $10^{-5}$ was chosen to increase the learning rate. The learning process monitors the change in classification accuracy and loss relative to the step number. The neural network maximum training period was limited to 100 epochs. If within five steps the model's classification accuracy has improved by less than 0.001, then the network training stops.

The transfer learning method is used for training the neural network [19]. It involves using the original convolutional section of the pre-trained network and the developed section of the classifier. The weights of the convolution section are taken from VGG16 and "frozen" (remain unchanged during training). Then the learning process does not affect the convolutional section. On the contrary, the own classifier weights are adjusted during network training.

The accuracy of the classifier after training is evaluated using two metrics Precision and Recall [20]. The Precision metric or Positive Predictive value (*PPV*) shows the fraction of the relevant rail joints classified as belonging to a given class:

$$PPV_K = TP_K/(TP_K + FP_K),$$

where $K = \overline{1, 8}$ is rail joint class number, $TP_K$ are true positive predictions, $FP_K$ is a false positive, i.e., test results that erroneously indicate that a rail joint belongs to class K.

The Recall metric or True Positive Rate (*TPR*) demonstrates the classifier's ability to detect a given class of rail joints in the set of all analyzed images:

$$TPR_K = TP_K/(TP_K + FN_K),$$

where $FN_K$ is a false negative, i.e., test results that erroneously classify rail joints as not belonging to class K.

The results of training the classifier on the VGG16 network for classes "**cut**" and "**missing**" are shown in Figs. 6 and 7. For the considered dataset, training ended after 8 epochs.

Sometimes it makes sense to evaluate both metrics using an aggregated quality criterion in the form of an $F_1$- score, which is the harmonic mean of the two values Precision and Recall:

$$F_1 = 2\frac{PPV \times TPR}{PPV + TPR}.$$

The results of neural network training for all classes of rail joints are shown in Table 2.

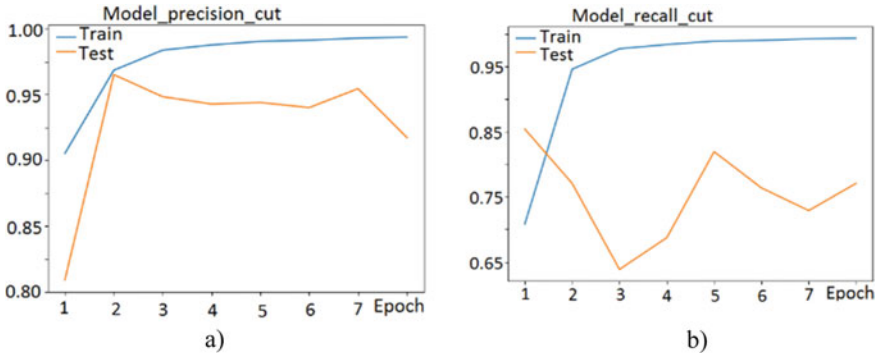



**Fig. 6** Class "**cut**": **a** *Precision*; **b** *Recall*

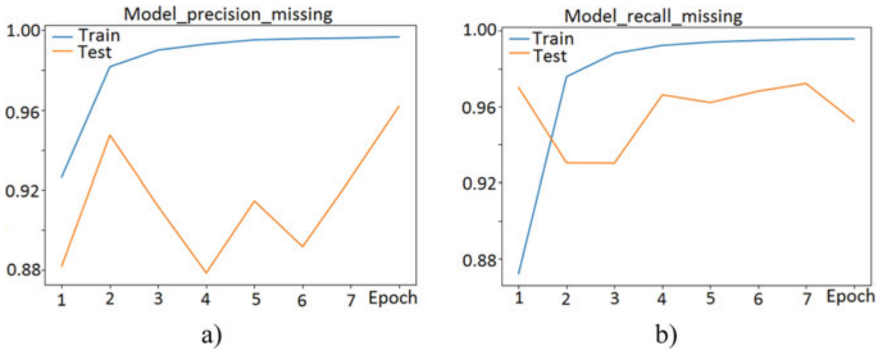



**Fig. 7** Class "missing": **a** *Precision*; **b** *Recall*

**Table 2** Accuracy metrics for rail joint classification

| Metrics | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 | Class 6 | Class 7 | Class 8 |
|---|---|---|---|---|---|---|---|---|
| *PPV* | 0.96 | 0.98 | 0.91 | 0.98 | 0.93 | 0.95 | 1.00 | 0.99 |
| *TPR* | 0.98 | 0.90 | 1.00 | 1.00 | 0.78 | 0.95 | 0.98 | 0.95 |
| $F_1$ score | 0.97 | 0.94 | 0.95 | 0.99 | 0.85 | 0.95 | 0.99 | 0.97 |

The biggest problem is separating the classes "**conn**" and "**cut**." The reason for this is the imbalance of the source dataset. There are five times fewer images of joints with a **"cut"** defect than frames with serviceable **"conn"** type rail joints. We created additional **"cut"** images according to the abovementioned method and received 7000 images. A new training of the neural network VGG16 was carried out, and the results for the **"cut"** class were obtained: $PPV = 0.97$; $TPR = 0.93$ and $F_1\ score = 0.95$.

Table 3 contains general data on the accuracy metrics of the developed classifier based on a deep neural network.

**Table 3** General statistics of the neural network classifier

| Metric | Accuracy | Loss | Precision | Recall |
|---|---|---|---|---|
| Mean | 0.963 | 0.199 | 0.964 | 0.962 |

Thus, the developed neural network can effectively classify the states of rail joints in general with an accuracy of 96%.

## 5 Using Accent Visualization and the Rail Joint Classifier to Train Track Measuring Car Operators

Modern technologies of augmented reality (AR) are used in industrial applications of cyber-physical systems to build interactive user interfaces. The main advantage of AR interfaces is their ability to place virtual objects in the same space and time as real objects. AR technology allows the development of interactive and context-sensitive user interfaces, which should provide computer vision and object recognition [21, 22].

To solve the problem of technical diagnostics of the rail track and other elements of the railway infrastructure using neural networks, high-quality and representative image samples for machine learning are required. To this end, it is necessary to train the track measuring car operators, who form training samples by identifying anomalous situations on the obtained images. This requires preliminary training of staff on typical examples. It is proposed to form an image set of rails with augmented reality, train operators to search for defective rail joints and fasteners, and then use the obtained data as training and test datasets for a neural network.

During his training, the operator of the track measuring car may be overloaded with a large amount of information, and some real objects that require his attention may overlap with virtual ones in one scene. An adaptive user interface has been built to solve this problem and improve the usability and efficiency of AR devices in training operators on the classification of rail joints. Identification of rail track elements and their classification is implemented using an artificial neural network.

## 6 Conclusion

The chapter presents a deep neural network VGG16, designed to classify rail joints and identify defects. The neural network operates as part of a rail track diagnostic complex that processes the rail images during the movement of a track measuring car. A modified architecture of a pre-trained deep network was proposed to improve decision-making efficiency, increasing the rail joint classification accuracy. The developed neural network training method is based on dataset balancing and the transfer learning method. Experimental studies have shown the possibility of achieving a classification accuracy of at least 96%.

Further development of the intelligent diagnostic system is supposed to be carried out using architectures of deep convolutional networks with a large number of layers. For example, it is advisable to build RetinaNet architecture with a backbone built on a neural network such as ResNet or EfficientNet [23, 24].

**Acknowledgements** Russian Foundation for Basic Research funded the work according to research project No. 20-08-00797.

## References

1. Platzer, A.: Logical Foundations of Cyber-Physical Systems. Springer International Publishing (2018)
2. Lee, E.A., Seshia, S.A.: Introduction to Embedded Systems, A Cyber-Physical Systems Approach, 2nd edn. The MIT Press, USA (2017)
3. Darragi, N., El-Koursi, E.M., Collart-Dutilleul, S.: Architecture description language for cyber-physical systems analysis: a railway control system case study. WIT Trans. Built Environ. **135**, 227–237 (2014)
4. Vágner, J., Zelenka, J., Hába, A., Kohout, M., Havlíček, P.: Stationary device for vibrodagnostics of passing vehicle. Vibroeng. Proc. **6**, 98–103 (2015)
5. Valsan, V.N., Patil, C.Y., Patekari, J.M.: Non-contact rail track parameter measurement. In: 2015 International Conference on Industrial Instrumentation and Control (ICIC). IEEE Xplore (2015)
6. Yue, G., Xu, Z., Wang, L.: Vibration analysis for slab track at different train speeds using Bayes wavelet denoising. Proc. Instit. Mech. Eng., Part F: J. Rail Rapid Transit. **231**(8), 892–901 (2016)
7. Wei, J., Liu, C., Ren, T., Liu, H., Zhou, W.: Online condition monitoring of a rail fastening system on high-speed railways based on wavelet packet analysis. Sensors **17**(2), 318 (2017)
8. Feng, H., Jiang, Z., Xie, F., Yang, P., Shi, J., Chen, L.: Automatic fastener classification and defect detection in vision-based railway inspection systems. IEEE Trans. Instrumen. Measure. **63**(4), 877–888 (2014)
9. Yi, H., Kai, L.: Inspection and Monitoring Technologies of Transmission Lines with Remote Sensing. Academic Press (2017)
10. Karakose, M., Yaman, O., Akin, E., Baygin, M., Murat, K.: A new computer vision based method for rail track detection and fault diagnosis in railways. Int. J. Mech. Eng. Robotics Res. **6**(1), 22–27 (2017)
11. Orlov, S., Girin, R.: Intelligent technologies in the diagnostics using object's visual images. In: Studies in Systems, Decision and Control. Cyber-Physical Systems: Advances in Design & Modelling, vol. 259, pp. 301–312. Springer International Publishing (2020)
12. Andronchev, I.K., Tarasov, E.M., Bulatov, A.A., et al.: A technique for diagnosis of the resistance of conductive rail track joints. Russ. Electr. Eng. **91**, 149–152 (2020)
13. Faghih-Roohi, S., Hajizadeh, S., Nunez, A., Babuska, R., De Schutter, B.: Deep convolutional neural networks for detection of rail surface defects. In: Proc. of the 2016 International Joint Conference on Neural Networks (IJCNN), pp. 2584–2589 (2016)
14. Orlov, S., Girin, R. Piletskaya A.: Intelligent information processing system for monitoring rail tracks. In: Proc. of 3rd International Conference on Control in Technical Systems (CTS). IEEE Xplore (2020)

15. Railway joints//Lokomo. Railway Track, https://lokomo.ru/zheleznodorozhnyy-put/relsovye-styki.html, last accessed 2021/12/27
16. Keras: The Python Deep Learning library, https://keras.io/applications, last accessed 2022/01/10
17. Simonyan, K., Zisserman, A.: Very deep convolution networks for large-scale image recognition. arXiv:1409.1556 (2014)
18. Kingma, D.P., Ba, J.L.: Adam: a method for stochastic optimization. arXiv:1412.6980v9[cs.LG] (2017)
19. Gupta, V., Murzova, A.: Keras Tutorial: transfer learning using pre-trained models. Learn OpenCV, https://learnopencv.com/keras-tutorial-transfer-learning-using-pre-trained-models, last accessed 2021/12/20
20. Powers, D.M.W.: Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Technol. **2**(1), 37–63 (2011)
21. Ivaschenko, A., Orlov, S., Krivosheev, A.: Accented visualization user interfaces in augmented reality. In: Studies in Systems, Decision and Control. Cyber-Physical Systems: Digital Technologies and Applications, vol. 350, pp. 213–223. Springer Nature Switzerland AG (2021)
22. Ivaschenko, A.V., Sitnikov, P.V., Diyazitdinova, A.R.: Accented visualization application in interactive manuals for technical training and support. J. Phys: Conf. Ser. **1691**, 012122 (2020)
23. Zeng, N.: RetinaNet Explained and Demystified, https://blog.zenggyu.com/en/post/2018-12-05/retinanet-explained-and-demystified, last accessed 2021/12/25
24. Mingxing, Tan, Quoc, V. Le.: EfficientNet: rethinking model scaling for convolutional neural networks. arXiv:1905.11946v5[cs.LG] (2019)

# Information Support for Maintenance and Repair of a Cyber-Physical System of a Machine-Building Profile

**Yulia V. Nemtinova, Andrey B. Borisenko, Vladimir A. Nemtinov, Alexey S. Protasov, and Kirill V. Nemtinov**

**Abstract** The issues of maintenance and repair management of cyber-physical systems of machine-building profile are considered on the example of numerical control machines (CNC) based on the use of the theory of expert systems construction. CNC machines maintenance includes a set of organizational and technical measures that ensure the maintenance of their output parameters at a given level during the entire period of operation; inspection and control of the technical condition of the machine and control system: cleaning, coolant replacement, machine mechanisms and control system elements washing and lubrication, worn-out parts, components, failed elements and control units of the machine replacement; hydraulic and electric drives check. The authors set and solved the problem of optimal selection of measures during maintenance and repair of complex CNC machine equipment. An information and logical model of decision support during its implementation is proposed, which allows, depending on the current state of the machine components and the technical conditions for manufacturing parts, to find the optimal option for CNC milling machining centers technical maintenance and adjustment. The production rules included in the model are constructed according to the type: if … (the conditions are met), then … (the implementation of the consequence). The authors have developed the software that implements the solution of the task, which has been tested during the maintenance and repair of a vertical milling machining center with CNCChallenger MCV-2418.

**Keywords** Cyber-physical system · Decision-making · Information and logic model · Maintenance and repair · Vertical milling machining center with CNC Challenger MCV-2418

Y. V. Nemtinova (✉)
Tambov State University named after G. R. Derzhavin, 392000 Tambov, Russia
e-mail: nemtinov.va@yandex.ru

A. B. Borisenko · V. A. Nemtinov · A. S. Protasov · K. V. Nemtinov
Tambov State Technical University, 392000 Tambov, Russia

 
A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_17

## 1 Introduction

With the development of science and technology, the technology of manufacturing products on the equipment of cyber-physical systems of machine-building profile, including numerical control machines (CNC), has become one of the irreplaceable important technologies in the manufacturing industry. CNC machines are pieces of modern production equipment, and their reliability directly affects the quality of products. The maintenance of this type of equipment plays a strategic role in machine-building enterprise functioning and can account for a significant part of a product's life cyclecost. The implementation of an effective maintenance management system allows to understand its significance on a global scale through the integration of knowledge and information obtained by various participants (for example, operators, maintenance team, CNC machines manufacturers) and sources (for example, automatic alarms, maintenance group reports, management).

The authors of many works have considered various approaches to the creation of automated systems for complex equipment maintenance cyber-physical systems. Thus, authors of [1–3] proposed an approach to support decision-making on automated planning of research center CNC equipment maintenance. The basic structure of the maintenance management system is based on: classification of maintenance work; distribution of these works between different service departments in accordance with the required knowledge and safety; planning preventive interventions; defining checklist options and customized responses to consolidate maintenance work; maintenance team training. The main results are: tracking maintenance activities to ensure that maintenance tasks are performed in accordance with CNC machines manufacturers requirements, as well as involving machine operators in their equipment maintenance to improve productivity and asset reliability. In fact, this approach provides a methodology for those companies that are considering for the first time the possibility of improving maintenance planning, allowing the end user to receive basic equipment maintenance as well as to begin accumulating maintenance data.

As the authors of [4–7] note, an unreasonable maintenance strategy increases the costs of its implementation and can significantly reduce the efficiency of CNC machines. Therefore, in order to obtain a scientific and reasonable maintenance strategy at the system level, it is necessary to take into account not only the state of equipment units wear, but also their mutual influence with other CNC machines units. Equipment wear affects the manufactured products quality, since the number of defects rises with increasing wear [5]. The authors propose a strategy of CNC machines group maintenance, taking into account economic dependence, structural and stochastic dependencies between critical units, and optimizes the group maintenance strategy. The model of CNC machines group maintenance consists of four submodels: submodels of units wear, sub-models of group maintenance decision, submodels of the maintenance process, submodels of maintenance costs. Using the model of CNC machines group maintenance, the time and purpose of its implementation can be determined depending on the degree of units' wear. Then you can calculate each technical service cost. In the group service model, economic dependence and structural dependence between components are determined by costs, while stochastic dependence is determined by the failure rate. In this case, the Monte Carlo method is used to simulate the machine operation process, and CNC machines long-term maintenance costs are calculated according to a certain failure rate threshold. Finally, the genetic algorithm is used to optimize the threshold values of the equipment failure rate during preventive maintenance and group maintenance.

Due to the lack of monitoring data and unclear understanding of CNC machines units wear process in the work [8–11], a method was proposed for assessing the condition and making decisions on CNC machines maintenance with several states, based on partially observable Markov decision process chains. On the one hand, taking into account the imperfect operational information on CNC machines and the impact of various maintenance measures on the process of machine wear, a dynamic model for assessing their working condition was built. On the other hand, a decision-making model for CNC machines maintenance has been developed in order to minimize the total cost of losses (including maintenance costs and potential failure costs).

Decision-making on maintenance in large production systems is a complex process, since it requires the integration of various information [12–14]. The limited control policy is popular in practice. Maintenance is performed when the CNC machine wear condition reaches a threshold value. The authors, developing a structure based on discrete-time Markov chain models, evaluated the performance of the system within the framework of limited control policy in production systems consisting of CNC machines with several states.

In advanced manufacturing systems, CNC machines are an important equipment for manufacturing components of high precision products, while from the point of view of equipment maintenance, they are considered as "products" provided by CNC machine manufacturers. In [15], a new knowledge-based maintenance planning system is presented to facilitate information and knowledge exchange between all stakeholders, including machine tool manufacturers, users (production systems), maintenance service providers and parts suppliers (for machine tools), within the emerging "Product-Service" business model. To improve maintenance planning efficiency, case-based justification principles have been implemented.

From the point of view of the product life cycle, expensive and complex CNC machines are both equipment and products, the reliability of which affects not only the quality of manufactured parts, but also the profit of equipment suppliers. Efficient maintenance and service play an important role in improving equipment reliability and improving production efficiency. Reuse of historical knowledge is an effective way to solve new problems and improve maintenance efficiency. The authors of the work [16] adopted a two-level ontology based on reasoning and real-life examples to represent expert knowledge using adaptation-oriented search. Based on semantic similarity and correlation, the most adaptive historical case was chosen to solve the current new problem.

Literature analysis has shown that several approaches to automating decision-making processes are currently being used to carry out complex CNC equipment technical maintenance and repair, depending on the specific conditions of machine-building products production. In this regard, the authors of this chapter propose an approach to CNC machines maintenance management based on the expert systems theory [17–20], which is illustrated by the example of vertical milling machining centers. The type of equipment is used to carry out the following operations:

- surface treatment with a milling cutter in the radial direction;
- milling of grooves, ledges, threading, etc.;
- drilling through and blind holes;
- deployment;
- processing with boring cutters in a given size.

It should be noted that no matter how perfect the machine tool equipment is, operations for its adjustment and control over technological parameters will always be present. This is due to the wear and tear of the working units, and the difference in the manufacturing technologies of parts. Therefore, maintenance specialists need in-depth knowledge not only of the design, theory of production processes in lathes, but also the ability to adjust them depending on the properties and condition of the processed material, technical condition, wear of parts, assemblies and mechanisms [21, 22].

## 2 Materials and Methods

### *2.1 Maintenance Systems for Vertical Milling Machining Centers*

CNC machines maintenance includes a set of organizational and technical measures that ensure the maintenance of their output parameters at a given level during the entire period of operation; inspection and control of the machine and the control system technical condition: cleaning, replacement of coolant, washing and lubrication of machine mechanisms and the control system elements, replacement of worn-out parts, components of mechanisms and out-of-order elements and control units of the machine; inspection of hydraulic and electric drives. Each machine tool equipment is characterized by deterioration of technical parameters during operation, which is expressed in a natural change in geometry. That means that machine parts, being subjected to mechanical and erosive influences, change in size over time. As a result, the parallelism of the structure is violated, which, of course, affects the rigidity of the machine in whole and leads to a breakdown of the machine.

Causes of accidents and breakdowns of vertical milling CNC machining centers:

1. Design flaws:
   a. design flaws in assemblies and their fastening; insufficient rigidity and vibration resistance, shortcomings of the lubrication system; lack of safety devices, etc.
   b. design flaws of parts (incorrect assignment of tolerances, heat treatment modes, dimensions and geometric shape, etc.); Manufacturing defects:
2. Manufacturing defects:
   a. metal defects (defects in the quality of the material, casting and workpieces);
   b. machine assembly defects (poor-quality fitting of mating parts; poor-quality wiring of the pneumatic and hydraulic systems of the machine; leakage of lubricating oils; the presence of foreign objects and chips);
3. Poor-quality maintenance of the machine:
   a. maintenance of the equipment in a dirty condition (ingress of moisture, dust and chips between the rubbing parts of the machine);
   b. absence or poor-quality lubrication (lack of oil in the tanks of the machine; malfunction of the pump or its drive; clogged oil pipeline; use of contaminated oil or oil of inappropriate viscosity);
4. Incorrect adjustment:
   a. poor-quality adjustment and adjustment (application of feed, cutting depth and speeds above the calculated;
   b. the use of dull cutting tools).

To implement all milling operations on CNC machines, their careful preparation and adjustment is necessary. High-quality performance of these works in the optimal time has a positive effect on the finished part.

## *2.2 The Task of Optimal Selection of Measures During Maintenance and Repair of CNC Machines*

In a formalized form, the formulation of optimal choice problem of CNC machines maintenance and repair measures, depending on their current state, is presented as follows: we need to find such a list of measures so that when the conditions of the operator:

$$Tx_i^{\min} \le Tx_i \le Tx_i^{\max}, \;\; i = 1, 2, \ldots N, \tag{1}$$

which represents a mathematical model of decision support during CNC machines maintenance and repair, including information about their parts and assemblies, current condition, troubleshooting methods, characteristics of maintenance and repair processes, are met, the following provisions are valid:

$$k_{opt} = \arg\min Q(k), \quad k \in K. \tag{2}$$

The optimality criterion Q is presented as the sum of the relative losses of criteria taken with certain values of their importance: the estimated costs of implementing the CNC machines maintenance and repair processes, the manufacturability and safety of the processes:

$$Q(k) = \sum_{j=1}^{3} (\rho_j \ \varpi_j(k)), \tag{3}$$

where: $\rho_1$, $\rho_2$, $\rho_3$—weighting factors set by experts.

$$\rho = \left\{ \rho_j : \rho_j \geq 0, \quad j = 1, 2, 3; \ \sum_{j=1}^{3} \rho_j = 1 \right\}. \tag{4}$$

Here: $Tx_i$, $Tx_i^{\min}$, $Tx_i^{\max}$—accordingly, are the current and limit values of the *i*-th characteristics of the CNC machine, *N*—number of characteristics; $\rho_j \varpi_j(k)$—weighted losses for each *j*-th criterion; $\rho_j \varpi_j(k) = \varpi_j(F_j(k)), \quad j = 1, ...3, \quad k \in K$—functional dependencies that transform the corresponding functions $F_j(k)$ to a dimensionless view; $F_1(k)$—estimated costs for the implementation of CNC machines maintenance and repair processes; $F_2(k)$, $F_3(k)$—accordingly, the values of the manufacturability and safety of the processes of their implementation. For the function $F_1$ *(k)* we will look for the minimum for the functions $F_2(k)$, $F_3(k)$—maximum.

Functional dependencies that transform functions $F_j(k)$ to the dimensionless form are:

$$\varpi_1(k) = \frac{F_1(k) - F_1^0}{F_1^{\max} - F_1^0}, \quad \varpi_l(k) = \frac{F_l^0(k) - F_l(k)}{F_l^0 - F_l^{\min}}, \quad l = 2, 3. \tag{5}$$

Here: $F_1^{\max}$—maximum value of the function $F_1(k)$, for which the minimum is being searched; $F_2^{\min}$, $F_3^{\min}$—minimum value of functions $F_2(k)$, $F_3(k)$, for which the maximum is being searched; $F_1^0$, $F_2^0$, $F_3^0$- optimal values of the corresponding $F_j(k)$, $j = 1, \ldots 3$.

When solving the problem (1–5), the values of the coefficients are set $\rho_j$, $j = 1, \ldots 3$, satisfying the relation (4) and reflecting the relative importance of the criteria $F_j(k)$.

For each *k*-th generated list of activities, the estimated cost $F_1(k)$ represents the sum of the costs of individual events, a $F_2(k)$ and $F_3(k)$ they are calculated, respectively, as value of the manufacturability multiplied by the value safety of the processes of individual events included in the *k*-th version of the generated list.

## 3 Results

To find the optimal variant of the list of CNC machines maintenance and repair measures that meet the regulatory requirements, an information and logical model for supporting decision-making has been developed, which includes many production rules [23–25].

As an example, we will give a number of rules with which a list of procedures for CNC machines maintenance and repair can be created. The production rules included in the model are constructed according to the type: *if … (the conditions are met), then … (the implementation of the consequence).*

Currently, the database contains more than 250 rules, the specific type of which is considered in the following examples in relation to vertical milling machining centers with CNC, in particular, the Challenger MCV-2418 machine (see Fig. 1). The rules are collected by experts and authors during contacts with operators-adjusters of vertical milling machining centers.

1. Rules for the selection of measures for defects elimination when working with a vertical milling machining center with a CNCChallenger MCV-2418:

    a. *if* "defect when boring holes" = "the hole turned oval", *then* "mandrel runout" = "mandrel replacement";
    b. *if* "defect when boring holes" = "the hole turned oval*", then* "incorrect cutting modes" = "correction of cutting presses";
    c. *if* "defect when boring holes" = "the hole turned oval", *then* "spindle runout" = "replacement of bearings and necks";
    d. *if* "defect when boring holes" = "the treated surface turned out to be uneven", *then* "the cutter was chipped" = "check the cutter, if necessary, re-cut, replace with a new one";
    e. *if* "defect when boring holes" = "the treated surface turned out to be uneven", *then* "large cutter feed" = "check the correctness of the feed selection, carry out finishing";
    f. *if* "defect when cutting ends and ledges" = "non-perpendicular arrangement of the end or ledge to the axis of the part", *then* "incorrect installation of the cutter" = "check the correct installation of the cutter";
    g. *if* "milling defect" = "the milled surface has a metal float", *then* "the milling cutter is blunted" = "sharpen the milling cutter";
    h. *if* "milling defect" = "the milled surface has a metal float", *then* "the milling cutter is blunted" = "put a new milling cutter";


**Fig. 1** General view of the vertical milling machining center with CNC Challenger MCV-2418

i. *if* “milling defect” = “the milled surface has incorrect dimensions”, *then* “incorrect binding of zeros in the coordinate system” = “check the binding, if necessary, reconnect”;
j. *if* “milling defect” = “the milled surface has incorrect dimensions”, *then* “incorrect binding of the tool” = “check the binding, if necessary, reconnect”;
k. *if* “milling defect” = “the milled surface has incorrect dimensions”, *then* “incorrect tool binding” = “check tool binding”;
l. *if* “milling defect” = “the milled surface has incorrect dimensions”, *then* “incorrect value of the tool corrector” = “check the tool corrector”;
m. *if* “milling defect” = “the milled surface has incorrect dimensions”, *then* “check the “base” for the presence of chips” = “blow the “base” from the chips”.

2. Rules for the selection of measures in case of malfunctions at the Challenger MCV vertical milling machining center-2418:

   a. n case of spindle malfunctions (see Fig. 2):
      (1) *if* “spindle malfunction” = “increased radial runout”, *then* “curvature of the working shaft” = “correction of the spindle shaft”;

   (2) *if* "spindle malfunction" = "increased radial runout", *then* "insufficient bearing tightening, gaps" = "bearing adjustment, tightening of the bearing tightening nut";
   (3) *if* "spindle malfunction" = "noise in the upper bearing", *then* "bearing diagnostics" = "wash the bearing from dirt with grease replacement or bearing replacement".

b. in case of malfunctions of the ball screw pair (ball screw) (see Fig. 3):
   (1) *if* "screw malfunction" = "slight deterioration of the positioning system", *then* "screw wear" = "machine axis compensation";
   (2) *if* "malfunction of the ball screw" = "serious deterioration of the positioning system", *then* "wear of the ball screw" = "replacement of the ball screw";
   (3) *if* "screw nut malfunction" = "slight deterioration of the positioning system", *then* "ball bearing wear" = "machine axis compensation";
   (4) *if* "screw nut malfunction" = "serious deterioration of the positioning system", *then* "diagnostics of ball bearings" = "cleaning of ball bearings with replacement of lubrication or replacement of ball bearings";

c. in case of malfunctions of the universal fixing support (see Fig. 4):
   (1) *if* "support malfunction" = "deterioration of the positioning system", *then* "bearing wear" = "bearing replacement";
   (2) *if* "support malfunction" = "increased radial runout", *then* "bearing wear" = "bearing replacement";
   (3) *if* "support malfunction" = "shaft axes misalignment", *then* "bearing wear" = "bearing replacement".

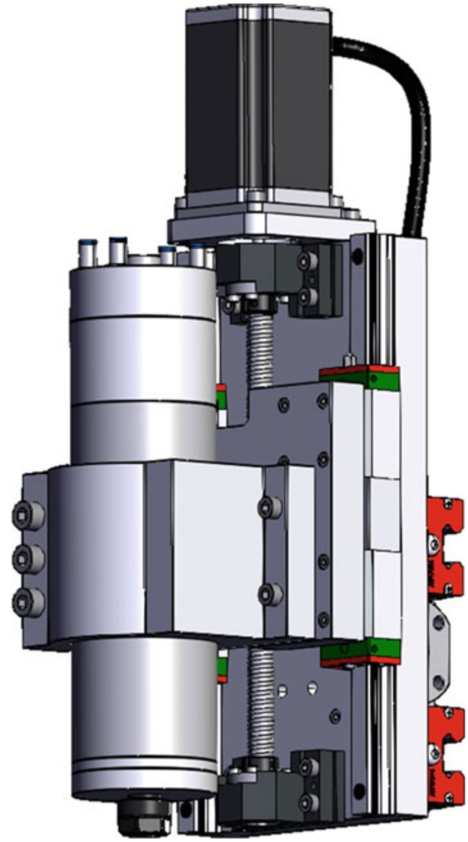
Information about the characteristics of each event is given in Table 1.

## 4 Discussion

As a result of the implementation of the decision support model for choosing the optimal options for CNC machines maintenance and repair, a variety of measures with different effectiveness can be formed. Currently, the dimension of the generated events set does not exceed a thousand, so the search for the optimal option is carried out by the complete search method.

The authors have developed a software package that implements the problems (1–5) solution and tested it during maintenance and repair of a vertical milling machining center with a Challenger MCV-2418 CNC.

**Fig. 2** Spindle model



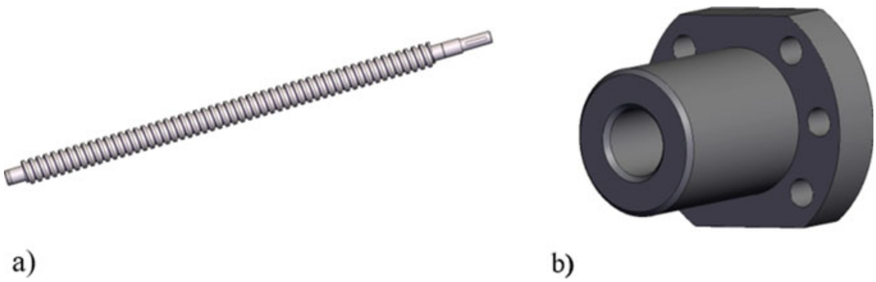



**Fig. 3** The model of the ball-screw pair: **a** the model of the screw; **b** the model of the nut

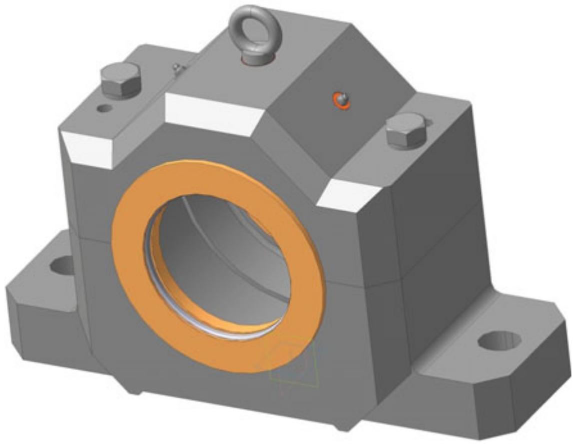


**Fig. 4** The model of the universal fixing support

**Table 1** Fragment of the measures characteristics database for CNC machines maintenance and repair

| Rule number | Manufacturability (scores from 1 to 10) | Estimated cost of the event (RUB) | Replacement safety (scores from 1 to 10, where 10 is complete safety, 1 is a serious injury to a worker) |
|---|---|---|---|
| … | … | … | . |
| 24 | 8 | 32,000 | 8 |
| 25 | 8 | 15,000 | 7 |
| 26 | 7.6 | 30,000 | 8 |
| 27 | 10 | 11,000 | 9.8 |
| 28 | 4 | 55,000 | 6 |
| 29 | 10 | 11,000 | 9.8 |
| 30 | 4 | 24,000 | 7 |
| 31 | 5 | 27,000 | 7 |
| 32 | 5 | 27,000 | 7 |
| … | … | … | … |

# 5 Conclusion

As a result of the conducted research, an approach to numerical control (CNC) machines maintenance and repair management has been developed, based on the use of the theory of building expert systems. The authors set and solved the problem of optimal selection of measures during complex CNC equipment maintenance and repair. The authors analyzed the issues of constructing an information and logical model of decision support during its implementation, which allows, depending on the current state of the machine components and the technical conditions for manufacturing parts, to find the best option for maintenance and commissioning of vertical milling machining centers with CNC. The approbation of the developed software package implementing the solution of the problem (1–5) during maintenance and repair of cyber-physical systems on the example of a vertical milling machining center with a Challenger MCV-2418 CNC showed the high efficiency of the proposed approach.

## References

1. Rossi, J., Bianchini, A., Cortesi, S., Naisson, P., Atieh, S.: Implementation of a maintenance management system for CNC machines. Proceedings of the Summer School Francesco Turco, 271549 (2021)
2. Sokolov, V., Krol, O., Baturin, Y.: Dynamics research and automatic control of technological equipment with electrohydraulic drive. In: Proceedings—2019 International Russian Automation Conference, RusAutoCon 8867652 (2019). https://doi.org/10.1109/RUSAUTOCON.2019.8867652
3. Pozdneev, B.M., Levchenko, A.N., Ovchinnikov, P.E., Sharovatov, V.I.: Integration of the automated control systems at a manufacturing enterprise. Russ. Eng. Res. **39**(12), 1057–1060 (2019). https://doi.org/10.3103/S1068798X19120165
4. Sun, J., Sun, Z., Chen, C., Yan, C., Jin, T., Zhong, Y.: Group maintenance strategy of CNC machine tools considering three kinds of maintenance dependence and its optimization. Int. J. Adv. Manuf. Technol. (2021). https://doi.org/10.1007/s00170-021-07752-6
5. Hajej, Z., Hajej, N., Rezg, A.: Gharbi quality issue in forecasting problem of production and maintenance policy for production unit. Int. J. Prod. Res. **56**(18), 6147–6163 (2018)
6. Nemtinov, V.A., Zazulya, A.N., Kapustin, V.P., Nemtinova, Yu.V.: Analysis of decision-making options in complex technical system design. J. Phys: Conf. Ser. **1278**(1), 012018 (2019). https://doi.org/10.1088/1742-6596/1278/1/012018
7. Krol, O., Sokolov, V.: Selection of worm gearing optimal structure for machine rotary table. Diagnostyka. **22**(1), 3–10 (2021). https://doi.org/10.29354/DIAG/129949
8. Zixuan, F., Xiaodong, W., Lifang, W.: Operational state evaluation and maintenance decision-making method for multi-state CNC machine tools based on partially observable Markov decision process. In: Proceedings of 2020 International Conference on Sensing, Diagnostics, Prognostics, and Control, SDPC 2020 9353134, pp. 120–124. (2020). https://doi.org/10.1109/SDPC49476.2020.9353134
9. Krol, O.S., Sokolov, V.I., Tsankov, P., Logunov, O.M.: Modelling of machining center vibration stability by the D-partitions method. J. Phys.: Conf. Series **1745**(1), 012085 (2021). https://doi.org/10.1088/1742-6596/1745/1/012085
10. Grieves, M., Vickers, J.: Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. In: Transdisciplinary Perspectives on Complex Systems, pp. 85–113. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-38756-7_4
11. Nemtinov V, Sergeev V, Nemtinova Y.: Design development of technological complex for friction welding of thin-walled products made of Armlen. IOP Conf. Series: Mat. Sci. Eng. **709**(3):033072 (2019). https://doi.org/10.1088/1757-899X/709/3/033072
12. Gu, X., Guo, W., Jin, X.: Performance evaluation for manufacturing systems under control-limit maintenance policy. J. Manuf. Syst. **55**, 221–232 (2020). https://doi.org/10.1016/j.jmsy.2020.03.003
13. Patil, R.B.: Integrated reliability and maintainability analysis of computerized numerical control turning center considering the effects of human and organizational factors. J. Qual. Maint. Eng. **26**(1), 87–103 (2020). https://doi.org/10.1108/JQME-08-2018-0063

14. Xing, J., Qiao, F., Lu, H.: A hybrid multi-objective algorithm for energy-efficient scheduling considering machine maintenance. IEEE Int. Conf. Auto. Sci. Eng. **8843144**, 115–120 (2019). https://doi.org/10.1109/COASE.2019.8843144
15. Wan, S., Li, D., Gao, J., Li, J.: A knowledge based machine tool maintenance planning system using case-based reasoning techniques. Robotics Comp.-Int. Manuf.
16. Wan, S., Li, D., He, F., Tong, Y.: Maintenance and service knowledge reusing for CNC machine tools. Comput. Integr. Manuf. Syst. **24**(9), 2254–2269 (2018). https://doi.org/10.13196/j.cims.2018.09.013
17. Redelinghuys, A.J.H., Kruger, K., Basson, A.H.: A six-layer architecture for digital twins with aggregation. In: Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future. Studies in Computational Intelligence, vol. 853, pp. 171–182. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-27477-1_13
18. Nemtinov, K., Eruslanov, A., Nemtinova, Yu.: Rationale construction of individual elements of technological complex. In: MATEC Web of Conferences: International Conference on Modern Trends in Manufacturing Technologies and Equipment (ICMTMTE 2018), vol. 224, 02036 (2018). https://doi.org/10.1051/e3sconf/201912600003
19. Krol, O., Sokolov, V.: 3D modelling of angular spindle's head for machining centre. J. Phys: Conf. Ser. **1278**, 012002 (2019). https://doi.org/10.1088/1742-6596/1278/1/012002
20. Karpushkin, S.V., Mokrozub, V.G., Glebov, A.O., Karpov, S.V.: Method for determining the parameters of heating systems for hydraulic presses. Chem. Pet. Eng. **56**(9–10), 744–754 (2021). https://doi.org/10.1007/s10556-021-00837-9
21. Pozdneev, B., Tolok, A., Ovchinnikov, P., Kupriyanenko, I., Levchenko, A., Sharovatov, V.: Digital transformation of learning processes and the development of competencies in the virtual machine-building enterprise environment. J. Phys: Conf. Ser. **1278**, 012008 (2019). https://doi.org/10.1088/1742-6596/1278/1/012008
22. Gromov, Y.Y., Ishchuk, I.N., Alekseev, V.V., Didrikh, V.E., Tyutyunnik, V.M.: Information support of finding a solution the problem of hidden objects. J. Theor. Appl. Inf. Technol. **95**(3), 615–620 (2017)
23. Alekseev, V.V., Yakovlev, A.V., Aminova, F.E., Sergeev, M.J.: The study of the complex information systems structure on the coefficient of structural uncertainty basis. J. Phys: Conf. Ser. **1399**(3), 033121 (2019). https://doi.org/10.1088/1742-6596/1399/3/033121
24. Shevchenko, S., Mukhovaty, A., Krol, O.: Research of the Modified Tooth-Belt Drive for the Machining Center. Lecture Notes in Mechanical Engineering. pp. 32–41, 033121 (2022). https://doi.org/10.1007/978-3-030-85233-7_4
25. Neydorf, R., Gaiduk, A., Gamayunov, N.: The multiplicative-isolating principle of significantly nonlinear mathematical models creation. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Cyber-Physical Systems: Modelling and Intelligent Control. Studies in Systems, Decision and Control, vol. 338, Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66077-2_3

# A System Dynamics Model for Float Glass Production Management

**Dmitry Petrov**

**Abstract** On the basis of the analysis of strategic goal map of multistage production of high-quality float glass the structure of system dynamics model accumulators is determined. The system of ordinary differential equations of system dynamics of multistage production of high-quality plate glass has been developed. The structure of the model of the system dynamics of sheet glass production has been developed. The equations connecting the accumulators and the values of constants are determined. A model of the system dynamics of the functioning of sheet glass production in the AnyLogic environment is presented. A computational experiment was carried out to simulate the functioning of the production of float glass. In the AnyLogic simulation environment, the Euler method was used to perform experiments with a model of system dynamics of multistage production of high-quality float glass to solve a system of differential equations. The simulation was performed over an interval of 10 years 3650 days. The graphs of the increase in the coefficients of glass defects depending on the seasonal components of the decrease in the quality of raw materials and wear of the main equipment, the dynamics of the filling capacity of the standard glass warehouse and the Jumbo glass warehouse in tons, taking into account the initial residues and seasonality of demand for these types of glass, the cost in rubles for equipment maintenance, the cost of gas, electricity and raw materials.

**Keywords** Simulation modeling · Strategic goal map · System dynamics · AnyLogic · Float glass production

D. Petrov (✉)
Institute of Precision Mechanics and Control, Russian Academy of Sciences, 24, Rabochaya str., Saratov 410028, Russia
e-mail: iac_sstu@mail.ru

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



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_18

## 1 Introduction

The modern level of automation of continuous technological processes allows us to speak about the presence of sufficiently effective cyberphysical systems of industrial enterprises with continuous nature of production [1, 2]. Quality management systems (QMS) in accordance with ISO 9000 series standards are used to manage production at such enterprises. Automated QMS based on ARIS, Business Studio, ELMA, etc. provide operational performance monitoring and forecasting by statistical methods fulfillment of the strategic goal map (SGM) of the enterprise, such systems belong to BPM-systems (Business Process Management) [3–5]. In order to realize the concept of Industry 4.0 [6–8], enterprises must develop digital twins in addition to BPM systems, which allow real-time simulation of production processes and make informed decisions on production management to meet individual customer requirements [9]. Traditionally, the optimal balance between the contradictions of "Fast-Cheap-Quality" in the form of the formulation of balanced indicators is determined by experts. In complex systems, the use of mathematical modeling can significantly improve the quality of decisions made by experts.

Multistage production of float glass by the float method involves a number of complex and interrelated chemical oxidation–reduction reactions and physical heat-mass transfer processes. It consists of the following stages: preparation of charge, glass melting, glass ribbon molding on molten tin, glass ribbon annealing, defect control, cutting glass ribbon into sheets of a given size and packaging the resulting sheets. The quality of float glass is significantly influenced by the technological process of charge mixture preparation, in which it is necessary to control the process of mixture formation of powders with different properties [10–13].

The initial and final stages of float glass production: preparation of charge, cutting glass ribbon into sheets and their packaging belong to the discrete technological processes (TP). The remaining stages are realized by continuous TPs. Thus, the multistage production of sheet glass is of a discrete–continuous nature. The analysis of such production was carried out in accordance with the stages of system analysis [14–16].

To analyze the functioning of such a complex nonlinear system, mathematical modeling is necessary [17–19]. The application of the "System Dynamics" simulation method gives an overall view of the functioning of the control object, focusing on global trends rather than on individual events. The system-dynamic model is defined by a system of differential equations that describe the interaction of storage units, flows between storage units, and dynamic variables.

The simulation modeling environment AnyLogic [20] based on the basic types of models: "system dynamics", "discrete-event" and "agent" allows us to build hierarchical models of complex systems. A system dynamics model of multistage production of high-quality float glass in AnyLogic environment provides the ability to optimize the cost of activities and determine the values of model parameters to minimize or maximize the value of the target functional. The goal of the work is mathematical modeling of the production process of multi-numbered float glass using the method of simulation “System dynamics” and computational experiment to study the effect on the cost of production of float glass parameters: the quality of raw materials, condition of the main technological equipment, the size of warehouses for glass of standard and “Jumbo” sizes.

## 2 Mathematical Model Development

The analysis of functioning of enterprises for the production of float glass allows us to determine the main goals of management of automated multi-stage production of high-quality float glass. Interconnections of the purposes are presented on the SGM (Fig. 1), developed in the Business Studio.

Based on the analysis of the objectives of the perspective “Internal Business Processes” of SGM multi-stage production of high quality float glass, the stock of the model of system dynamics are defined: $S_1$—“Enterprise finances”, $S_2$—“Charge production”, $S_3$—“Glass melting”, $S_4$—“Molded glass”, $S_5$—“Standard glass warehouse”, $S_6$—““Jumbo” glass warehouse”, $S_7$—“Normal state of equipment”, $S_8$—“Emergency state of equipment”.

Initial values, units and designations of stocks in the system dynamics model are given in Table 1.

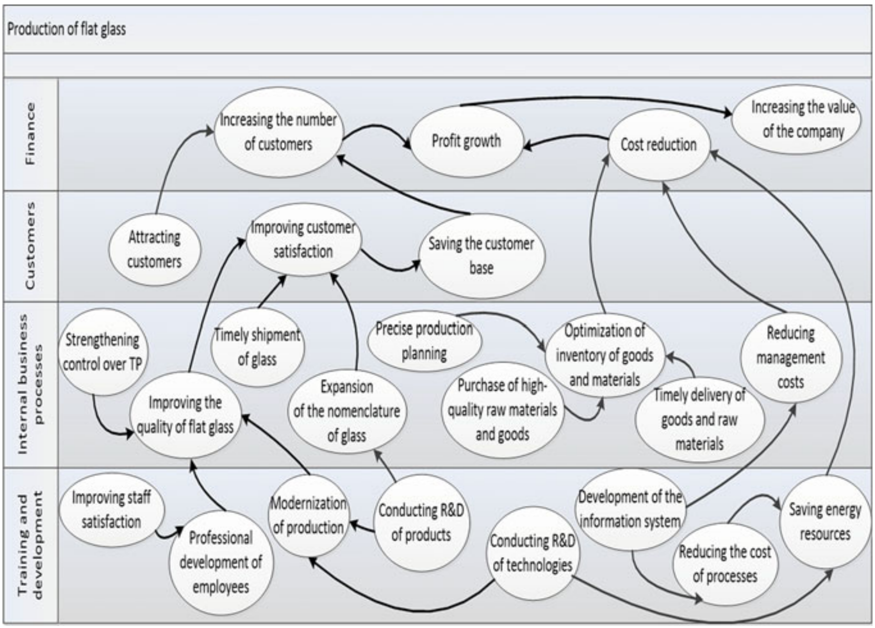



**Fig. 1** SGM of multistage production of high-quality float glass

**Table 1** Stocks of the system dynamics model

| No | Stock (S) | Designations in model | Initial values | Units |
|---|---|---|---|---|
| 1 | Enterprisefinances | Finance | 10,000,000 | rubles |
| 2 | Charge production | Charge | 700 | ton |
| 3 | Glass melting | Glass_furnace | 700 | ton |
| 4 | Molded glass | Glass_ | 700 | ton |
| 5 | Standard glass warehouse | Glass_warehouse_st | 40,000 | ton |
| 6 | "Jumbo" glass warehouse | Glass_warehouse_jumbo | 10,000 | ton |
| 7 | Normal state of equipment | Normal_state | 100 | % |
| 8 | Emergency state of equipment | Emergency_state | 0 | % |

Change of states of stocks in system dynamics is described by ordinary differential equations. The basic system of equations of system dynamics of multistage production of high-quality float glass has the form:

$$\begin{cases} dS_1/dt = r_1 - \mathrm{r}_2 \\ dS_2/dt = r_3 + \mathrm{r}_7 - \mathrm{r}_4 \\ dS_3/dt = r_4 - \mathrm{r}_5 - \mathrm{r}_6 \\ dS_4/dt = \mathrm{r}_6 - \mathrm{r}_7 - \mathrm{r}_8 - \mathrm{r}_9 \\ dS_5/dt = \mathrm{r}_8 - \mathrm{r}_{10} \\ dS_6/dt = r_9 - \mathrm{r}_{11} \\ dS_7/dt = r_{13} - \mathrm{r}_{12} \\ dS_8/dt = r_{12} - \mathrm{r}_{13} \end{cases}, \tag{1}$$

where $r_1$—flow "Income", $r_2$—flow "Consumption", $r_3$—flow "Supply volume of raw materials", $r_4$—flow "Charge loading", $r_5$—flow "Waste of raw materials", $r_6$—flow "Produced glass", $r_7$—flow "Quality control", $r_8$—flow "Standard glass production plan", $r_9$– flow ""Jumbo" glass production plan", $r_{10}$—flow "Standard glass shipment", $r_{11}$—flow ""Jumbo" glass shipment", $r_{12}$—flow "Equipment wear", $r_{13}$—flow "Equipment maintenance".

The stocks bind the flows. Their designations in the system dynamics model and units of measurement are given in Table 2.

Flow values: "Income", "Consumption", "Supply volume of raw materials", "Charge loading", "Waste of raw materials", "Produced glass", "Quality control", "Standard glass production plan", ""Jumbo" glass production plan", "Standard glass shipment", ""Jumbo" glass shipment", "Equipment wear", "Equipment maintenance" are determined from the following relations:

$$r_1 = v_1 + v_2, \tag{2}$$

**Table 2** System dynamics model flows

| No | Flow (r) | Designations in model | Units |
|---|---|---|---|
| 1 | Income | Income | rubles |
| 2 | Consumption | Consumption | rubles |
| 3 | Supply volume of raw materials | Supply_of_raw_materials | ton |
| 4 | Charge loading | Charge_loading | ton |
| 5 | Waste of raw materials | Waste | ton |
| 6 | Produced glass | Flow | ton |
| 7 | Quality control | Quality_control | % |
| 8 | Standard glass production plan | Plan_st | ton |
| 9 | "Jumbo" glass production plan | Plan_jumbo | ton |
| 10 | Standard glass shipment | Shipment_st | ton |
| 11 | "Jumbo" glass shipment | Shipment_jumbo | ton |
| 12 | Equipment wear | Equipment_wear | % |
| 13 | Equipment maintenance | Maintenance | % |

$$r_2 = v_3 + v_4 + v_5 + v_6, \tag{3}$$

$$r_3 = v_7 + r_5, \tag{4}$$

$$r_4 = S_2, \tag{5}$$

$$r_5 = S_3 * c_1, \tag{6}$$

$$r_6 = S_3 - r_5, \tag{7}$$

$$r_7 = S_4 * v_8, \tag{8}$$

$$r_8 = v_9, \tag{9}$$

$$r_9 = v_{10}, \tag{10}$$

$$r_{10} = v_{11}, \tag{11}$$

$$r_{11} = v_{12}, \tag{12}$$

$$r_{12} = v_{10} * c_{2,} \quad (13)$$

$$r_{13} = delay(r_{12}), \quad (14)$$

where $v_1$—variable "Cost per ton of standard glass", $v_2$—variable "Cost per ton of "Jumbo" glass", $v_3$—variable "Cost of electricity per day", $v_4$—variable "Cost of natural gas per day", $v_5$—variable "Cost of raw materials per day", $v_6$—variable "Cost of equipment maintenance per day", $v_7$—variable "Volume of raw materials", $v_8$—variable "Glass defect", $v_9$—variable "Standard glass production plan", $v_{10}$—variable ""Jumbo" glass production plan", $v_{11}$—variable "Standard glass ordering", $v_{12}$—variable ""Jumbo" glass ordering", $v_{13}$—variable "Equipment wear", $c_1$—constant "Raw material waste percentage", $c_2$—constant "Equipment wear factor", $delay()$– function of flow opening delay per period.

Table 3 shows the designations of dynamic variables in the system dynamics model and their measurement units.

The values of the dynamic variables: "Cost per ton of standard glass", "Cost per ton of "Jumbo" glass", "Cost of electricity per day", "Cost of natural gas per day", "Cost of raw materials per day", "Cost of equipment maintenance per day", "Volume of raw materials", "Glass defect", "Standard glass production plan", ""Jumbo" glass

**Table 3** A system dynamics model dynamic variable

| No | Dynamic variable (v) | Designations in model | Units |
|---|---|---|---|
| 1 | Cost per ton of standard glass | Cost_st | rubles |
| 2 | Cost per ton of "Jumbo" glass | Cost_jumbo | rubles |
| 3 | Cost of electricity per day | Cost_enrgy_d | rubles |
| 4 | Cost of natural gas per day | Cost_gas_d | rubles |
| 5 | Cost of raw materials per day | Cost_raw_dd | rubles |
| 6 | Cost of equipment maintenance per day | Maintenance_costs | rubles |
| 7 | Volume of raw materials | Rav_volume | ton |
| 8 | Glass defect | Defect_glass | % |
| 9 | Standard glass production plan | Plan_glass_st | ton |
| 10 | "Jumbo" glass production plan | Plan_glass_jumbo | ton |
| 11 | Standard glass ordering | Order_st | ton |
| 12 | "Jumbo" glass ordering | Order_jumbo | ton |
| 13 | Equipment wear | Maintenance | % |
| 14 | Electricity for glass production per day | Energy_volume | MW |
| 15 | Gas volume for glass production per day | Gas_volume | $м^3$ |
| 16 | Raw material quality factor | Quality_raw1 | % |
| 17 | Salary per day | Salary | rubles |
| 18 | Tax per day | Tax | rubles |

production plan", "Standard glass ordering", ""Jumbo" glass ordering", "Equipment wear" are determined from the following relations:

$$v_1 = v_{11} * c_3, \tag{15}$$

$$v_2 = v_{12} * c_4, \tag{16}$$

$$v_3 = v_{14} * 24/700 * c_6, \tag{17}$$

$$v_4 = v_{15} * c_7, \tag{18}$$

$$v_5 = v_7 * v_{16} * c_5, \tag{19}$$

$$v_6 = S_8 * r_{13} * c_{10}, \tag{20}$$

$$v_7 = r_8 + r_9, \tag{21}$$

$$v_8 = S_8/100 + (1 - v_{16}) + c_{11}, \tag{22}$$

$$v_9 = S_4 * v_{11}/(v_{11} + v_{12} + r_7), \tag{23}$$

$$v_{10} = S_4 * v_{12}/(v_{11} + v_{12} + r_7), \tag{24}$$

$$v_{11} = Demayd\_st(t), \tag{25}$$

$$v_{12} = Demayd\_jambo(t), \tag{26}$$

$$v_{13} = Maintence\_F(t), \tag{27}$$

$$v_{14} = S_4 * c_9, \tag{28}$$

$$v_{15} = c_8 * 24/700 * S_2, \tag{29}$$

$$v_{16} = Quality\_raw1\_F(t), \tag{30}$$

$$v_{17} = c_{13} * Salary\_F(t), \tag{31}$$

$$v_{18} = r_1 * c_{14}, \tag{32}$$

where $c_3$—constant "Price of standard glass per ton", $c_4$—constant "Price of "Jumbo" glass per ton", $c_5$—constant "Price of raw material per ton", $c_6$—constant "Price of electricity per megawatt", $c_7$—constant "Price of gas per 1000 cubic meters", $c_8$—constant "Gas consumption per hour per 700 tons", $c_9$—constant "Electricity consumption per hour per 700 tons", $c_{10}$—constant "Equipment repair rate", $c_{11}$—constant "Percentage of cullet glass in the charge", $c_{12}$—constant "Volume of glass furnace", $c_{13}$—constant "Average salary of employees", $c_{14}$—constant "Tax deductions coefficient", Demand_st—function of getting demand for standard glass, Demand_jumbo—function of getting demand for "Jumbo" glass, Quality_raw1_F—function of raw material quality change, Maintenance_F—function of equipment wear, Salary_F—function of staff labor compensation change.

On the basis of a system of differential Eqs. (1), equations of flows (2–14) and expressions for dynamic variables (15–32) the structure of mathematical model of the system dynamics of multistage production of high-quality float glass (Fig. 2) in simulation software AnyLogic was developed.

Table 4 shows the designations of constants in the system dynamics model, their values and units of measurement.

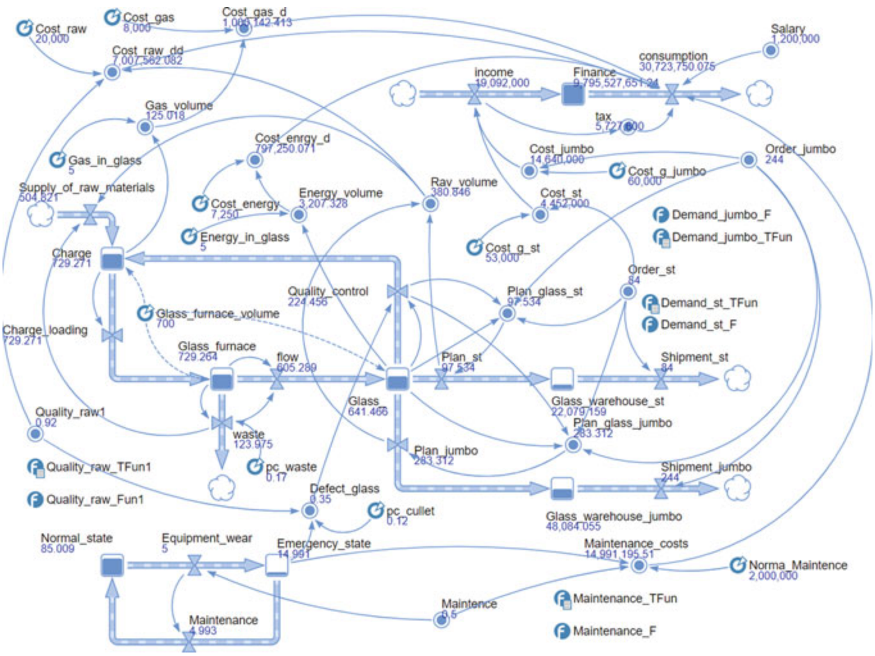



**Fig. 2** The structure of the system dynamics model

**Table 4** Dynamic variables of the system dynamics model

| No | Constant (c) | Designations in model | Values | Units |
|---|---|---|---|---|
| 1 | Percentage of raw material waste | Pc_waste | 18 | % |
| 2 | Coefficient of equipment wear | EE | 10 | coefficient |
| 3 | Price of standard glass per ton | Cost_g_st | 53,000 | rubles |
| 4 | Price of "Jumbo" glass per ton | Cost_g_jumbo | 60,000 | rubles |
| 5 | Price of raw material per ton | Cost_raw | 20,000 | rubles |
| 6 | Price of electricity per megawatt | Cost_energy | 7250 | rubles |
| 7 | Price of gas per1000 cubic meters | Cost_gas | 8000 | rubles |
| 8 | Gas consumption per hour per 700 tons | Gas_in_glass | 5 | 1000 $м^3$ |
| 9 | Electricity consumption per hour per 700 tons | Energy_in_glass | 5 | MW |
| 10 | Equipment repair rate | Norma_Maintence | 2,000,000 | rubles |
| 11 | Percentage of cullet glass in the charge | Pc_cullet | 12 | % |
| 12 | Volume of glass furnace | Glass_furnace_vol | 700 | ton |
| 13 | Average salary of employees | Salary_c | 1,200,000 | rubles |
| 14 | Tax deductions coefficient | Tax_c | 0.3 | Coefficient |

## 3 Mathematical Modeling

In the AnyLogic software to perform experiments with the system dynamics model of multistage production of high-quality float glass, the Euler method is used to solve the system of differential equations (1). Newton's method is used to solve algebraic Eqs. (2–32). The unit of model time chosen is a day. The simulation is performed on an interval of 10 years—3650 days. Numerical simulations were performed for these simulation parameters, initial values from Table 1, and constant values from Table 4.

Figure 3 shows graphs of increase of glass defect coefficients depending on seasonal components of decrease in quality of raw materials and wear of the main equipment. The chemical composition of raw materials is not stable and varies from batch to batch. Condition-based maintenance of equipment, flexible control of charge recipes and the charge program can reduce the cost of float glass production and reduce the reject rate by 14%.

Figure 4 shows the graphs of the dynamics of filling of the standard glass warehouse and the "Jumbo" glass warehouse in tons, taking into account the initial balances and the seasonality of demand for these types of glass. Application of automatic identification of float glass defects based on the use of artificial neural networks and wavelet transformations allows to reduce the number of unidentified defects and with the help of optimal cutting increase the production of standard size and "Jumbo" size glass.

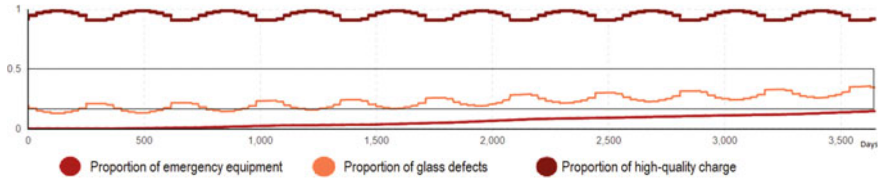



**Fig. 3** Results of simulation of float glass defects

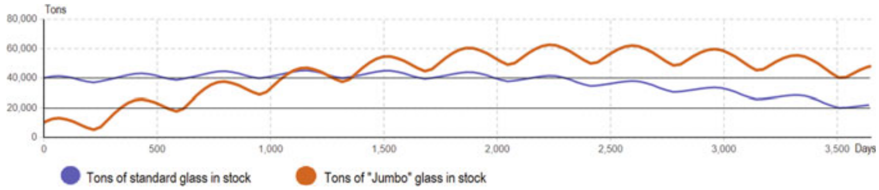



**Fig. 4** Dynamics of occupancy of the standard glass warehouse and the "Jumbo" glass warehouse

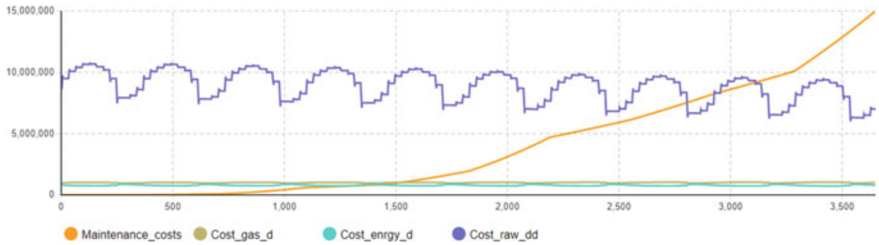



**Fig. 5** Dynamics of increase in equipment maintenance costs in rubles

Figure 5 shows graphs of costs in rubles for equipment maintenance, gas, electricity and raw materials. Control, equipment condition management on the basis of "condition-based maintenance" allows to reduce equipment maintenance costs and the percentage of float glass rejects.

## 4 Conclusion

In this chapter the issues of system analysis of float glass production are considered. Based on the objectives of the strategic goal map of the enterprise developed a system of differential equations for the functioning of multi-stage production of high-quality float glass. The equations linking the stocks and the values of the constants are defined. In simulation modeling software AnyLogic the structure of system dynamics model was developed. A computational experiment on modeling the functioning of float glass production has been carried out. The results of simulation modeling show that in order to reduce the cost of production of high-quality float glass it is necessary: flexible charge production control to correct changes of raw material quality, intelligent control of the condition of the main process equipment, automatic defect identification and optimization of float glass sheet cutting.

**Acknowledgements** The work was performed in Institute of Precision Mechanics and Control, Russian Academy of Sciences, under the state assignment of the Ministry of Science and Higher Education of the Russian Federation (theme No FFNM-2022-0010 "Development of intellectual models and methods of control of complex man-machine systems in critical situations").

## References

1. Meshalkin, V.P., Butusov, O.B., Reverberi, A., Kolmakov, A.G., Sevostyanov, M.A., Garabadzhiu, A.V., Alexandrova, A.G.: Multifractal analysis of the mechanical PropOKerties of the texture of biopolymer-inorganic composites of chitosan-silicon dioxide. Energies **15**(19), 7147 (2022). https://doi.org/10.3390/en15197147
2. Petrov, D.Yu. The architecture of the information system for managing the life cycle of a digital twin for continuous production. In: Proceedings of the St. Petersburg State Technological Institute (Technical University), No. 57, pp. 98–104 (2021)
3. Makarov, R.I., Khorosheva, E.R.: Information analysis of the sheet-glass utilization factor in production. Glass Ceram **77**, 295–297 (2020). https://doi.org/10.1007/s10717-020-00292-7
4. Makarov, R.I., Khorosheva, E.R.: Salient aspects of the implementation of digital economics in glass plants in Russia. Glass Ceram **75**, 438–440 (2019). https://doi.org/10.1007/s10717-019-00107-4
5. Viet, N.T., Kravets, A.G.: The new method for analyzing technology trends of smart energy asset performance management. Energies **15**(18), 6613 (2022). https://doi.org/10.3390/en15186613
6. Lim, K.Y.H., Zheng, P., Chen, C.J.: A state-of-the-art survey of digital twin: Techniques, engineering product lifecycle management and business innovation perspectives. Intell. Manuf. **31**, 1313–1337 (2020). https://doi.org/10.1007/s10845-019-01512-w. Access date 23 Jan 2022
7. Davis, J., Edgar, T., Porter, J., Bernaden, J., Sarli, M.: Smart manufacturing, manufacturing intelligence and demand-dynamic performance. Comput. Chem. Eng. **47**, 145–156 (2012). https://doi.org/10.1016/j.compchemeng.2012.06.037. Access date 22 Jan 2022
8. Redelinghuys, A.J.H., Basson, A.H., Kruger, K.: A six-layer architecture for the digital twin: a manufacturing case study implementation. J. Intell. Manuf. **31**, 1383–1402 (2020). https://link.springer.com/article/10.1007%2Fs10845-019-01516-6. Access date 12 Feb 2022
9. Bizubac, D., Popa, M.S., Hörmann, B.O., Faur, A.S.: Research of industrial processes in automation. Acta Technica Napocensis Ser. Appl. Math. Mech. Eng. **61**(2) (2018) [Electronic resource]: https://atna-mam.utcluj.ro/index.php/Acta/article/view/990/918. Access date 12 Mar 202
10. Geng T., Sau L.L., Xiaochuan Ya., Moo S.H.: A dimensionless analysis of residence time distributions for continuous powder mixing. Powder Technol **315**, 332–338 (2017). https://doi.org/10.1016/j.powtec.2017.04.007
11. Ghaderi, A.: Continuous mixing of particulate materials. In: Proceedings of the 4th International Conference for Conveying and Handling of Particulate Solids, At Budapest, Hungary, vol. 2 (2003). https://doi.org/10.13140/2.1.3487.6801
12. Kapranova, A.B., Verloka, I.I., Bahaeva, D.D.: About preparation of the analytical platform for creation of a cyber-physical sys-tem of industrial mixture of loose components. Cyber-Phys. Syst. Adv. Des. Model. Stud. Syst. Dec. Control. **259**, 81–91 (2020). https://doi.org/10.1007/978-3-030-32579-4_7

13. Bauman, I., Curic, D., Boban, M.: Mixing of solids in different mixing devices. Sadhana **33**(6), 721–731 (2008). https://doi.org/10.1007/s12046-008-0030-5
14. Glushkov, V.M.: Introduction to ACS, 2nd edn. K.: Technique, 320 p. (1974)
15. Polyakova, M., Shiriaeva, E., Iansaitova, M.: Application of system analysis for technological processes investigation. Key Eng. Mater. **834**, 24-31 (2020)
16. Chistyakova, T., Novozhilova, I., Kozlov, V., Shevchik, A.: Resource and energy saving control of the steelmaking converter process, taking into account waste recycling. Energies **16**(3), 1302 (2023). https://doi.org/10.3390/en16031302
17. Haddod, F., Dingli, A.: Intelligent Digital Twin System in the Semiconductors Manufacturing Industry. Artificial Intelligence in Industry 4.0, pp. 99–113 (2021). https://doi.org/10.1007/978-3-030-61045-6_8
18. Grega, W., Pilat, A., Tutaj, A.: Modelling of the glass melting process for real-time implementation. Int. J. Model. Optim. **5**, 366–373 (2015). https://doi.org/10.7763/IJMO.2015.V5.490
19. Groot, H., Laevsky, K., Farina, A., Klar, A., Mattheij, R., Mikelić, A., Siedow, N.: Mathematical Modelling of Glass Forming Processes (2010). https://doi.org/10.1007/978-3-642-15967-1_1
20. AnyLogic website [Electronic resource]: System Dynamics. https://www.anylogic.ru/use-of-simulation/system-dynamics/. Access date 21 Apr 2022

# Society 5.0: Cyber-Solutions Security

# Interval Analysis of Security for Information and Telecommunication Resources of Critical Infrastructures

**Igor Kotenko and Igor Parashchuk**

**Abstract** The object of the study is a new methodological approach to solving the problem of interval analysis of the security for information and telecommunication resources of critically important infrastructures. This approach is one of the options for the practical application of the theory of interval averages (interval calculations). The analysis of the peculiarities of this approach, which determine its suitability and usefulness for assessing the security indicators of such complex technical systems over a time interval, is carried out. Theoretical aspects of constructing the algorithms for calculating interval average values of security levels for information and telecommunication resources of critical infrastructures, the features of calculating the upper and lower average values of security levels are considered. A sequence of calculations and analytical expressions for such calculations are proposed on the example of a specific security indicator. At the same time, interval analysis does not have great mathematical and computational complexity, but it allows obtaining interval estimates of resource security that are adequate for control and management tasks, saves computing resources and, ultimately, works out to increase the reliability of security control of modern critical infrastructures.

**Keywords** Interval analysis · Security · Method · Interval averages · Indicator · Information and telecommunications resources · Critical infrastructures

I. Kotenko (✉) · I. Parashchuk
St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 39, 14-ya Liniya, St Petersburg 199178, Russia
e-mail: ivkote@comsec.spb.ru

I. Parashchuk
e-mail: parashchuk@comsec.spb.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_19

## 1 Introduction

The creation and improvement of critical infrastructures of any country is difficult, and often impossible, without the simultaneous and, in some cases, advanced development of information and telecommunication systems and complexes, which are the key elements of such infrastructures and provide the material basis for timely, reliable and secure data exchange between their users.

The critical Infrastructures (CIs) are understood as an economic, industrial, energy, information and/or telecommunication infrastructures, the termination or disruption of which leads to an emergency or significant negative consequences for a long period of time for defense, security, international relations, economy, other spheres of the economy of the country, or for the life of the population living in the corresponding territory [1, 2].

At the same time, the CI information resources include the entire set of data organized to effectively obtain reliable information in the interests of ensuring the CI smooth functioning. This is a collection of individual documents and arrays of documents, as well as many documents and arrays of documents in the CI information subsystems—libraries, archives, funds, data banks and other information subsystems [3, 4].

CI telecommunication resources can include all subscriber numbers, IP addresses and domain addresses available in telecommunication networks, systems and complexes, the number and capacity of communication lines (wire, optical, radio and satellite lines), channels and paths for transmitting information, routers, switching stations and nodes, as well as the radio frequency resource [5, 6].

At the same time, the tasks of analyzing the security of information and telecommunication resources (I&TR) of modern CIs deserve special attention. They are a key issue in the theory and practice of managing data protection in CIs and ensuring their secure functioning [7, 8].

At the same time, the existing difficulties in creating secure CI I&TR, the complexity of solving the problem of multicriteria analysis of their security are due, among other things, to the lack of universal methods that take into account a number of properties and features of the protection subsystems of the elements of these resources, the incompleteness and heterogeneity of the initial information on the quality of information protection in CI I&TR as a whole. Therefore, the development of new methods for calculating and multi-criteria interval analysis of the I&TR security of modern CIs with incomplete initial information is an important, urgent task.

## 2 Related Works

A lot of modern scientific research is devoted to the development of methods for analyzing the I&TR security of complex managed information and organizational and technical systems [9–23]. In these studies, various approaches to security analysis are formulated and described in detail. At the same time, the practical application of these approaches for a comparative analysis of the security level of various CI resources and for the development of promising directions for improving the protection of information in CIs becomes problematic.

In particular, this is due to taking into account the transient processes occurring in the operating CIs, the multi-criteria nature of modern requirements for the security of the CI I&TR and the management of the process of ensuring information security, which determine the formulation of not only the vector [9–11], but also the dynamic task of analyzing the security for resources of this class infrastructures, the approaches to the solution of which have not been studied within the framework of existing methods.

In [12], it is proposed to carry out security analysis taking into account risk management algorithms, but the advantages of this method are compensated by the large computational costs required for its practical implementation. In addition, the particular techniques of security analysis proposed in [13–15], in the overwhelming majority, do not take into account the parameters of the bodies and processes for managing the protection of CI I&TR when developing systems of security indicators (SIs) of this class resources. In the case when such a problem was nevertheless solved, for example, in [16], the absolute SI values were analyzed, but the deviations of these SI from the requirements for them were not taken into account, which is essential for analyzing the security of resources of such complex infrastructures.

The analysis of works [17–19] shows that the collection of a large amount of heterogeneous data on the state of the SI for complex technical systems, provided for in existing methods, leads to large time costs for collecting statistics, which negatively affects the efficiency of the analysis. This, in turn, affects the response time of the I&TR protection subsystem, increases its “response time” to the threats implemented by the intruder and the introduced control actions, focused on the initiation of certain means of protecting the CI I&TR.

The analysis of relevant works deserves special attention in terms of the accuracy of security analysis. So, for example, in the existing methods [20, 21], the traditional criterion of accuracy—the variance of the estimation error, is equal to the a priori variance of the studied I&TR security process itself, which cannot be acceptable for modern high-precision control systems. Artificial neural networks are used to increase the reliability of the SI analysis [22, 23]. They make it possible to take into account one important facet of the uncertainty of the initial data—the incompleteness and inconsistency of the measuring (observed) information about the SI values, but, unfortunately, they are not able to take into account the fuzziness of such information and its “noisiness”.

On the other hand, studies in the interests of analyzing the security of complex systems under conditions of uncertainty, although already carried out, were not universal. They were mainly devoted to the current assessment of information security indicators within the framework of one type of uncertainty (for example, fuzziness) [24, 25].

Thus, the analysis of relevant works allows us to speak about the objective need to develop existing methods for analyzing the security of CI I&TR in the case of taking into account modern requirements of the CI security management subsystem, the requirements of officials (auditors, administrators) related to the duration of cycles of security monitoring and security management of this class systems.

In other words, there is an objective need to obtain not point, but interval estimates of SIs. Security analysis can and should be carried out with a certain frequency, should take into account all types of uncertainty in the initial data. The mathematical and methodological basis of such an analysis can be algorithms for calculating the interval averages of SIs for the CI I&TR. This will save computing resources, increase the reliability and adequacy of the analysis.

## 3 Theoretical Part—Aspects of Constructing Algorithms for Calculating Interval Average Values of Security Levels

The inertia of I&TR security control cycles and procedures for managing the security of information circulating and stored in the CIs emphasizes the need to obtain not point (instantaneous), but interval security assessments. An important element of the scientific and practical novelty of these studies, in our opinion, is the fact that they are distinguished from existing works by the presence of a complex, dual, both "horizontal" and "vertical" interval analysis.

In other words, the assessment of the SIs for the CI I&TR is carried out both at time intervals ("horizontally"), and interval averages (lower and upper limits) of the SI values ("vertically") are determined. At the same time, the time interval assessment ("horizontal") is ensured by the use of mathematical models based on continuous Markov chains, when the change of discrete states (values) of the SIs occurs at random times, but is estimated at a predetermined time interval. And the interval analysis of SIs ("vertical") is due to the use of methods of the theory of interval averages, which allow calculating the upper and lower average values (levels) of a specific SI for the CI I&TR in this time interval.

A rational methodological and analytical tool for solving such problems, in our opinion, can be interval calculation algorithms (interval estimation, calculation of interval averages), which are one of the key mathematical methods in the theory of interval averages. Algorithms for computing interval averages (ACIA) are able to mathematically correctly combine methods for describing various types of uncertainty, considering them as special cases. With the help of ACIA, it is possible to process heterogeneous incomplete, fuzzy and contradictory information [25–27].

ACIA make it possible to perform interval calculations and obtain the final interval assessment of the security indicators for the CI I&TR, and the physical meaning and operation of ACIA is as follows.

The initial data for the operation of ACIA is the predetermined time intervals Δt, within which the interval analysis of the security indicators for the CI I&TR will be carried out. These intervals Δt can range from one to 20 min, depending on the duration of the control cycle (control inertia) of the security for the CI I&TR.

The next step in the work of ACIA is determination of the SIs for the CI I&TR for each security state ψ(Θn) on the set of all identified SI values, by analyzing the statistics for the time Δt of the instantaneous (point, step) values of SIs. At the same time, the numerical values of parameters characterizing various aspects of information security can serve as the SIs for the CI I&TR: availability, confidentiality, integrity. For example, the time $I_{hcp}(\Delta t)$ of cracking and compromising passwords for access to I&TR can be the SI for confidentiality, the time $I_{spr}(\Delta t)$ of secure provision of I&TR to users and CI subsystems—for availability, the time $I_{ra}(\Delta t)$ to restore access of users and CI subsystems to I&TR after hacking and compromising passwords—for integrity [7].

The next step in the work of ACIA is the calculation of the upper and lower average values (levels) of a specific SI for the CI I&TR on the time interval Δt. The calculation is carried out on the basis of a statistical analysis of the security parameters measured and observed on the interval Δt using mathematical expressions underlying the methods of the theory of interval averages [26, 27].

For example, for a specific SI that characterizes the integrity of resources—the time to restore access to I&TR for users and CI subsystems after hacking and compromising passwords, the exact upper $\underline{I}_{ra}(\Delta t)$ and lower $\underline{I}_{ra}(\Delta t)$ values of its average level observed over the time interval Δt are calculated. Analysis, for example, of the recovery time of access to I&TR for users and CI subsystems after hacking and compromise of passwords, using ACIA, has a number of peculiarities. In particular, it is assumed that the SI initially has N possible states (values): {Θ1, Θ2, …, Θn, …, ΘN}. In this case, the SI for each state ψ(Θn) on the time interval Δt is determined on the set of all possible states of this SI for the CI I&TR.

## 4 Methodological Part (Determination of Upper and Lower Average Values of Security Levels) and Discussion

The search for the upper and lower average values of $SI_{LI}$ for the CI I&TR at a certain time interval is also subject to the rules for calculating interval averages: the next step in the work of ACIA is, within our example, the calculation of for a specific SI $I_{ra}(\Delta t)$ that characterizes the integrity of telecommunication resources (for the recovery time of users and CI subsystems access to I&TR after cracking and compromising passwords), on the time interval Δt using the identified states (values):

$$
\begin{aligned}
I_{ra}(\Delta t) &= \sum_{n=1}^{N} \psi(\Theta_n)\ p_n(\Delta t) \\
&= (\psi(\Theta 1) p1(\Delta t))\ +\ (\psi(\Theta 2) p2(\Delta t)) + \ldots + (\psi(\Theta n) pn(\Delta t)) \\
&\quad + \ \ldots + (\psi(\Theta N) pN(\Delta t)),
\end{aligned} \tag{1}
$$

where ψ(Θn) is an indicator that characterizes the n-th (n = 1, …, N) state on the time interval Δt; $p_n(\Delta t)$ is the probability that a specific SI $I_{ra}(\Delta t)$ for the CI I&TR is in the state (has an identified value) Θn on the time interval Δt.

Even if the probabilities $p_n(\Delta t)$ for this particular SI $I_{ra}(\Delta t)$ are not defined (unknown), there is always the possibility to determine a priori on the basis of expert analysis the upper $\underline{I}_{ra}(\Delta t)$ and lower $\underline{I}_{ra}(\Delta t)$ boundaries of the average level of security, in our case—in terms of the recovery time of access to I&TR for users and CI subsystems after hacking and compromise of passwords, observed on the time interval Δt.

In this case, the SI ψ for the CI I&TR will be considered as a certain "feature" (an identified security parameter), and the functions $\underline{I}_{ra}(\Delta t)$ and $\underline{I}_{ra}(\Delta t)$ as the upper and lower averages of this feature. From the point of view of ACIA, this can be written mathematically as

$$
\{\underline{I}_{ra}(\Delta t)\ = \underline{R}(\psi); \underline{I}_{ra}(\Delta t) = \underline{R}(\psi), \tag{2}
$$

where $\underline{R}(\psi)$ and $\underline{R}(\psi)$ are the upper and lower interval means.

At the same time, the algorithms for calculating interval averages provide that in the case of the complete absence of any a priori information about the SI values, the upper $\underline{I}_{ra}(\Delta t)$ and lower $\underline{I}_{ra}(\Delta t)$ limits of the average level of protection are equal to one and zero, respectively [26]:

$$
\{\underline{I}_{ra}(\Delta t)\ = 1; \underline{I}_{ra}(\Delta t) = 0. \tag{3}
$$

Taking into account the fact that, within the framework of ACIA, it is customary to call an arbitrary numerical function f (y) as a "sign" of a random event (a random variable) (moreover, y ∈ Y, and Y is the space of elementary outcomes of SI values), in our case, the "sign" is the level (value) of a specific SI for the CI I&TR in the time interval Δt.

Then the concept of "interval averages" adequately characterizes the upper and lower average levels of security for the CI I&TR over the time interval Δt, and for a specific SI, for example, the time $I_{ra}(\Delta t)$ it takes for users and CI subsystems to access the I&TR after hacking and compromising passwords, it is possible to determine its "exact" average $R_{ra}(f)$ [26]:

$$
R_{ra}(f) = \int_{Y} f(y) dF(y), \tag{4}
$$

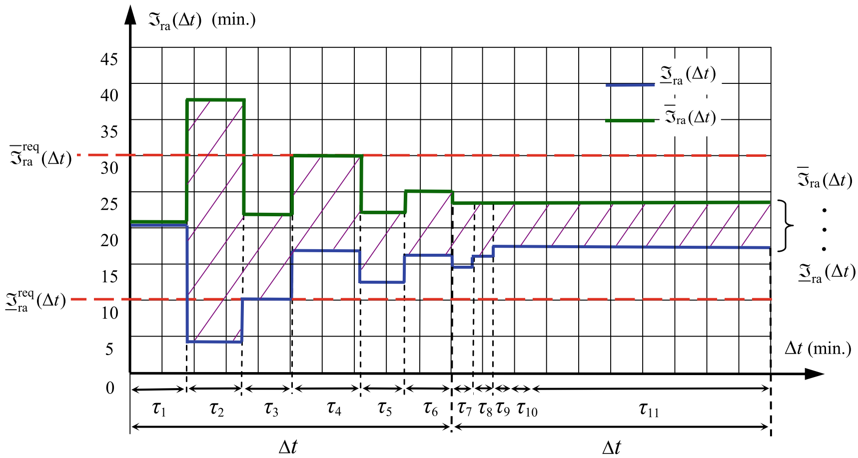



**Fig. 1** Graphical interpretation of the determination of the interval average of time $I_{ra}(\Delta t)$ of recovering the access to the CI I&TR for users and subsystems after hacking and compromising passwords

where F(y) is the distribution function of a random variable—the SI values, and the interval mean of our particular "sign" is the interval $[\underline{R}_{ra}(f) \ldots \underline{R}_{ra}(f)]$, which, without taking into account the concept of "sign", can be written as $[\underline{I}_{ra}(\Delta t) \ldots \underline{I}_{ra}(\Delta t)]$. An example of the application of the theory of interval averages in terms of determining the average within the interval $[\underline{I}_{ra}(\Delta t) \ldots \underline{I}_{ra}(\Delta t)]$ for the time to restore access after hacking and compromising passwords is shown in Fig. 1.

Figure 1 illustrates that with an a priori unknown distribution function F(y), it is possible to talk only about "boundaries" $\underline{F}(y) \leq F(y) \leq \underline{F}(y)$ for all values y ∈ Y. At the same time, within the "exact" limits

$$\underline{R}_{ra}(f) \leq R_{ra}(f) \leq \underline{R}_{ra}(f), \tag{5}$$

characterizing the upper $\underline{I}_{ra}(\Delta t)$ and lower $\underline{I}_{ra}(\Delta t)$ interval averages,

$$\underline{I}_{ra}(\Delta t) \leq I_{ra}(\Delta t) \leq \underline{I}_{ra}(\Delta t), \tag{6}$$

the "exact" average is found for a specific SI for the CI I&TR, in our case, for the time $I_{ra}(\Delta t)$ of restoring the access of users and CI subsystems to I&TR after hacking and compromising passwords.

Analysis of the graphs in Fig. 1 allows one to see that, taking into account the initial, starting time for restoring user access $I_{ra}(\Delta t) \approx 20$ min (given as an example), the state can take values from 0 to 45 min, and its estimate reaches a stable state already at the eighth step (τ8) of operation and at the second step (time interval $\Delta t$) of interval analysis.

Graphs in Fig. 1, in addition to observing the change of estimated interval values $I_{ra}(\Delta t)$ (i.e., its transition from one state to another in random time intervals τ), they allow one to get an answer to a question that is important, even fundamental for security analysis: are the obtained interval lower $\underline{I}_{ra}(\Delta t)$ and the upper $\underline{I}_{ra}(\Delta t)$ values of the analyzed SI for the CI I&TR (shaded sector) within the limits of acceptable, required average lower $\underline{I}_{ra}^{req}(\Delta t)$ and upper $\underline{I}_{ra}^{req}(\Delta t)$ values specified by the security policy or the CI security administrator.

To determine other security indicators, for example, within the framework of determining the level of confidentiality—for the time $I_{hcp}(\Delta t)$ of hacking and compromising passwords for access to I&TR; within the framework of determining the level of accessibility—for the time $I_{spr}(\Delta t)$ of secure provision of I&TR to users and CI subsystems on the time interval Δt, the task of finding their interval average values is solved in a similar way.

## 5 Conclusions

The considered method shows that there is a practical possibility of analyzing the security of the CI I&TR, taking into account the modern requirements of CI security management subsystems, the requirements of officials (auditors, security administrators) associated with the inertia of decision-making processes, with the duration of cycles of security monitoring and security management of systems of this class. This approach is based on algorithms for calculating interval averages and makes it possible to obtain not point (instantaneous), but interval estimates of SIs, while security analysis is carried out at a predetermined frequency and takes into account the uncertainty of the initial data.

The proposed methodological approach does not have great mathematical and computational complexity, but it allows obtaining interval estimates of the security of CI resources that are adequate to the tasks of control and management. This takes into account the uncertainty of the observed and controlled security parameters, and saves computing resources, which ultimately leads to an increase in the reliability of the security analysis of modern critical infrastructures.

The practical application of the proposed approach to interval computing is possible both in the framework of research work and in systems for automated security control for resources of complex managed critical infrastructures. The direction of further research may be the development of security analysis methods that combine interval calculations and traditional filtering algorithms.

**Acknowledgements** This research is being supported by the grant of RSF #21-71-20078 in St. Petersburg Federal Research Center of the Russian Academy of Sciences.

## References

1. Setola, R., Luiijf, E., Theoharidou, M.: Critical Infrastructures, Protection and Resilience. Managing the Complexity of Critical Infrastructures. In: Studies in Systems, Decision and Control, pp. 1–18. Springer-Verlag Berlin Heidelberg GmbH, Berlin (2016)
2. Cogwell, M.T.: Critical Infrastructures. Nova Publishers, New York (2003)
3. Jordan, J.M.: Information, Technology, and Innovation. Resources for Growth in a Connected World. John Wiley & Sons Limited, New Jersey (2012)
4. Basch, E.M., Thaler, H.T., Shi, W., Yakren, S., Schrag, D.: Use of information resources by patients with cancer and their companions. Cancer Interdisc. Int. J. Am. Cancer Soc. **100**(11), 2476–2483 (2004)
5. Ortiz, J.H.: Telecommunications Networks: Current Status and Future Trends. InTechOpen (2012)
6. Bouras, C.J.: Trends in Telecommunications Technologies. InTech, Patras (2010)
7. Kotenko, I.V., Parashchuk, I.B.: Evaluation of information security of industrial automation systems using fuzzy algorithms and predicates. In: Paper presented at the International Russian Automation Conference (RusAutoCon), Sochi, Russia (5–11 Sept. 2021). IEEE Xplore Digital Library: Browse Conferences, vol. 9537332, pp. 261–266 (2021)
8. Kotenko, I., Stepashkin, M., Doynikova, E.: Security analysis of information systems taking into account social engineering attacks. In: Paper presented at the 19th Euromicro International Conference on Parallel, Distributed and network-based Processing (PDP 2011), pp. 611–618 (2011)
9. Kamara, M.K.: Securing Critical Infrastructures. Xlibris US, Bloomington (2020)
10. Complying with the European NIS Directive. In: Cybersecurity for critical infrastructures. KPMG (2019)
11. Erbach, G.: Cybersecurity of Critical Energy Infrastructure, EPRS, European Parliament (2019)
12. Arnold, R.: Cybersecurity: A Business Solution: An Executive Perspective on Managing Cyber Risk. Threat Sketch, LLC, Winston-Salem (2017)
13. O'Neil, M.J., Dempsey, J.X.: Critical infrastructure protection: threats to privacy and other civil liberties and concerns with government mandates or industry. Depaul Bus. Law J. **12**, 97–111 (2000)
14. Kotenko, I., Saenko, I., Branitskiy, A.: Machine learning and big data processing for cybersecurity data analysis. In: Sikos, L.M., Raymond Choo, K.-K. (eds.) Data Science in Cybersecurity and Cyberthreat Intelligence. Intelligent Systems Reference Library book series (ISRL), vol. 177, pp. 61–85. Springer, Cham (2020)
15. Al-Mhiqani, M.N.: Cyber-security incidents: a review cases in cyber-physical systems. Int. J. Adv. Comput. Sci. Appl. **1**, 499–508 (2018)
16. Doynikova, E., Fedorchenko, A., Kotenko, I.: A semantic model for security evaluation of information systems. J. Cyber Secur. Mob. **9**(2), 301–330 (2019)
17. NIST Special Publication 800-61: Revision 2, Computer Security Incident Handling Guide, 16 Jan 2020. https://www.nist.gov/privacy-framework/nist-sp-800-61. Accessed 14 Jan 2022
18. ISO/IEC 27043:2015: Information Technology. Security Techniques. Incident Investigation Principles and Processes, 2015-03. https://www.iso.org/ru/standard/44407.html. Accessed 15 Jan 2022
19. Ekpo, U.: Introduction to Cyber Security. Fundamentals. Independently published, New York (2018)
20. Gabber, H.: The 2020 CyberSecurity & Cyber Law Guide. Independently published, New York (2020)
21. Arthur, C.: Cyber Wars. Hacks that Shocked the Business World. Kogan Page, London (2018)
22. Desnitsky, V.A., Kotenko, I.V., Parashchuk, I.B.: Neural network based classification of attacks on wireless sensor networks. In: Paper presented at the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). 27–30 Jan. 2020, St. Petersburg and Moscow, Russia, IEEE Xplore Digital Library (19 March 2020), pp. 284–287 (2020)

23. Meeuwisse, R.: Cybersecurity Exposed: The Cyber House Rules. Cyber Simplicity Ltd., London (2017)
24. Buldakova, T.I., Mikov, D.A.: Comprehensive approach to information security risk management. In: CEUR Workshop Proceedings, vol. 2081, pp. 21–26 (2017)
25. Buckley, J.J., Eslami, E.: An Introduction to Fuzzy Logic and Fuzzy Sets. Springer-Verlag, Berlin Heidelberg GmbH, Berlin (2002)
26. Guerra, M.L., Magni, C.A., Stefanini, L.: Interval and fuzzy average internal rate of return for investment appraisal. In: Fuzzy Sets and Systems, vol. 257, pp. 208–250 (2012)
27. Alefeld, G., Mayer, G.: Interval analysis: theory and applications. In: J. Comput. Appl. Math. **121**, 421–464 (2000)

# Assigning Access Rights to Applications in the Corporate Mobile Network: Software Development

**Alla G. Kravets, N. A. Salnikova, and E. L. Shestopalova**

**Abstract** The chapter presents the study and implementation of a method for assigning access rights to applications in a corporate mobile network. The existing solutions for assigning access rights to applications and services of the intranet are analyzed, the need to create a method to ensure the security of information when accessing corporate networks with different requirements for the level of security is substantiated, and the progress and results of development are described.

**Keywords** Access rights · Differentiation of powers · Application · Corporate mobile network · Information system · Web server

## 1 Introduction

When creating a single information space of a large company, one of the most difficult problems is managing access rights to applications and services of the intranet. Each Web server or application server has its own permissions controls [1]. In some cases, these are just configuration files edited using a text editor, in others—a developed graphical interface, in others—built-in operating system tools. It occurs when the configuration of the system of differentiation of powers is placed in the source code of the application, i.e. any change in the settings entails the restructuring of the software modules [2, 3]. When first entering or deleting a user account from the system, when

A. G. Kravets (✉)
Volgograd State Technical University, 28 Lenin Av, Volgograd 400005, Russia
e-mail: agk@gde.ru

Dubna State University, 19 Universitetskay St., Moscow Region, Dubna 141982, Russia

N. A. Salnikova
Volgograd Institute of Management—Branch of the Russian Presidential Academy of National Economy and Public Administration, 8 Gagarin St., Volgograd 400131, Russia

E. L. Shestopalova
Volgograd State Medical University, Square of the Fallen Fighters, 1, Volgograd 400131, Russia



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_20

publishing or deleting a resource, the administrator has to make changes to each of the systems, which leads to errors, incorrect or outdated settings.

In simple cases, when the set of applications is small, the task of centralized control can be solved using our own developments or mechanisms built into Web servers or application servers [4]. For example, user information stores can be consolidated using meta-directory software or native applications into one central master directory, after which Web and application servers (with additional tools) can place access control information in the same directory [5, 6].

However, when the number of applications increases and the system becomes more complex, the cost of implementing such a solution can be equal to the cost of migrating systems to a new, unified platform [7, 8].

Mobile technologies have been in business for a long time and employers everywhere use them for work purposes. Emerging problems in the areas of security, reliability of data storage have to be solved constantly. The most serious risk for any company, of course, is the increased likelihood of business information leaks, which, in turn, leads to various losses—financial, technical, reputational. As part of corporate mobility, strict security rules must be introduced, at least two of which must be absolutely mandatory: mobile device management (MDM—Mobile Device Management), which ensures their configuration in accordance with specified policies, as well as the separation of personal and business data into mobile devices using appropriate corporate applications designed to work with e-mail, files and integrated with the corporate directory service. The main risks of companies are connected precisely with the insufficiency of the applied protection measures and separation of access rights to applications.

The relevance of this work is obvious, since the risks are especially high in the case of providing mobile access to business information to company executives. The higher the status and position of a person, the higher the value of confidential information that he uses in his work. Mobile access directly affects the quality of company management and, ultimately, the efficiency and effectiveness of its work. Therefore, it is important to clearly define the list of possible threats and, based on this, build a policy for the use of personal mobile devices with separation of access levels.

The aim of the work is to develop a method for assigning access rights to applications in corporate networks with different requirements for the level of security, which will take into account the characteristics of the activities of various users and assign access to the applications necessary for work.

Based on the analysis of trends and development prospects of modern corporate mobile networks, a contradiction has been revealed between the requirements for information security when accessing protected services and information using universal mobile devices, and the technical capabilities of information security systems that allow ensuring the security of information when such access is made in corporate networks with different security requirements.

## 2 Analysis of the Assignment of Access Rights in Corporate Networks

Modern conditions for organizing the activities of most companies need information technology in order to remain competitive, and therefore create conditions for the widespread use of high-tech solutions [9, 10].

The Information Technology Department is engaged in the support and development of the IT infrastructure of the enterprise, whose employees are usually engaged in the following activities:

- support of external and internal servers of the company;
- system administration;
- user support.

Let's take a closer look at user support.

The User Support Specialist is responsible for installing, assigning access, and performing diagnostics on software and hardware; provides technical support and advice to end users; is responsible for organizing the repair of computer equipment; ensures the availability of consumables for computer and office equipment; advises users on technical issues [11, 12].

The most time-consuming process is associated with the timely assignment of access rights to corporate network applications, since this requires an understanding of the functions and tasks performed by the user in accordance with job descriptions. Consequently, there is a significant expansion of the duties of a user support specialist [13, 14].

With the traditional approach to organizing access, employees who often access several applications must not only receive an account for each of them, but also go through the authorization procedure many times [15, 16]. The result is discomfort at work. In addition, passwords are sometimes written down on a "piece of paper" lying next to the computer, which already creates a weak spot in the security system; the support service is very often contacted by users who have forgotten one of their passwords. The diagram of the model of work of the IT department is shown in Fig. 1.

At the moment, the activity of a user support specialist is reduced to receiving an application for permissions to access each application separately. At the next stage, the appeal is analyzed and the need for an account is determined, which gives the user a login and password to complete the authorization procedure in the application [17, 18].

When such a situation is detected, an account is created that gives the user access rights. This procedure is acceptable for small companies, the number of employees and applications in which is limited to a few dozen [19].

For companies with more employees than this number of users, the support specialist may have a number of time-consuming tasks:

- when an employee is dismissed or transferred from one department to another, there is a need to change access to certain applications;

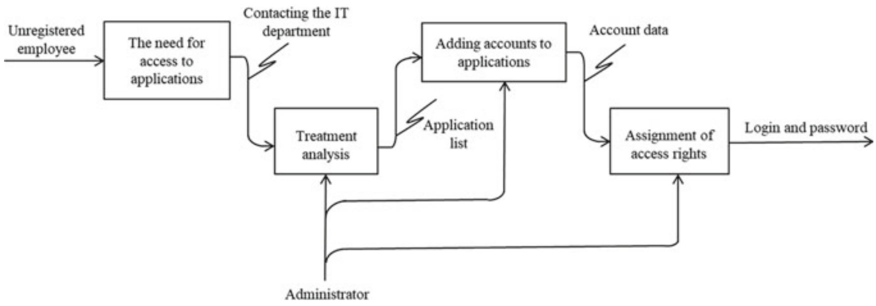



**Fig. 1** Structural and functional model of the process of assigning access rights to applications in a corporate mobile network

- when the staff of the company increases, the process of creating accounts and authorization takes up most of the working time, which prevents the performance of other duties within the position.

## 3 Review and Analysis of Existing Solutions

Currently, there are many different systems for restricting access to corporate applications on the market. Below is an overview of some systems that gives an idea of the state of the field.

### *3.1 Oracle Mobile Security Suite*

The new suite, combined with Identity and Access Management solutions, provides an integrated Oracle platform that allows organizations to manage access to all applications across all types of devices, including mobile gadgets, desktops and laptops. The solution meets the modern concepts of COPE (the device belongs to the company, is used personally by the employee, including for personal purposes) or BYOD (the device belongs to the employee and is used by him personally, including for work purposes) [20]. The structure of the complex is shown in Fig. 2.

The new suite, while being available as a stand-alone solution, is tightly integrated with the Oracle Mobile Suite for application development, integration, and deployment.

Among the main features of Oracle Mobile Security Suite is mobile security management. Among the management and control tools:

- Geo-Fencing mode to restrict access or functionality based on location;
- managing application usage policies, including restricting copy/paste/print functions to prevent data leakage;

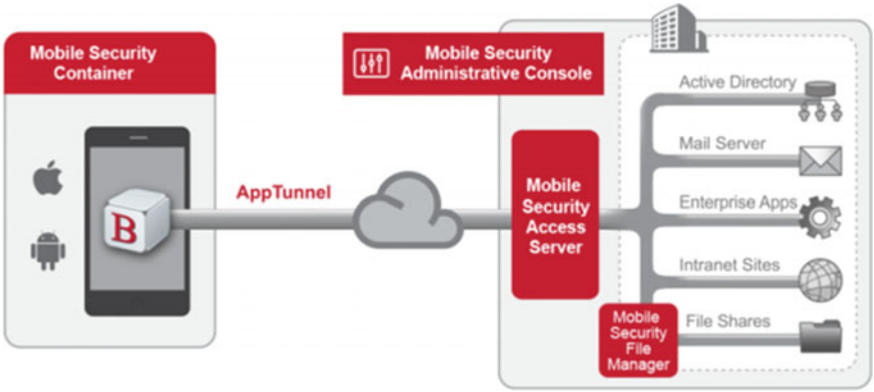



**Fig. 2** Structure of Oracle Mobile Security Suite

- secure network tunnels for each application, ensuring the protection of information without the use of VPN connections through virtual private networks at the level of devices that drain their batteries;
- encryption of local data with support for the possibility of remote deletion of corporate information and applications from the employee's personal device upon his dismissal.

### *3.2 MS Azure Active Directory*

Azure Active Directory (Azure AD) helps you manage cloud applications, on-premises applications, and resources through enterprise groups. Resources can either be directory-specific, such as permissions to manage objects through roles in the directory, or not, such as SaaS applications, Azure services, SharePoint sites, and local resources [21].

Azure AD allows you to grant access to corporate resources by granting access rights to a single Azure AD user or group. Using groups, the resource owner (or the owner of the Azure AD directory) can assign a set of access permissions to all members of the group instead of granting rights individually. The owner of a resource or directory can also grant rights to manage the member list to someone else, such as a department manager or help desk administrator, allowing that person to add and remove members as needed. The structure of Azure AD is shown in Fig. 3.

There are four ways to assign users access rights to a resource:

- direct appointment. The resource owner directly assigns the user to the resource.
- group assignment. The resource owner assigns an Azure AD group to the resource, which automatically grants all its members access to the resource. Group membership is managed by both the group owner and the resource owner, allowing each to add and remove group members.

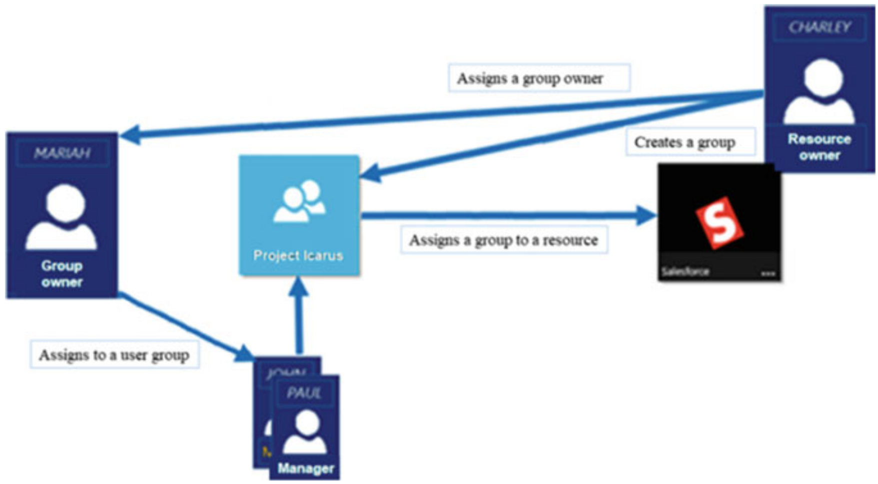



**Fig. 3** Structure of Azure AD

- rule-based assignment. The resource owner creates a group and uses a rule to determine which users are assigned to a particular resource. The rule is based on the attributes assigned to individual users. The resource owner manages the rule to determine which attributes and values are required to grant access to the resource.
- assignment by an external source. Access is provided by an external source such as a local directory or a SaaS application. In this situation, the resource owner assigns a group to provide access to the resource, and the members of the group are managed by an external source.

## 3.3 Cisco DNA Center

Comprehensive network management system to support business growth and innovation. Today, the network plays a more important role in business than ever before. The system automates infrastructure deployment, manages its lifecycle and connects it to the network, and maintains application quality and security, allowing IT staff to focus on networking projects that are strategically important to the business. With the Cisco® DNA Center™, you no longer have to spend too much time provisioning network resources and performing time-consuming troubleshooting tasks [22].

Features such as automatic network connection of devices and software image management allow you to install and update devices in minutes, not hours. And with a ready-to-use Cisco appliance, you can easily connect new remote offices to the network. DNA Center Assurance turns every point on the network into a sensor that generates a continuous stream of telemetry data about application performance and user connectivity in real time.

In addition, automatic path tracing data and a guided recovery feature are available. All this allows you to troubleshoot the network in minutes, before they lead to serious consequences. Through integration with Cisco Stealthwatch® security solutions, threats can be detected and remediated even in encrypted traffic.

DNA Center also provides an open, extensible platform that supports a wide variety of external applications and systems for sharing data and analytics with built-in functionality. DNA Center is the only centralized network management system that combines all of these features.

DNA Center offers a single control panel for all the basic functionality of your network. This platform empowers IT departments to respond faster and more efficiently to change and meet emerging challenges. DNA Center is the network management system, core controller, and analytics platform upon which the Cisco intent-based network is built.

In addition to device and configuration management, DNA Center offers a suite of software solutions that provide an entire network management platform; software-defined network controller for automating virtual devices and services; control mechanism to ensure a guaranteed quality of service for all users.

The DNA Center software runs on the DNA Center appliance and controls all of your Cisco devices, both physical and virtual (in and out of the fabric). The main menu of the DNA Center contains four general sections. The network is designed using visual physical maps and logical topologies. The functionality of the platform allows you to import existing maps, images, and topologies directly from the Cisco Prime Infrastructure and Application Policy Infrastructure Controller Enterprise Module (APIC-EM). This makes it easier and faster to perform updates.

Device discovery is performed automatically. To do this, you can either use the Cisco Discovery Protocol (CDP) or simply enter a range of IP addresses. DNA Center translates the information collected in the policy into specific network and device configurations according to the different types, builds, models, operating systems, roles, and resource limits of your network devices.

With DNA Center, you can create virtual networks, access control policies, traffic copy policies, and application policies. Once policies are created in DNA Center, provisioning is very easy using drag and drop. Each identity category (users, devices, applications, etc.) in the DNA Center registry is assigned a policy that will always be associated with that identity. This process is fully automated. New devices added to the network are assigned a policy based on their identities. This makes it much easier to set up a remote office. DNA Center Assurance is an end-to-end solution that delivers higher and more consistent service levels to meet growing business needs. DNA Center Assurance does more than just network monitoring and troubleshooting.

Proactive response and predictive network operations help ensure high performance for clients, applications, and services. As a result, service consistency is achieved, and the network optimization process becomes proactive. At the same time, less time is spent on troubleshooting and troubleshooting. The structure of the Cisco DNA Center is shown in Fig. 4.

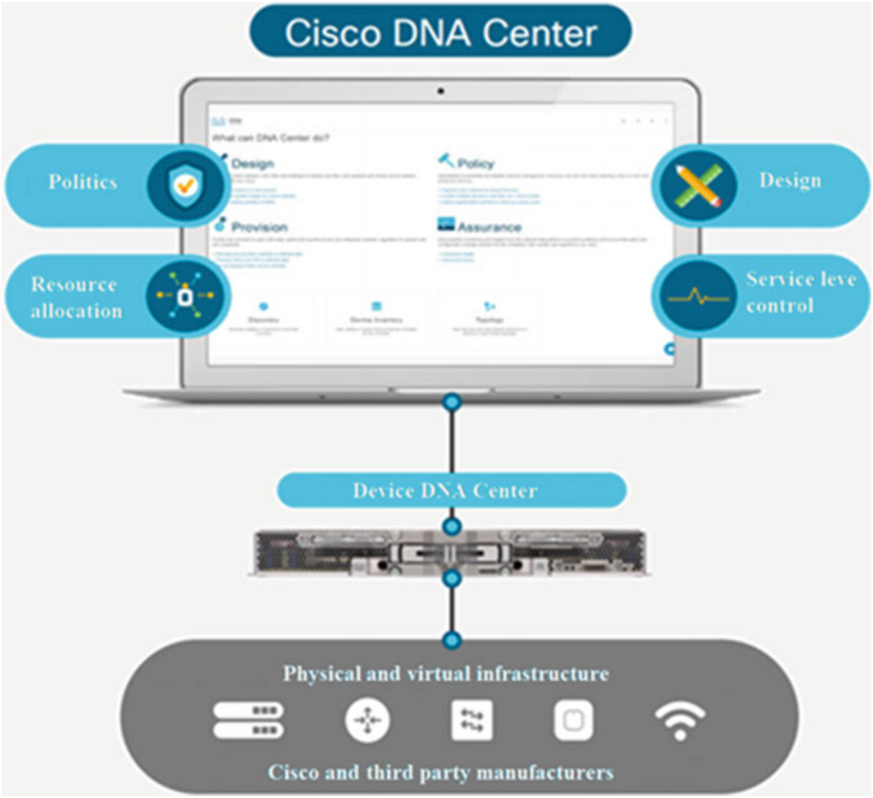



**Fig. 4** Structure of the Cisco DNA Center

## 3.4 G Suite Google

Google's mobile device management features are designed to administer, secure, and track corporate mobile devices. Can control various devices including phones, tablets and smart watches. Employees of an organization can work with its resources on personal or company-provided devices [23].

There are two levels of management: basic and advanced. The available options vary depending on the mode you are using. Some advanced features, such as device auditing and policy setting, are not available in all versions.

In basic control mode, you can do the following:

- set rules that require you to set up a screen lock or password to protect corporate data on devices.
- remove corporate data from lost or stolen devices;
- give users access to recommended enterprise apps for Android devices.
- publish and distribute private applications;

- view the list of devices with access to corporate data in the Google Admin Console.

If you need more features to control devices with access to corporate data, you can use the advanced management mode, which allows you to:

- establish a rule requiring the use of stronger passwords;
- provide users with access to enterprise applications in the Recommended Software Catalog (for Android and iOS);
- use work profiles on Android devices to separate personal and corporate apps;
- deny access to certain device settings and functions. For example, you can prohibit connecting to mobile networks or via Wi-Fi, creating screenshots, etc.;
- monitor policy compliance and receive reports on users, devices, and OS versions.

### *3.5 Comparative Analysis of Access Control Systems for Corporate Applications*

A comparative analysis of the systems for delimiting access rights to corporate applications is given in Table 1.

After analyzing ready-made solutions, it was decided to develop the module on its own, without involving paid software. This is due to the fact that already existing developments have either limited functionality, which is not enough to implement all the functions of the module, or redundant, i.e. the finances spent on the purchase of a ready-made solution will remain unused in full. Also, when purchasing software, it is necessary to allocate resources for training personnel in preparing, setting up and maintaining the stable operation of the purchased product. Also, in connection with the imposition of sanctions by foreign companies on the use of the results of intellectual activity and in order to increase the security of access to applications in

**Table 1** Comparative analysis of systems for delimiting access rights to corporate applications

| Function | Analyzed systems | | | |
|---|---|---|---|---|
| | Oracle Mobile Security Suite | MS Azure Active Directory (Azure AD) | Cisco DNA Center | G Suite Google |
| Free version | – | – | – | + |
| Convenient interface | + | + | + | + |
| Open source | – | – | – | – |
| Support for data encryption | + | + | + | + |
| Additional plugins | + | – | + | – |
| Multi-language support | + | + | +/– | + |

corporate networks, it is necessary to develop a unique method for assigning rights with different requirements for the level of security.

## 4 Design and Development of a Method for Assigning Access Rights to Applications in a Corporate Mobile Network

As part of the problem being solved, a multi-user system is used that runs on the entire computer fleet of the organization. Therefore, client–server DBMS are considered.

The client–server DBMS allows the client and server to exchange the minimum required amount of information. In this case, the main computational load falls on the server. The client can perform pre-processing functions before transmitting information to the server, but its main functions are to organize user access to the server [24].

In most cases, a client–server DBMS is much less demanding on computer network bandwidth than a file-server DBMS, especially when performing a search operation in the database according to user-specified parameters, because for search, there is no need to receive the entire array of data on the client: the client passes the request parameters to the server, and the server searches the local database for the received request. The result of the query, which is usually several orders of magnitude smaller than the entire array of data, is returned to the client, which provides a display of the result to the user. MS Access was chosen as the database. Embarcadero RAD studio was chosen as the programming language and development environment.

### *4.1 Algorithm for Implementing the Information System Method for Assigning Access Rights to Corporate Mobile Network Applications*

Building a model of an information system begins with a description of the functioning of the system as a whole. Consider the processes that occur during interaction in corporate networks. If the user wants to get access rights, he must contact the administrator directly, who, after analyzing the user's job descriptions and directions of work, will either add the user's data to the module or refuse it if work with corporate mobile network applications is not required to perform the work. If the user wants to log in to the network, he enters his login and password in the module's Web interface, after which he receives a list of available applications. The administrator can add, delete, edit and view the database.

The functional requirements for the developed method are that it must provide the ability to perform user authorization, provide the administrator with the ability to

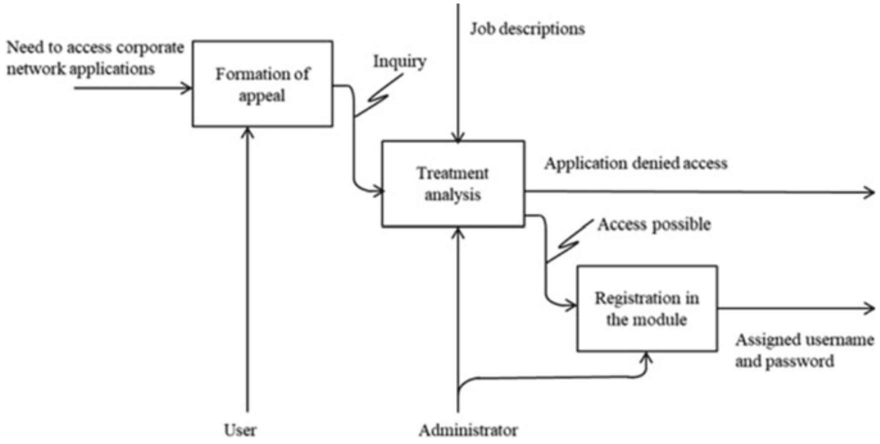



**Fig. 5** Decomposition diagram (IDEF0 TO BE)

manage user access rights to various applications, manage application storage, log user operations and keep records.

From the diagram shown in Fig. 5, it can be seen that the process of assigning access rights should be reduced to analyzing user requests and registering them using the module for assigning access rights.

On Fig. 6 shows the algorithm of the method for accepting and processing requests for assigning access rights to mobile corporate network applications.

The proposed method regulates the sequence of the process of applying for access to applications. The user directly contacts the administrator, who analyzes the employee's job descriptions and either grants access rights to the required corporate mobile network applications or refuses to receive them. To log in to the network, you must enter a username and password, after which the employee gets access to the necessary applications. The work is controlled by an administrator who can edit the database.

## *4.2 The Concept of Software Implementation of the Method of Assigning Access Rights to Applications of a Mobile Corporate Network*

When performing a software implementation of the proposed method for assigning access rights to applications, it must be taken into account that the system must provide the following functions:

- authorization: when using the module, users (administrator, employee) must perform authorization in the user interface;

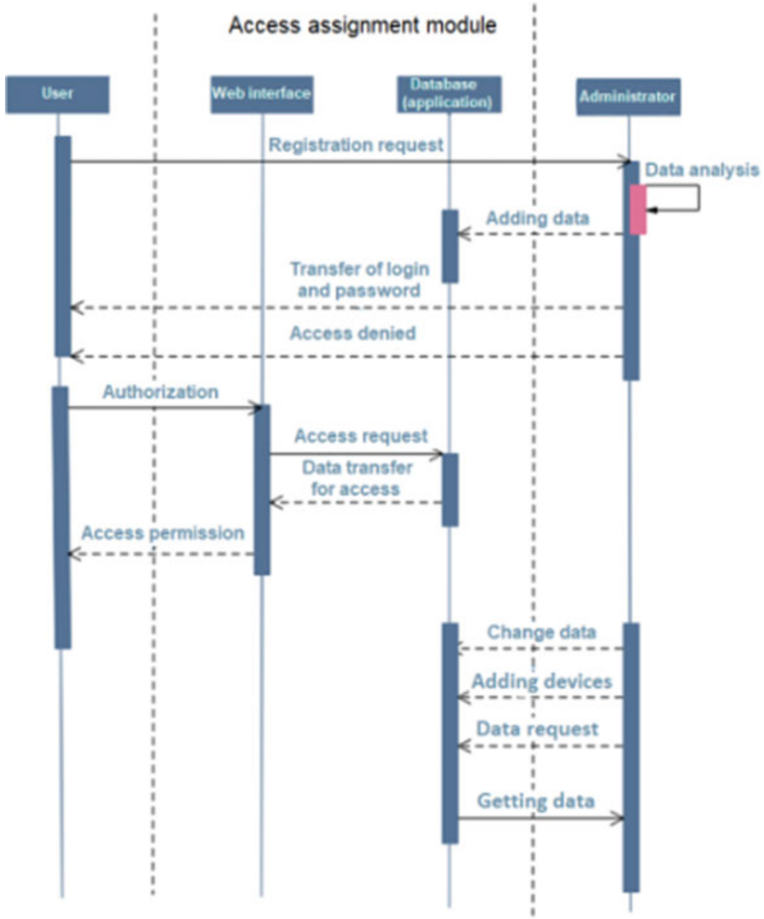



**Fig. 6** Algorithm of the method for assigning access rights to applications of a mobile corporate network

- user management: administrator maintains, adds and edits employee information (login, password, etc.);
- corporate application repository management: the administrator maintains, adds and edits the list of applications in the repository;
- mobile application management: assigning access rights to applications to company employees, data about devices stored in the database;
- record keeping: registration of transactions performed by users.

The input data should contain information entered into the database by the administrator:

- information about the user—all the necessary information about the user of mobile devices and applications (Login, Password, …);

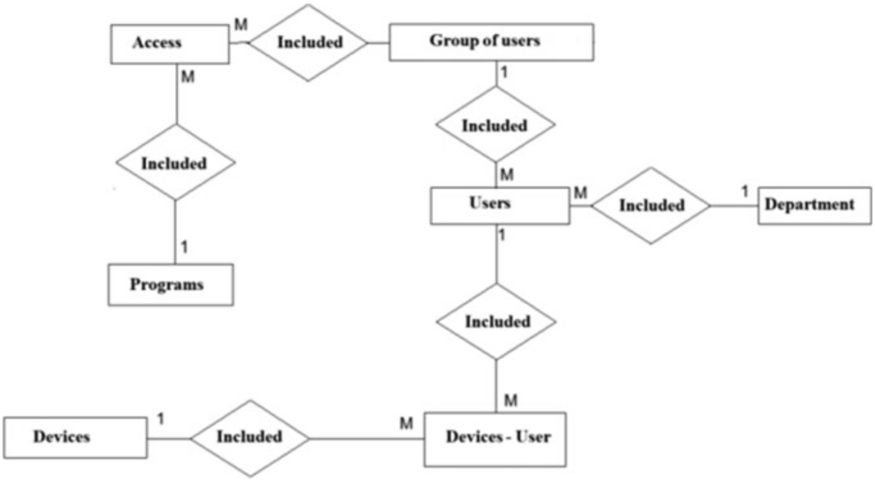



**Fig. 7** The data model

- mobile device—all necessary information about the mobile device (Name, ID, IMEI, Model, …);
- a list of applications and all the necessary information about the application (Name, Description, Version, …).

On Fig. 7 shows the data model [25].
During the analysis of the subject area, the following entities were identified:

- entity "Devices" (Code, IMEI, Device name, IP address, OS version, Registration date);
- entity "Department" (Code, Department name);
- entity "Users" (Code, Last Name, First Name, Login, Password, Group Code, E-mail, Comment, Department Code);
- entity "User groups" (Code, Group name, Comment);
- the essence of the "Program" (Code, Program name, Link to the program, Program version, Date added, Description);
- entity "Devices-users" (Code, Device code, User code);
- entity "Access" (Code, Group code, Program code).

The main program window (Fig. 8) consists of the following elements:

- tab for viewing information about users;
- tab for viewing information about user groups;
- tab for viewing information about programs.

The main menu (Fig. 8) provides the addition of new data about users, workgroups, programs and applications used, and also provides the ability to edit the entered data.

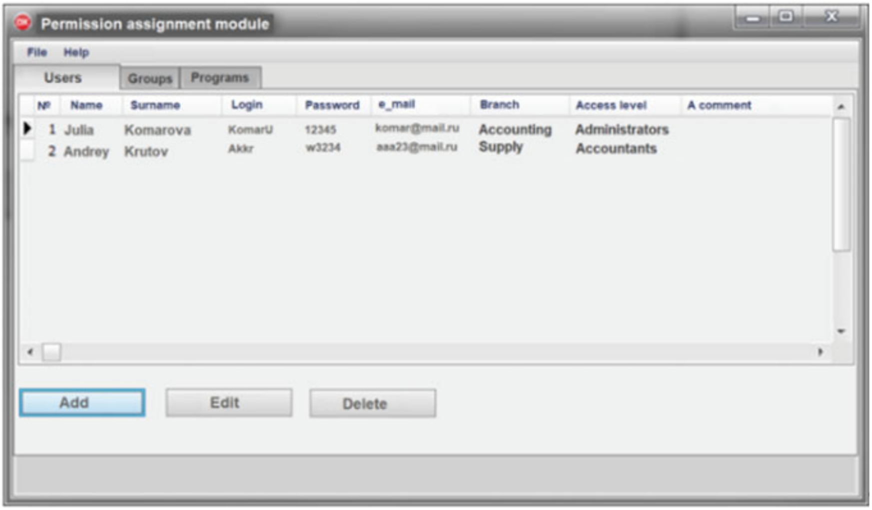



**Fig. 8** Screen form of the main software window

## 5 Conclusion

During the implementation of this project, an analysis was made of the activities of the department's employees on the issue of assigning access rights to applications; input, constant and output data are defined; the functions and purpose of individual components of the designed module are considered; analysis of existing developments, a review of existing systems and the choice of a module design strategy were carried out, as a result of which it was decided to develop this module on its own, without involving paid software; the choice of a DBMS and a user interface development environment has been made; the basics of design decisions were investigated and substantiated, the optimal DBMS—MS Access and the development environment—Embarcadero Delphi 10.2 were determined.

The design of the database is described step by step, a model of information flows is built, the physical scheme of interaction of individual parts of the module is considered, on the basis of which the database is created in the selected DBMS. A user interface has been developed with forms that display information stored in the database. Forms provide the ability to input, output data, edit them, process and store data about users of the corporate mobile network (employees of the company) and mobile applications, with the subsequent assignment of access rights to employees to use applications.

## References

1. Kravets, A., Kozunova, S.: The risk management model of design department's PDM information system. In: Communications in Computer and Information Science (CCIS-2017), vol. 754, pp. 490–500 (2017)
2. Kravets, A.G., Egunov, V.: The software cache optimization-based method for decreasing energy consumption of computational clusters. Energies **15**(20), 7509 (2022). https://doi.org/10.3390/en15207509
3. Kravets, A.G., Salnikova, N.A., Mikhnev, I.P., Orudjev, N.Y., Poplavskaya, O.V.: Web portal for project management in electronics design software development. In: 2019 International Seminar on Electron Devices Design and Production (SED 2019), Proceedings, pp. 8798472 (2019)
4. Access rights management in corporate web systems. Access mode: http://www.iksmedia.ru/articles/23910-Upravlenie-pravami-dostupa-v-korpor.html. Accessed 25 Jan 2022
5. Kravets, A., Salnikova, N., Dmitrenko, K., Lempert, M.: Industrial cyber-physical systems: risks assessment and attacks modeling. In: Studies in Systems, Decision and Control, vol. 260, pp. 197–210 (2020)
6. Bolshakov, A.A., Klyuchikov, A.V.: Decision Support System for Selecting Designs of Autostereoscopic Displays. In: Cyber-Physical Systems: Design and Application for Industry 4.0 (Studies in Systems, Decision and Control), vol. 342, pp. 73–88 (2021)
7. Kravets, A., Shumeiko, N., Lempert, B., Salnikova, N., Shcherbakova, N.: "Smart queue" approach for new technical solutions discovery in patent applications. In: Communications in Computer and Information Science (CCIS-2017), vol. 754, pp. 37–47, Volgograd, Russia (2017)
8. Kizim, A.V., Matokhina, A.V., Vayngolts, I.I., Shcherbakov, M.V.: Intelligent platform of monitoring, diagnosis and modernization of technical systems at various stages of life cycle. In: Proceedings of the 5th International Conference on System Modeling and Advancement in Research Trends (SMART 2016), vol. 5, pp. 145–150 (2016)
9. Mobile devices in a corporate network: management and control. Access mode: https://www.kaspersky.ru/blog/mobil-ny-e-ustrojstva-v-korporativnoj-seti-upravlenie-i-kontrol/14809/. Last accessed 25 Jan 2022
10. Kravets, A.G., Salnikova, N.A., Shestopalova, E.L.: Development of a module for predictive modeling of technological development trends. In: Studies in Systems, Decision and Control, vol. 350, pp. 125–136 (2021)
11. Tyukov, A., Shcherbakov, M., Brebels, A.: Automatic two way synchronization between server and multiple clients for HVAC system. In: ACM International Conference Proceeding Series. 13th International Conference on Information Integration and Web-Based Applications and Services (IIWAS 2011), Proceedings, pp. 467–470 (2011)
12. Khanova, A.A., Protalinsky, O.M., Dorokhov, A.F., Bolshakov, A.A.: Socio-economic systems strategic management concept based on simulation. World Appl. Sci. J. **24**(24), 74–79 (2013)
13. Kravets, A.G., Skorobogatchenko, D.A., Salnikova, N.A., Orudjev, N.Y., Poplavskaya, O.V.: The traffic safety management system in urban conditions based on the C4.5 algorithm. In: 2018 Moscow Workshop on Electronic and Networking Technologies, MWENT 2018—Proceedings, 2018-March, art. no. 8337254, pp. 1–7 (2018)
14. Bolshakov, A., Veshneva, I., Lushin, D.: Mathematical model of integration of cyber-physical systems for solving problems of increasing the competitiveness of the regions of the Russian Federation. In: Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, pp. 129–139. Springer, Cham (2021)
15. Kravets, A.G., Fomenkov, S.A., Kravets, A.D.: Component-based approach to multi-agent system generation. In: Communications in Computer and Information Science, vol. 466 CCIS, pp. 483–490 (2014)
16. Kravets, A., Salnikova, N., Medintseva, I.P., Shinkaruk, V.: Development of a method for intellectual support of inventive activity based on deep machine learning. In: Communications in Computer and Information Science, vol. 1448 CCIS, pp. 29–44 (2021)

17. Bolshakov, A., Veshneva, I., Lushin, D.: Mathematical model of integration of cyber-physical systems for solving problems of increasing the competitiveness of the regions of the Russian Federation. In: Society 5.0: Cyberspace for Advanced Human-Centered Society (Studies in Systems, Decision and Control), vol. 335, pp. 129–140 (2021)
18. Shcherbakov, M.V., Shcherbakova, N.L., Kamaev, V.A., Brebels, A., Gerget, O.M., Devyatykh, D.: Outlier detection and classification in sensor data streams for proactive decision support systems. J. Phys. Conf. Ser., 012143 (2017)
19. Korobkin, D.M., Fomenkov, S.A., Kolesnikov, S.G., Orlova, Y.A.: A multi-stage algorithm for text documents filtering based on physical knowledge. World Appl. Sci. J. **24**(24), 91–97 (2013)
20. Oracle Mobile Security Suite. Access mode: https://docs.oracle.com/cd/E52734_01/omss/MOBAD/understanding.htm#MOBAD162. Last accessed 25 Jan 2022
21. Azure Active Directory. Access mode: https://docs.microsoft.com/ru-ru/azure/active-directory/fundamentals/active-directory-manage-groups. Last accessed 25 Jan 2022
22. Cisco DNA. Access mode: https://www.cisco.com/c/damglobal/ru_ru/products/cloud-systems-management/dna-center/nb-09-dna-center-so-cte-ru.pdf. Last accessed 25 Jan 2022
23. G Suite by Google Cloud. Access mode: https://gsuite.google.com/features/. Last accessed 25 Jan 2022
24. Korobkin, D.M., Kolesnikov, S.G., Fomenkov, S.A., Golovanchikov, A.B.: Evaluating an operation principle practicability of new technical system. In: 6th International Conference on Information, Intelligence, Systems and Applications (IISA 2015), vol. 6, p. 7388068 (2015)
25. Infological data model "entity-relationship" [Electronic resource]. Access mode: https://spravochnick.ru/bazy_dannyh/infologicheskaya_model_dannyh_suschnost-svyaz_osnovnye_ponyatiya/. Last accessed 25 Jan 2022

# Scaling Networks and Capturing Keys Using Combined Systems of Sets

**Alexander Frolov** and **Natalya Kochetova**

**Abstract** This chapter discusses a united approach to detecting blocks containing an uncompromised key in wireless sensor networks with a structure of combinatorial designs and scaling those networks. It consists in using both individual blocks of combinatorial designs and collections of blocks containing a certain element. Such collections of blocks are called dual blocks. If the elements are interpreted as key identifiers, then knowledge of the composition of the dual block, combined with information about the placement of blocks, allows planning operations to capture blocks with the key of interest to the invader. On the other hand, the chapter proposes to construct a network with a combinatorial design structure consisting of fully or partially connected subnets corresponding to blocks, and possessing connections corresponding to dual blocks, thereby providing a scaling step. The resulting network has a structure of a combined combinatorial design that, along with blocks, contains dual blocks; its nodes correspond to pairs of blocks and dual blocks. The chapter gives examples and estimates of scaling when converting a combinatorial design to a combined combinatorial design.

**Keywords** A system of sets · Combinatorial design · Combined combinatorial design · Combinatorial network structure · Internal subnet · External subnet · Selective node capture · Network scaling

## 1 Introduction

In the era of Industry 4.0, the Internet of Things has emerged and is developing based on the widespread use of secure Wireless Sensor Network (WSN) structures. There is established practice and technology for building a WSN, in particular for

A. Frolov (✉) · N. Kochetova
National Research University Moscow Power Engineering Institute, Moscow, Russian Federation
e-mail: frolovab@mpei.ru

N. Kochetova
e-mail: kochetovanp@mpei.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_21

a “smart” home, based on the IEEE 802.15.4 standard [1], the routing protocol in ZigBee [2–4]. At the same time, resources are used that are limited both in terms of energy consumption and the amount of memory required [5, 6]. The security of such systems can be ensured by using symmetric encryption systems [7] based on the preliminary distribution of keys in the network nodes [8–10]. Sensor network key management includes the following steps [11]: Key Pre-distribution, Shared Key Discovery, and Determining the key delivery path (Path-key Establishment). Currently, the algorithms of these stages are carried out using certain types of combinatorial block designs, which are specific systems of sets defined by a set of elements and a set of certain subsets (blocks) of it [11]. In parallel, also with the use of set systems, network structures are studied in [12, 13]. Combinatorial block-designs as mathematical models of computer nets are implemented in [14, 15].

In this chapter, we consider the problems of detecting blocks containing an uncompromised key in a wireless sensor network possessing a combinatorial design structure and scaling those networks. We propose a united approach to solving them using both combinatorial design blocks and collections of blocks containing a common element. Those collections are called dual blocks. During selective node capture at each step, the attacker knows the set K of compromised keys and set T of key numbers that have not already been compromised [16, 17]. To compromise some key that number belongs to T, the attacker should capture the node containing this key. Thus it can capture any appropriate node from the dual block corresponding to that key number. Considering a number of well-known approaches to scaling networks by modifying known combinatorial designs, we propose to supplement such structures with structures based on dual blocks and evaluate the resulting structure compared to the original one.

Section 2 is devoted to combinatorial designs used in key pre-distribution systems in WSN. In Sect. 3, we consider combined counterparts of such combinatorial designs. In Sect. 4, we continue the discussion on capturing an uncompromised key. Section 5 is devoted to highly scalable networking structures and evaluations of our proposals. Section 6 is a conclusion.

## 2 Some Combinatorial Designs Used in WSN

In this section, we recall basic concepts of combinatorial designs and consider examples of combinatorial designs and their known modifications proposed for scaling WSN.

The set system ($\mathbf{X}$, $\mathbf{B}$) is specified by a finite set $\mathbf{X}$ of elements and the collection $\mathbf{B}$ of its subsets [11]. Varieties of set systems currently used for pre-key distribution are Balanced Incomplete Block Designs (BIBD) with parameters ($v$, $b$, $r$, $k$), where $v$ is the number of elements, $b$ is the number of blocks, $r$ is the number of blocks containing any given element, $k$ is the number of elements of any block. Additionally, the fifth parameter $\lambda$ is used as the number of blocks, each containing a given pair of elements. If $v = b$, BIBD is symmetric with parameters ($v$, $k$) and additional parameter λ. It is abbreviated SBIBD. SBIBD with parameters $\left(n^2+n+1, n+1, 1\right)$ is called a projective plane (or geometry) of order $n$ [12]. It is denoted PG(2, $n$). There are known algorithms for construction PG(2, $n$) for any prime power $n$. BIBD $\left(n^2, n^2+n, n, 1\right)$ is called an affine plane of order $n$; it is residual BIBD obtained from PG(2, $n$) by removing any block and removing elements of this block from the remaining blocks. So from each PG(2, $n$), one can get $n^2+n+1$ differing affine planes. Union of all affine planes corresponding to given PG(2, $n$) form the set of blocks of modified design, introduced in [18].

***Example 1*** Let us consider cyclic projective geometry PG(2, 2) = (**X**, **B**) [12], where

$$\begin{aligned}\mathbf{X} &= \{0, 1, 2, 3, 4, 5, 6\},\\ \mathbf{B} &= \{\{0, 1, 5\}, \{1, 2, 6\}, \{2, 3, 0\}, \{3, 4, 1\}, \{4, 5, 2\}, \{5, 6, 3\}, \{6, 0, 4\}\}.\end{aligned}$$

There are 21 blocks of modified residual design MRPG(2, 2), corresponding to PG(2, 2) on the same set **X** [18]

$$\begin{aligned}\mathbf{B} = \{&\{0, 1\}, \{0, 2\}, \{0, 3\}, \{0, 4\}, \{0, 5\}, \{0, 6\}, \{1, 2\},\\ &\{1, 3\}, \{1, 4\}, \{1, 5\}, \{1, 6\}, \{2, 3\}, \{2, 4\}, \{2, 5\},\\ &\{2, 6\}, \{3, 4\}, \{3, 5\}, \{3, 6\}, \{4, 5\}, \{4, 6\}, \{5, 6\}\}.\end{aligned}$$

The number of blocks of MRPG(2, $n$) is defined by formula [18]

$$|\mathbf{B}| = (n^2+n+1)(n+1)$$

(instead $(n^2+n+1)$ blocks of PG(2, $n$)).

Next, consider the transversal combinatorial design. The linear transversal combinatorial design TD($k$, $n$) [11] is constructed from elements of $k$ pairwise disjoint sets $w_i$ containing $n$ elements each, and consists of $n^2$ sets $Y_j, j=\overline{(1, n^2)}$, each containing $k$ elements. Moreover, each set $Y_j$ has exactly one element from each set $w_i$, $i = 0$, …, $k-1$, and if $j \neq s$, then $Y_j$ and $Y_s$ have at most one common element. Thus, the combinatorial design TD($k$, $n$) is defined on the set **X** of cardinality $kn$. If $n$ is a prime or a prime power, then TD($k$, $n$) exists. Then, in the algebraic interpretation, the set X is the set of pairs ($y$, $x$), $x \in X_1$, where $\mathbf{X}_1$ is the set of $k$ elements of the field $F_n$, $y \in F_n$; the blocks of the linear transversal combinatorial design TD($k$, $n$) are determined by a pair of elements $(a, b) \in GF(n^2)$ [11]:

$$B_{a,b} = \{(ax+b, x) : x \in \mathbf{X}_1\}.$$

To each block from TD($k$, $n$), we associate the set of non-empty subsets of lower cardinality obtained from other blocks by extracting this block. Combining the resulting sets, we obtain the residual transversal combinatorial block design RTD($k$, $n$), corresponding to TD($k$, $n$) [19].

***Example 2*** Let us consider the transversal combinatorial design TD(3, 3) = (**X, B**), where

$$\mathbf{X} = \{00,\ 01,\ 02, 10, 11, 12, 20, 21, 22\};$$
$$\mathbf{B} = \{\{00, 10, 20\}, \{01, 11, 22\}, \{02, 12, 21\}, \{00, 11, 21\}, \{01, 12, 20\}, \{02, 10, 22\}, \{00, 12, 22\}, \{01, 10, 21\}, \{02, 11, 20\}\}.$$

The set of blocks of corresponding residual transversal combinatorial block design RTD(3, 3) (there are used only blocks of the same rank) is the following

$$\mathbf{B} = \{\{11, 21\}, \{01, 12\}, \{02, 22\}, \{12, 22\}, \{01, 21\}, \{02, 11\}, \{00, 21\}, \{12, 20\}, \{02, 10\}, \{00, 12\}, \{10, 21\}, \{02, 20\}, \{10, 20\}, \{01, 22\}, \{02, 12\}, \{01, 10\}, \{00, 10\}, \{11, 22\}, \{02, 21\}, \{00, 22\}, \{00, 20\}, \{01, 11\}, \{12, 21\}, \{11, 20\}, \{01, 20\}, \{00, 11\}, \{10, 22\}\}.$$

The number of blocks of so defined residual transversal combinatorial block design RTD($k$, $n$) corresponding to TD($k$, $n$) equals $n^3$ instead of $n^2$ blocks of initial TD($k$, $n$).

Note that the use of projective geometry keys for keys pre-distribution is limited by the fact that the number of elements in the block is greater than the square root of the number of nodes. This is the reason for the preference for using transversal combinatorial block designs. In order to increase the resistance to the compromise of keys, in [16, 17] the network is divided into cells in which different key pools are used. At the same time, a decrease in the number of keys in the nodes is achieved. Projective geometry is used for key distribution in cells and transversal combinatorial block designs for communication between cells.

## 3 Combined Combinatorial Designs

The system of sets (**X**, **B**) is associated with a combined system of sets (**X**, **B**, **D**, **M**), in which two more sets are considered. The set **D** of duality blocks is in one-to-one correspondence with the set **X**. Duality blocks are sets of blocks containing an element corresponding to a given duality block. The main set **M** is the set of all distinct pairs of a form (a block, an element of the block) [20].

***Example 3*** Combined projective geometry CPG(2, 2) = (**X**, **B**, **D**, **M**) corresponding to PG(2, 2) is defined by sets

$$\mathbf{X} = \{0, 1, 2, 3, 4, 5, 6\};$$
$$\mathbf{B} = \{\{0, 1, 5\}, \{1, 2, 6\}, \{2, 3, 0\}, \{3, 4, 1\}, \{4, 5, 2\}, \{5, 6, 3\}, \{6, 0, 4\}\}.$$
$$\mathbf{D} = \{\{\{0, 1, 5\}, \{2, 3, 0\}, \{6, 0, 4\}\}, \{\{0, 1, 5\}, \{1, 2, 6\}, \{3, 4, 1\}\}, \{\{1, 2, 6\}, \{2, 3, 0\}, \{4, 5, 2\}\}, \{\{2, 3, 0\}, \{3, 4, 1\}, \{5, 6, 3\}\},$$

$$\{\{3,4,1\},\{4,5,2\},\{6,0,4\}\},\{\{0,1,5\},\{4,5,2\},\{5,6,3\}\},$$
$$\{\{1,2,6\},\{5,6,3\},\{6,0,4\}\}\}.$$

$$\begin{aligned}\mathbf{M} = \{&(\{0,1,5\},0),(\{0,1,5\},1),(\{0,1,5\},5),\\&(\{1,2,6\},1),(\{1,2,6\},2),(\{1,2,6\},6),\\&(\{3,4,1\},3),(\{3,4,1\},4),(\{3,4,1\},1),\\&(\{4,5,2\},4),(\{4,5,2\},5),(\{4,5,2\},2),\\&(\{5,6,3\},5),(\{5,6,3\},6),(\{5,6,3\},3),\\&(\{6,0,4\},6),(\{6,0,4\},0),(\{6,0,4\},4)\}.\end{aligned} \tag{1}$$

The set **D** of duality blocks of combined modified residual design CMRPG(2, 2) corresponding to modified residual design MRPG(2, 2) from Example 1

$$\begin{aligned}\mathbf{D} = \{&\{0,1\},\{0,2\},\{0,3\},\{0,4\},\{0,5\},\{0,6\}\},\{\{0,1\},\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{1,6\}\},\\&\{\{0,2\},\{1,2\},\{2,3\},\{2,4\},\{2,5\},\{2,6\}\},\{\{0,3\},\{1,3\},\{2,3\},\{3,4\},\{3,5\},\{3,6\}\},\\&\{\{0,4\},\{1,4\},\{2,4\},\{3,4\},\{4,5\},\{4,6\}\},\{\{0,5\},\{1,5\},\{2,5\},\{3,5\},\{4,5\},\{5,6\}\},\\&\{\{0,6\},\{1,6\},\{2,6\},\{3,6\},\{4,6\},\{5,6\}\}.\end{aligned}$$

The main set of this combined modified residual design CMRPG(2, 2) corresponding to MRPG(2, 2) is the following

$$\begin{aligned}\mathbf{M} = \{&(\{0,1\},0),(\{0,1\},1),(\{0,2\},0),(\{0,2\},2),(\{0,3\},0),\\&(\{0,3\},3),(\{0,4\},0),(\{0,4\},4),(\{0,5\},0),(\{0,5\},5),(\{0,6\},0),\\&(\{0,6\},6),(\{1,2\},1),(\{1,2\},2),(\{1,3\},1),(\{1,3\},3),\\&(\{1,4\},1),(\{1,4\},4),(\{1,5\},1),(\{1,5\},5),(\{1,6\},1),(\{1,6\},6),\\&(\{2,3\},2),(\{2,3\},3),(\{2,4\},2),(\{2,4\},4),(\{2,5\},2),(\{2,5\},5),\\&(\{2,6\},2),(\{2,6\},6),(\{3,4\},3),(\{3,4\},4),(\{3,5\},3),(\{3,5\},5),\\&(\{3,6\},3),(\{3,6\},6),(\{4,5\},4),(\{4,5\},5),(\{4,6\},4),\\&(\{4,6\},6),(\{5,6\},5),(\{5,6\},6)\}.\end{aligned} \tag{2}$$

Applied to the main set of combined modified residual design CMRPG(2, $n$), there is a valid formula

$$|\mathbf{M}| = \left(n^2+n+1\right)\left(n^2+n\right).$$

***Example 4*** Combined transversal combinatorial design CTD(3, 3) = (**X**, **B**, **D**, **M**) is defined by the sets

$$\mathbf{X} = \{00,01,02,10,11,12,20,21,22\};$$
$$\begin{aligned}\mathbf{B} = \{&\{00,10,20\},\{01,11,22\},\{02,12,21\},\{00,11,21\},\{01,12,20\},\\&\{02,10,22\},\{00,12,22\},\{01,10,21\},\{02,11,20\}\};\end{aligned}$$
$$\mathbf{D} = \{\{\{00,10,20\},\{00,12,22\}\},\{\{01,11,22\},\{01,12,20\}\},$$

$$
\begin{aligned}
&\{\{02, 12, 21\}, \{02, 10, 22\}\}, \{\{00, 10, 20\}, \{02, 10, 22\}\},\\
&\{\{01, 11, 22\}, \{02, 11, 20\}\}, \{\{02, 12, 21\}, \{00, 12, 22\}\},\\
&\{\{00, 10, 20\}, \{02, 11, 20\}\}, \{\{02, 12, 21\}, \{01, 10, 21\}\},\\
&\{\{02, 10, 22\}, \{00, 12, 22\}\}\}.\\
\mathbf{M} = &\{(\{00, 10, 20\}, 00), (\{00, 10, 20\}, 10\}), (\{00, 10, 20\}, 20), \ldots,\\
&(\{02, 11, 20\}, 02), (\{02, 11, 20\}, 11), (\{02, 11, 20\}, 20)\}. \qquad (3)
\end{aligned}
$$

The set of 9 duality blocks of combined residual transversal combinatorial block design CRTD(3, 3) corresponding to residual transversal combinatorial block design RTD(3, 3) from Example 2 is

$$
\begin{aligned}
\mathbf{D} = \{&\{\{00, 22\}, \{00, 21\}, \{00, 12\}, \{00, 10\}, \{00, 20\}, \{00, 11\}\},\\
&\{\{01, 12\}, \{01, 21\}, \{01, 22\}, \{01, 10\}, \{01, 11\}, \{01, 20\}\},\\
&\{\{02, 22\}, \{02, 11\}, \{02, 10\}, \{02, 20\}, \{02, 12\}, \{02, 21\}\},\\
&\{\{02, 10\}, \{10, 21\}, \{10, 20\}, \{01, 10\}, \{00, 10\}, \{10, 22\}\},\\
&\{\{11, 21\}, \{02, 11\}, \{11, 22\}, \{01, 11\}, \{11, 20\}, \{00, 11\}\},\\
&\{\{01, 12\}, \{12, 22\}, \{12, 20\}, \{00, 12\}, \{02, 12\}, \{12, 21\}\},\\
&\{\{11, 21\}, \{01, 21\}, \{00, 21\}, \{10, 21\}, \{02, 21\}, \{12, 21\}\},\\
&\{\{02, 22\}, \{12, 22\}, \{01, 22\}, \{11, 22\}, \{00, 22\}, \{10, 22\}\}\}.
\end{aligned}
$$

The main set of this combined residual transversal combinatorial block design CRTD(3, 3) consists of 54 pairs

$$\mathbf{M} = \{(\{11, 21\}, 11), (\{11, 21\}, 21), \ldots, (\{10, 22\}, 10), (\{10, 22\}, 22)\}.$$

Applied to the main set of combined modified residual transversal combinatorial design corresponding to transversal combinatorial design TD(2, $k$, $n$), there is a valid formula

$$|\mathbf{M}| = kn^3.$$

## 4 Capturing of Uncompromised Key

To implement the attack of capturing an uncompromised key described in the introduction, it is necessary to know the compositions of the dual blocks. This is achieved by algorithms for block numbering, calculating a block with a given number, and calculating a set of block numbers that make up a dual block corresponding to a given element (key). These algorithms, as applied to projective geometries, Steiner triplets, and linear and quadratic transversal combinatorial designs, are given in the author's papers [20, 21]. Examples of dual blocks that can be used to implement this attack are given in Sect. 3.

## 5 Scaling Based on Combined Combinatorial Designs

In Sect. 2, a well-known scaling technique was implemented based on the use of residual combinatorial block designs. Let us show that a significant increase in scaling is possible based on combined combinatorial designs. We will build a network on a set of nodes corresponding to the elements of the main set of the combined block design. Then the number of nodes will be $k$ times greater than the number of blocks in the original block design, where $k$ is the number of block elements. We construct the network structure as a composition of fully or partially connected networks of two types. Subnets of the first type, let's call them internal networks, are built on a set of nodes represented by pairs with the same first elements, and subnets of the second type (external) are built on sets of nodes with the same second elements. Those internal networks correspond to blocks of the original design, and external subnets correspond to its duality blocks.

***Example 5*** On the main set (1) of combined projective geometry CPG(2, 2), one can construct the network (see Fig. 1) with 7 internal subnets with a structure of $K_3$ (italicized numbers)

$$\{(\{0, 1, 5\}, 0), (\{0, 1, 5\}, 1), (\{0, 1, 5\}, 5)\},$$
$$\{(\{1, 2, 6\}, 1), (\{1, 2, 6\}, 2), (\{1, 2, 6\}, 6)\},$$
$$\{(\{2, 3, 0\}, 2), (\{2, 3, 0\}, 3), (\{2, 3, 0\}, 0)\},$$
$$\{(\{3, 4, 1\}, 3), (\{3, 4, 1\}, 4), (\{3, 4, 1\}, 1)\},$$
$$\{(\{4, 5, 2\}, 4), (\{4, 5, 2\}, 5), (\{4, 5, 2\}, 2)\},$$
$$\{(\{5, 6, 3\}, 5), (\{5, 6, 3\}, 6), (\{5, 6, 3\}, 3)\},$$
$$\{(\{6, 0, 4\}, 6), (\{6, 0, 4\}, 0), (\{6, 0, 4\}, 4)\}.$$

and 7 external subnets with the same structure (in Fig. 1, their nodes are connected by dotted lines)

$$\{(\{0, 1, 5\}, 0), (\{2, 3, 0\}, 0), (\{6, 0, 4\}, 0)\},$$
$$\{(\{0, 1, 5\}, 1), (\{1, 2, 6\}, 1), (\{3, 4, 1\}, 1)\},$$
$$\{(\{1, 2, 6\}, 2), (\{2, 3, 0\}, 2), (\{4, 5, 2\}, 2)\},$$
$$\{(\{2, 3, 0\}, 3), (\{3, 4, 1\}, 3), (\{5, 6, 3\}, 3)\},$$
$$\{(\{3, 4, 1\}, 4), (\{4, 5, 2\}, 4), (\{6, 0, 4\}, 4)\},$$
$$\{(\{0, 1, 5\}, 5), (\{4, 5, 2\}, 5), (\{5, 6, 3\}, 5)\},$$
$$\{(\{1, 2, 6\}, 6), (\{5, 6, 3\}, 6), (\{6, 0, 4\}, 6)\}.$$

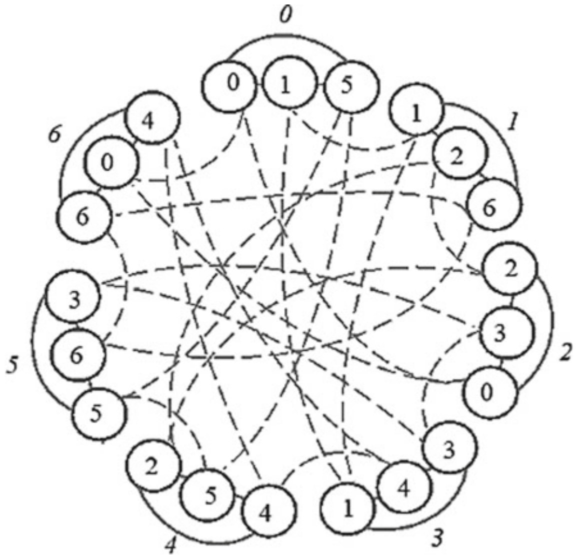



**Fig. 1** The network with the structure of combined projective geometry CPG(2, 2) (Example 3). It contains 21 nodes. Numbers of internal subnets are shown in italics, external subnets are shown by dotted lines. In cycles are given the numbers of external subnets containing this node

***Example 6*** On the main set (2) of combined modified residual design CMRPG(2, 2) from Example 3, one can get 21 internal subnets with structure $K_2$

$$\{(\{0,1\},0),(\{0,1\},1)\},\{(\{0,2\},0),(\{0,2\},2)\},$$
$$\{(\{0,3\},0),(\{0,3\},3)\},\{(\{0,4\},0),(\{0,4\},4)\},$$
$$\{(\{0,5\},0),(\{0,5\},5)\},\{(\{0,6\},0),(\{0,6\},6)\},$$
$$\{(\{1,2\},1),(\{1,2\},2)\},\{(\{1,3\},1),(\{1,3\},3)\},$$
$$\{(\{1,4\},1),(\{1,4\},4)\},\{(\{1,5\},1),(\{1,5\},5)\},$$
$$\{(\{1,6\},1),(\{1,6\},6)\},\{(\{2,3\},2),(\{2,3\},3)\},$$
$$\{(\{2,4\},2),(\{2,4\},4)\},\{(\{2,5\},2),(\{2,5\},5)\},$$
$$\{(\{2,6\},2),(\{2,6\},6)\},\{(\{3,4\},3),(\{3,4\},4)\},$$
$$\{(\{3,5\},3),(\{3,5\},5)\},\{(\{3,6\},3),(\{3,6\},6)\},$$
$$\{(\{4,5\},4),(\{4,5\},5)\},\{(\{4,6\},4),(\{4,6\},6)\},$$
$$\{(\{5,6\},5),(\{5,6\},6)\}.$$

and 7 external subnets of structure $K_6$

$$\{(\{0,1\},0),(\{0,2\},0),(\{0,3\},0),(\{0,4\},0),(\{0,5\},0),(\{0,6\},0)\},$$
$$\{(\{0,1\},1),(\{1,2\},1),(\{1,3\},1),(\{1,4\},1),(\{1,5\},1),(\{1,6\},1)\},$$
$$\{(\{0,2\},2),(\{1,2\},2),(\{2,3\},2),(\{2,4\},2),(\{2,5\},2),(\{2,6\},2)\},$$
$$\{(\{0,3\},3),(\{1,3\},3),(\{2,3\},3),(\{3,4\},3),(\{3,5\},3),(\{3,6\},3)\},$$
$$\{(\{0,4\},4),(\{1,4\},4),(\{2,4\},4),(\{3,4\},4),(\{4,5\},4),(\{4,6\},4)\},$$

$$\{(\{0,5\},5),(\{1,5\},5),(\{2,5\},5),(\{3,5\},5),(\{4,5\},5),(\{5,6\},5)\},$$
$$\{(\{0,6\},6),(\{1,6\},6),(\{2,6\},6),(\{3,6\},6),(\{4,6\},6),(\{4,6\},6)\}.$$

***Example 7*** On the main set (3) of combined transversal combinatorial design CTD(3, 3) from Example 4, one can get 9 internal subnets with structure $K_3$

$$\{(\{00,10,20\},00),(\{00,10,20\},10),(\{00,10,20\},20)\},$$
$$\ldots,$$
$$\{(\{02,11,20\},02),(\{02,11,20\},11),(\{02,11,20\},20)\}$$

and 9 external subnets with the same structure

$$\{(\{00,10,20\},00),(\{02,10,22\},00),(\{00,12,22\},00)\},$$
$$\ldots,$$
$$\{(\{01,11,22\},22),(\{02,11,20\}22,),(\{00,12,22\},22)\}.$$

***Example 8*** On the main set (4) of combined modified residual transversal combinatorial design CRTD(3, 3) from Example 4, one can get 27 internal subnets with structure $K_2$

$$\{(\{11,21\},11),(\{11,21\},21)\},\{(\{01,12\},01),(\{01,12\},12)\},\ldots$$
$$\ldots,\{(\{00,11\},00),(\{00,11\},11)\},\{\{10,22\},10),(\{10,22\},22)\}$$

and 9 external subnets of structure $K_6$

$$\{(\{00,21\},00),(\{00,12\},00),(\{00,10\},00),(\{00,22\},00),(\{00,20\},00),(\{00,11\},00)\},$$
$$\{(\{01,12\},01),(\{01,21\},01),(\{01,22\},01),(\{01,10\},01),(\{01,11\},01),(\{01,20\},01)\},$$
$$\{(\{02,22\},02),(\{02,11\},02),(\{02,10\},02),(\{02,20\},02),(\{02,12\},02),(\{02,21\},02)\},$$
$$\{(\{02,10\},10),(\{10,21\},10),(\{10,20\},10),(\{01,10\},10),(\{00,10\},10),(\{10,22\},10)\},$$
$$\{(\{11,21\},11),(\{02,11\},11),(\{11,22\},11),(\{01,11\},11),(\{11,20\},11),(\{00,11\},11)\},$$
$$\{(\{01,12\},12),(\{12,22\},12),(\{12,20\},12),(\{00,12\},12),(\{02,12\},12),(\{12,21\},12)\},$$
$$\{(\{12,20\},20),(\{02,20\},20),(\{10,20\},20),(\{00,20\},20),(\{11,20\},20),(\{01,20\},20)\},$$
$$\{(\{11,21\},21),(\{01,21\},21),(\{00,21\},21),(\{10,21\},21),(\{02,21\},21),(\{12,21\},21)\},$$
$$\{(\{02,22\},22),(\{12,22\},22),(\{01,22\},22),(\{11,22\},22),(\{00,22\},22),(\{10,22\},22)\}.$$

The network structures described in this section have important properties:

- sets of nodes of internal subnets form a partition of the set of all nodes into subsets of the same cardinality;
- similarly, the sets of nodes of external subnets also form a partition of the set of all nodes into subsets of the same cardinality;
- each network element is an element of some internal and some external subnets.

**Table 1** Numbers of net and subnet nodes

| | PG(2, $n$) | CPG(2, $n$) | MRPG(2, $n$) | CMRPG(2, $n$) |
|---|---|---|---|---|
| Net | $(n^2+n+1)$ | $(n^2+n+1)(n+1)$ | $(n^2+n+1)(n+1)$ | $(n^2+n+1)(n^2+n)$ |
| Internal subnet | – | $n+1$ | – | $n$ |
| External subnet | – | $n+1$ | – | $(n^2+n)$ |

**Table 2** Numbers of net and subnet nodes

| | TD(2, $k$, $n$) | CTD($k$, $n$) | RTD($k$, $n$) | CRTD($k$, $n$) |
|---|---|---|---|---|
| Net | $n^2$ | $kn^2$ | $n^3$ | $kn^3$ |
| Internal subnet | – | $k$ | – | $k-1$ |
| External subnet | – | $n$ | – | $(k-1)n$ |

From a key pre-distribution perspective, these properties allow different key pools and different combinatorial models to distribute keys across subnets. For example, smaller subnets can use projective planes, and higher subnets use transversal combinatorial designs (as in [16, 17]). Each node uses two sets of keys to communicate on its respective internal and external subnets. This achieves a decrease in the amount of memory required for storing keys and also protection against the compromise of keys. At this stage of scaling, the nodes use key pools of the same capacity, and key pools are protected from mutual compromise. In Tables 1 and 2, there are given numbers of nodes in considered nets and their separate subnets.

## 6 Conclusion

In this chapter, we consider two important practice problems of capturing an uncompromised key and scaling a computer network whose structure corresponds to a combinatorial design. For the first of them, a set of nodes containing the key of interest to the invader is determined. For the second, there are used fully or partially connected, subnets the composition of which is determined by blocks of combinatorial design. The inside connections are supplemented by external connections of computers corresponding to the same elements of the blocks. Algorithmically, the solution is achieved based on the concept of a combined combinatorial block design when a set of duality blocks is used along with a set of blocks. This made it possible, by scaling the sets of elements of two well-known combinatorial schemes, to build new, larger-scale network structures that have properties characteristic of the practice of preliminary key distribution. We also obtained the number of nodes in such network structures and in their subnets corresponding to blocks and dual blocks.

## References

1. Gholami, K.El., Yassine, M. Y., Fatani, I.: The IEEE 802.15.4 Standard in industrial applications: a survey. J. Theor. Appl. Inf. Technol. **99**(15), 1–17 (2021)
2. Kim, S.H., Chong, P.K., Kim, T.: Performance study of routing protocols in ZigBee wireless mesh networks. Wireless Pers. Commun. **95**, 1829–1853 (2017)
3. Kim, T., Kim, S.H., Yang, J., Yoo, S.e., Kim, D.: Neighbor table-based shortcut tree routing in zigbee wireless networks. IEEE Trans. Parallel Distrib. Syst. **25**(3), 706–716 (2014)
4. Gomez, C., Paradells, J.: Wireless home automation networks: a survey of architectures and technologies. IEEE Commun. Mag. **48**(6), 92–101 (2010)
5. Solari Esfehani, N., Haj Seyyed Javadi, H.: A survey of key pre-distribution schemes based on combinatorial designs for resource-constrained devices in the IoT network. Wirel. Netw. **27**(4), 3025–3052 (2021). https://doi.org/10.1007/s11276-021-02629-8
6. Bahrami, P.N., Javadi, H.H.S., Dargahi, T., Dehghantanha, A., Choo, K.K.R.: A hierarchical key pre-distribution scheme for fog networks. Concurrency Comput. Pract. Experience **31**(22) (2019). https://doi.org/10.1002/cpe.4776
7. GOST R (Russian: ГОСТ Р) 34.12-2015. Information technology. Cryptographic data security. Block ciphers. Standartinform, Moscow (2015)
8. Camtepe, S.A., Yener, B.: Combinatorial design of key distribution mechanisms for wireless sensor networks. IEEE/ACM Trans. Netw. (TON) **15**, 293–308 (2004). https://doi.org/10.1007/978-3-540-30108-0_18
9. Du, W., Deng, J., Han, Y., Varshney, P., Katz, J., Khalili, A.: A pairwise key pre distribution scheme for wireless sensor networks. ACM Trans. Inform. Syst. Secur. **8**, 228–258 (2005). https://doi.org/10.1145/948109.948118
10. Urivskiy, A.: Key predistribution scheme using affine planes and Blom's scheme. In: Proceedings of International Mathematical Conference "50 Years of IITP" (2011). ISBN 978-5-901158-15-9
11. Lee, J., Stinson, D.R.: On the construction of practical key predistribution schemes for distributed sensor networks using combinatorial designs. ACM Trans. Inf. Syst. Secur. **11**(2). https://doi.org/10.1145/1330332.1330333
12. Stinson, D.R.: Combinatorial designs: constructions and analysis. Springer, Berlin (2007)
13. Tarannikov, Yu.V.: Combinatorial Properties of Discrete Structures and Applications to Cryptology. MCCME, Moscow (2014). ISBN 978-5-4439-2045-0 (In Russian)
14. Parkhomenko, P.P.: Algorithmizing design of a class of combinatorial block diagrams. Autom. Remote. Control. **77**(7), 1216–1224 (2016). https://doi.org/10.1134/S0005117916070080
15. Mozharov, G.P.: Fault-tolerant computer networks constructed on the basis of combinatorial block designs. Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Priborostr [Herald of the Bauman Moscow State Tech. Univ., Instrum. Eng.] **6**, 41–53 (2016). https://doi.org/10.18698/0236-3933-2016-6-41-53. (In Russian)
16. Ruj, S., Roy, B.: Key predistribution using combinatorial designs for grid-group deployment scheme in wireless sensor networks. ACM Trans. Sens. Netw. **6**(1) (2009). https://doi.org/10.1145/1653760.1653764
17. Kumar, A., Pais, A.R.: A new combinatorial design based key pre-distribution scheme for wireless sensor networks. J. Ambient. Intell. Humaniz. Comput. **10**(6), 2401–2416 (2019). https://doi.org/10.1007/s12652-018-0902-4
18. Modiri, V., Javadi, H.H.S., Anzani, M.: A novel scalable key pre-distribution scheme for wireless sensor networks based on residual design. Wireless Pers. Commun. **96**(2), 2821–2841 (2017). https://doi.org/10.1007/s11277-017-4326-9

19. Masaeli, N., Javadi, H.H.S., Erfani, S.H.: Key pre-distribution scheme based on transversal design in large mobile fog networks with multi-clouds. J. Inf. Secur. Appl. **54** (2020). https://doi.org/10.1016/j.jisa.2020.102519
20. Frolov, A., Kochetova, N.: Networks with the structure of a combined combinatorial block design. In: International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), pp. 1–6 (2021). https://doi.org/10.1109/ICECCME52200.2021.9591136
21. Frolov, A., Klyagin, A., Kochetova, N., Temnikov, D.: On the synthesis of systems possessing the structure of some combinatorial designs. J. Comput. Syst. Sci. Int. **60**(4), 585–594 (2021). https://doi.org/10.1134/S1064230721040067

# Video Conferencing Subsystem of the Secure Serverless Internet (TheOoL.Net)

**A. V. Nenashev and A. Yu. Tolstenko**

**Abstract** The article discusses prerequisites for the development of the private serverless internet with a built-in super secure video and audio communications subsystem featuring the possibility of holding conferences with three or more participants. A brief description of the principles underlying the private serverless internet architecture is given, which, in addition to its overlay data transfer protocol, includes serverless hosting, cloud computing, serverless search engine, as well as an integrated payment system. A simplified graphical model of the architecture and operation of the serverless super secure video and audio communications subsystem is presented. The basic modes of its operation and the methods for ensuring the security and privacy of conversations and correspondence are reviewed.

**Keywords** Parametric synthesis · Automatic control system · PID controller · Modal method · Poles · Transient characteristic

## 1 Introduction

The recent processes of merging the virtual space of the internet reflecting the mental space of individuals and society [1], not only with the global financial system and the social management system, but also through IoT (the Internet of Things) and IoE (the Internet of Everything), with control systems for technical objects in the physical world, result in the merging of virtual space and real space [2, 3]. This has been especially visible since 2020 thanks to COVID-19 pandemic and the global lockdown, which had substantially accelerated the process of creating remote jobs and their automation [4]. As a result, there has been an explosive growth in the demand for the internet communication services and secure remote access to corporate information

A. V. Nenashev (✉) · A. Yu. Tolstenko
Samara State Technical University, Samara, Russia
e-mail: alexvlnenashev@gmail.com

A. Yu. Tolstenko
e-mail: tolstenko.ay@samgtu.ru



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_22

systems, and significantly increased demand for information security [5]. In fact, the information security market has been growing continuously for at least the past twenty years at an annual average of 7%, which was the economic rationale for the development of the secure serverless internet (TheOoL project) and the serverless secure videoconferencing system as its integral part.

The technical prerequisites for the project were, for one thing, the inherent vulnerabilities of networks based on TCP/IP stack of protocols, which were initially open and decentralized [6–8] and which offer a wide array of built-in capabilities for web intelligence at the third to seventh layers of the OSI/ISO [7, 9–14] model implementation and acquisition of information about operating systems and applications running on specific network nodes and their hardware in order to identify potentially vulnerable components. For another, their inconsistence with the architecture of most mod ern-day data storage and processing systems that are designed and engineered as centralized systems [15]. It is obvious that such possession of information about the software- and hardware composition of network nodes and network topology implies significant information security risks [3, 16–18], and these risks materialize with amazing regularity [19–22].

Several dozen companies and informal development teams have been working on solving the security problems of networking as part of Web 2.0 [23] concept. In our view, the developers of I2P [24] ‘invisible internet’ project came up with the best solution for the development of private, overlay internet traffic network, which despite a number of well-known vulnerabilities [24, 25] ensures high level of internet traffic security, unlike its direct competitor—TOR project [26], the vulnerabilities of which allow to solve the problem of secure connection compromise using limited resources [27–31]. Despite popular appeal of remote work, which is rather impossible without videoconferencing, there are virtually no public systems for private video communications. Existing commercial services are uniformly designed with centralized architecture with complete traffic decryption on the server side [32], which in itself is a vulnerability from the standpoint of a customer, especially in case of confidential conversations. Effectively, owners of the most widely used systems, including ZOOM, Skype, etc., as well as their employees have full access to conversations of their customers. We believe the second source of threat to customer privacy in such videoconferencing systems is the centralized services for the generation and management of cryptographic key and certificates they use. The problem is that, in addition to customers involved in conversations, the owners and maintenance teams of such cryptography services, to a varying degree, have the certificates and keys as required to the decryption of video conference traffic, and this substantially raises the possibility of interception of such conversation by the relevant owners or maintenance staff. In general our analysis of currently available solutions for private networks revealed a number of shared problems: Slow access to content as compared with the ‘open internet’; lack of trusted private hosting, which means that for their protected resources one has to use servers and data links with inferior capacity; lack of trusted protocols for setting up search directories, which results in the use of centralized services, the registration in which entails disclosure of information about topology of the network and geographic location of the servers; the problem of security of audio and video communications has no solution at the time of writing this article. In our view, the solution to the above mentioned problems lies in the transition from Web 2.0 paradigm to Web3/Web 3.0 [33, 34] paradigm, i.e. the transition from topology-oriented, centralized service architecture to its complete decentralization. In this paradigm, implementation of this video and audio communications system is a separate challenge, as it requires a high capacity and security subsystem for distributed computing, complex and time-efficient cryptography, as well as a trusted traffic dispatch and routing system. Yet if implemented this subsystem would allow for restricting access to cryptographic keys and certificates to participants of the respective videoconferences thereby significantly reducing the likelihood of interception of conversations by eliminating every other party from the operation of such video and audio communications, except for the immediate participants of conversations. The use of the distributed computing system makes it possible to further enhance conversation privacy due to planned migration of the components of a decentralized video server to randomized network nodes.

## 2 The Decentralized Internet by TheOoL

As the private video and audio communications system ("SoPVAC") is not a standalone product but rather a subsystem within the ecosystem of TheOoL private decentralized internet [35] it is difficult to describe this system without a brief over view of the architecture and functionality of TheOoL ecosystem. TheOoL is a variation of the implementation of the internet in Web3 paradigm [34] with Web 3.0 elements [33], and it is primarily focused on ensuring privacy for the network members, which implies the possibility of complete suppression of digital footprint or generation of a false trail under complete control of the user, the user's total and exclusive possession of all the cryptographic keys and the user's sole right to govern any data as generated in the course of their activity. In TheOoL network there are no server owners or service owners; it is a completely serverless, peer-to-peer distributed system managed by algorithms, which implement the processes of operation and supply of the network resources in the form of a decentralized autonomous organization or DAO [36]. On the engineering side, TheOoL is a geographically distributed, high-end automated queuing network [37]. At the logical level, it is a data and computing cloud based on the private and secure decentralized ecosystem ("PSDE") (Fig. 1) under pinned by the following principles:

- Direct user interaction with user information via rank-order information card file.
- Robustness of user information and impossibility of its dissemination, modification or deletion without consent of the information owner or persons authorized by the owner.
- Availability of user information via the automated, decentralized, information custom card file.

- Controllability of user information, which implies total control over the user's digital footprint on the user side and impossibility of data acquisition by anyone else.
- Storage of user information 'everywhere and nowhere.' This means maximized decentralization and fragmentation between the executing nodes of user data and the algorithms executed.
- Abstraction of users from metadata, technical information, topology or the network operation processes.
- Concealment of metadata and technical information from users.

TN TheOoL software is the same, and its meaning at a particular point of time is automatically defined by PSDE algorithms depending on requirements of the network. TN can simultaneously play several roles (Table 1).

TN TheOoL software is the same, and its meaning at a particular point of time is automatically defined by PSDE algorithms depending on requirements of the network. TN can simultaneously play several roles. TN can simultaneously play

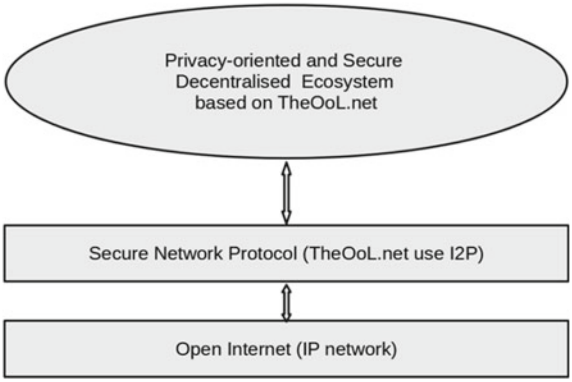



**Fig. 1** The architecture of private and secure decentralized internet

**Table 1** Roles of the nodes

| # | Role | Function |
|---|---|---|
| 1 | AN (abonent node) | Designed for direct interaction of a user and PSDE |
| 2 | TN (technical nodes) | |
| 2.1 | SCN (storage & computational node) | Direct execution of distributed computational problems and distributed data storage |
| 2.2 | MN (metadata node) | Working with metadata of a distributed database stored and processed in SCN network |
| 2.3 | DLCN (distributed ledger consensus node) | Nodes executing algorithms of the distributed ledger consensus [38–40] |

several roles. The architecture of PSDE node software includes four (4) layers (Fig. 2). The components are located in nodes of the system in accordance with the location map of software components of PSDE four-layer architectures (Table 2). Software components are distributed as operation of PSDE user functions and as a possibility to control and manage TN hardware on the side of their owners.

That said, execution of the main principles of PSDE development is ensured, namely complete concealment of metadata and topology of data storage, and fully automated maintenance of PSDE software and the data serviced.

**Table 2** Location map of the software architecture components on PSDE nodes

| Layer No. | Architecture layer | Component ID | Component | Abonent node | Technical node | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | SCN | MN | DLCN SSN | |
| 1 | User application management layer | 1.1 | User workflows controller (UWFC) | Y | SA | SA | SA | SA |
| 2 | Local node management layer | 2.1 | User data controller (UDC) | Y | SA | SA | SA | SA |
| | | 2.2 | User computations controller (UCC) | Y | N | N | N | SA |
| 3 | Network data access layer | 3.1 | Serverless security engine (SSE) | N | N | Y | Y | N |
| | | 3.2 | Serverless data search engine (SDSE) | N | N | Y | N | Y |
| | | 3.3 | Cloud computations engine (CCE) | N | Y | Y | Y | N |
| 4 | Network data processing layer | 4.1 | Serverless network autoconfiguration & automanagement engine (SNA&AE) | N | LM | Y | Y | LM |
| | | 4.2 | Data transmission & storage protocol (DT&SP) | N | Y | Y | Y | Y |
| | | 4.3 | Serverless cloud computations protocol (SCCP) | N | Y | Y | Y | N |

*Y* Available, *N* Not available, *SA* Specialized application, *LM* Local module

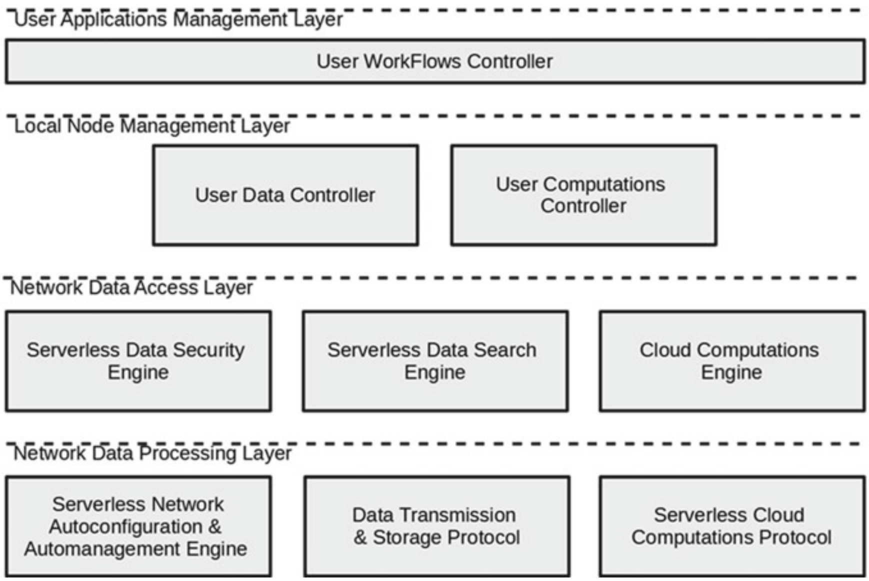



**Fig. 2** PSDE software architecture

The engine of PSDE serverless security and management system is TheOoL distributed ledger (blockchain). It contains two interrelated distributed database: Data base security descriptor for the Serverless Data Security Engine (Fig. 2) that ensures control over data access; PS database and database for smart contracts—scripts which manage user interaction with PSDE. Every task PSDE receives from users is a smart contract (special script) that defines roles and rewards of specific PSDE nodes in the scope of that task. In accordance with this script a major user task is divided into a set of minor process tasks (TT ("Tech Task")) which are assigned a node executor and rewards for execution of each TT. In the real world, rewards are given to owners of the corresponding hardware nodes that have no access to the execution process, as well as any information about TT.

MN network services the distributed ledger database. This network not only contains blocks and smart contract scripts but is also responsible for managing data ad dresses and their conversion from address space of a certain subsystem into address space of a different subsystem. In fact, data addresses in the subsystems of 4th layer architecture differ from their presentation in the subsystem of 3rd layer architecture. This is done in order to enhance implementation of the sixth and seventh principles of PSDE. Here MN network plays the role of a secure transcriber that ensures maximum abstraction and concealment of the topology of data storage and processing from any network member.

To some extent MN represents points of vulnerability from the standpoint of the possibility of breaching the aforesaid principles. This is why assignment to any TN the function of MN takes place in a random manner; MN infinitely migrates into TN networks; data on MN are encrypted, while the very address mappings are stored on different MN, i.e. SDSE encrypted link and its matching DT&SP encrypted link are stored on different MN, and these MN are, with high probability, owned by different persons or entities in the real world.

DLCN network has been set up in a comparable way and for the same reason; in this network algorithms of the consensus responsible for ensuring legitimacy of data alteration in distributed ledgers, are executed. In order to apply and alter security descriptors we use secure and reliable Proof Of Work (PoW) consensus algorithm [38] which is modified to as to eliminate competitive growth of computational complexity, which is typical of other PoW networks and which eliminates excessive power consumption as required to support the consensus in TheOoL network. One of the implementations of Proof of Stake (PoS) algorithm [39, 40] was selected as the consensus algorithm for the distributed ledger of smart contracts in payments. However, it should not be understood that two separate consensus algorithms are used in TheOoL network. In fact, we use a hybrid of PoW and PoS, which operates as an inseparable mechanism ensuring double validation of the most important entries in the distributed ledger, namely entries about payments and changes in the security descriptors. Technical metadata (entries about physical location of parts of data or task executor nodes as part of distributed computing) are serviced by PoW only. Double validation is re quired in order to significantly reduce the probability of success of distributed ledger attacks [41]. Encryption and computer generated signatures in TheOoL distributed ledger are executed using the well-proven (for instance, in Bitcoin network cryptography [38]) elliptic curves (ECC) [42, 43] featuring SECP521R1 elliptic curve.

## 3 The Serverless Videoconferencing System

In order to implement SoPVAC (Fig. 3) we use the components 1.1 and 2.2 in AN and the components 3.1, 3.3 and 4.3 in TN network. In AN (UI) SoPVAC user interface is executed and interacted with a group of virtual servers—SoPVAC routers, which are automatically created and executed in TN PSDE network for each specific videoconference within the scope of one specific smart contract. TN network can be represented as a multitude of nodes:

$$S_t \equiv S_m \cup S_d \cup S_s \cup S_{ss} \equiv \left\{X_1 \ldots X_i \ldots X_\xi\right\}, X_i = \{a, b, c\}, \quad i = \overline{1 \ldots \xi} \quad (1)$$

where $S_m$—MN multitude, $S_d$—DLCN multitude, $S_s$—SCN multitude, $S_{ss}$—SSN multitude, $\xi$—total number of TN in PSDE, $X_i$—multitude of parameters of a specific TN [37] which are essential for the operation of SoPVAC and which includes $a$—number of white spaces in the queue, $b$—node performance index, and $c$—available space in disk subsystem of the node.

Decentralized applications or DApps responsible for operation of the search system ("SDSE") and the data storage system ("DT&SP") ensure uploading on AN

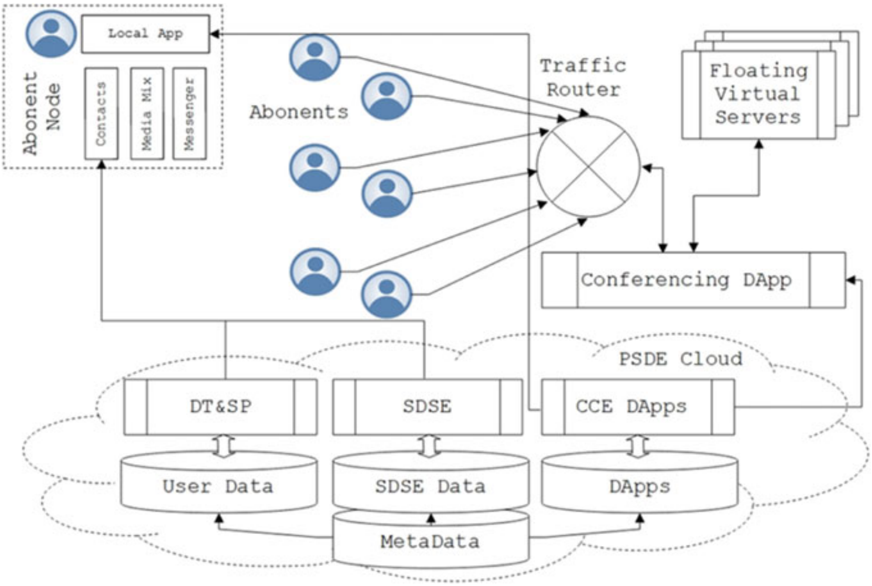



**Fig. 3** SoPVAC architecture

the list of contacts and current version of the local application SoPVAC ("AN UI") which interacts via UWFC (Table 2) with the videoconference virtual router and SoPVAC DApp ("CDApp").

In order to support SoPVAC 2 indexes are maintained for each user—the subscriber list in SDSE: $A$—the list of abstract subscribers (AA) registered as SoPVAC users; $A_o$—the list of AA with the status 'online' and 'available.' Let's define the sets $A$ and $A_o$:

$$A = \{a_1 \ldots .a_n\}, A_o = \{a_1 \ldots a_k\}, a_i = \left\{id_i, t_i, Q_i^s, t_i^{max}\right\} \quad (2)$$

where $a_i$—subscriber in the list attributed by lifetime $t_i$, $t_i^{max}$—maximum storage time, steady-state public key ECC $Q_i^s$ and the unique identifier $id_i$, and the range of $i$ index change is defined by appurtenance to one or another list:

$$i = \overline{1 \ldots n} \big| a_i \in A, i = \overline{1 \ldots k} \big| a_i \in A_0$$

AN UI node contains a multitude of subscribers registered in the node, which we define as:

$$U = \{u_1 \ldots u_r\}, u_j = \left\{id_j, d_j^s, Q_j^s\right\}, Q_j^s = d_j^s \cdot G, j = \overline{1 \ldots r} \quad (3)$$

where $id_j$—unique identifier of a subscriber, $d_j^s$—steady-state private key of a subscriber, $Q_j^s$—steady-state public key of a subscriber, $G$—anchor point of the elliptic curve which is common for PSDE network, $r$—number of subscribers serviced by specific AN UI.

Only SDSE and AN UI are responsible for managing the first list. For each subscriber the following functionality is executed:

$$\left\{t_j = 0 | id_j \in A\right\}, \left\{f : \mathrm{U} \to \left\{a = \left\{id_j, t_j = 0, Q_j^s, t_j^{max}\right\} \in A\right\} | id_j \notin A\right\}, j = \overline{1 \ldots r}$$

in the multitudes (2, 3). Put it in other way, when AN UI is launched the node checks availability of the entries about $u_j$ in the multitude $A$ registered in the node. If subscribers are registered AN UI messages SDSE to clear time taking since the last connection. Otherwise the procedure of registration of new AA is launched in SDSE indexes. During registration a user sets the maximum value for connection timeout. This parameter defines the period, during which SDSE stores information about the registered AA and their related data in the cloud storage. When this value is reached all the information is automatically deleted, and subsequent AA connections are executed as new ones.

Since the second list requires consistent monitoring of subscriber activity it is manager—in addition to SDSE and AN UI—by computation processes launched within CDApp. AN UI activated by a user gets connected with CDApp modules on MN known to it, then it updates the list of available MN and notifies CDApp module in MN about conversion of AA registered in it into the status ‘online’ and ‘available’ specifying the user groups that will have access to this information. CDApp synchronizes information about subscriber activation via SDSE indexes which are publicly accessible in the corresponding groups or in SDSE public index, if a subscriber needs to interact with a wide range of counterparties unknown to the user. An entry in SDSE remains active for 2 min, after which SDSE automatically deletes the entry if its lifetime is not extended, while the status of abstract subscribers registered in the corresponding AN UI is converted to ‘offline’ status. After that CDApp assigns two monitoring MN (CMN) from (1) to the node with AN UI, the task of which is to extend lifetime of the subscriber’s entry in SDSE:

$$f : S_m \to \left\{\{CMN_1, CMN_2\} | \max_{a,b} S_m\right\}$$

The entry about CMN assigned to the subscriber is also entered in SDSE. This is necessary as it is CMN that ensures communications of the subscriber with other SoPVAC network members. For each n subscriber in the multitudes (2, 3) the following functionality is executed:

$$f : U \to \left\{a = \left\{id_j, t_j = 0, Q_j^s, t_j^{max} = 2, \{CMN_1, CMN_2\}\right\} \in A_o\right\} | id_j \notin A_o, \\ j = \overline{1 \ldots k}$$

CMN obtain the designation of the subscriber once per 1 min and, if connection is established with the corresponding AN UI, it extends lifetime of the entry about activity of abstract subscribers in SDSE indexes. Moreover, CMN under no circum stances receives or processes audio and video streams or any text messages of the messenger. CMN objectives include setting initial connection; call initialization; and initialization of the videoconference virtual routers.

Even though CMN, out of necessity, has information about a client identifier on the transport layer of ISO/OSI model it does not provide for any possibility of user compromise and deanonymization because TheOoL network uses I2P as its transport layer, since CMN in this case does not know the client's I2P ID but I2P ID tunnel router [24] which is the closest to CMN. This router is generally an I2P network node which, firstly, is not necessarily a TheOoL node and, secondly, is automatically replaced once every 10 min (a default value). Therefore CMN remain unchanged for a subscriber as long as there are active, except when CMN operation is terminated. The following two scenarios are possible in this case:

1. At least one (1) CMN is available. In this case PSDE automatically allocates a new CMN to replace the one that failed.
2. All the CMN are switched off. In this case an AN UI client initiates the algorithm of initial connection, receives the new CMN pair, and generates the new entry about availability.

In order to maintain SoPVAC availability one must only have one CMN, an additional one is introduced into the system in order to enhance its fault tolerance. Once AN UI is launched and CMN is allocated a user is ready to send text messages to users via the local messenger in p2p mode from the list permitted to them in SDSE index in the status 'available' and the user is able to make voice and video calls to them. That being said, no call or conversation data can be saved, including locally, and there is also no possibility for delayed communication, for example, for a user to send text messages in 'offline' mode.

Implementation of [44] Real-time Transport Protocol (RTP) [45] in p2p mode is responsible for processing multimedia in SoPVAC. This protocol is widely used for the transmission of multimedia files in 'online' mode; it is well documented and has no essential features for the application in SoPVAC. By comparison, the procedure of videoconference initialization, data transfer, and data processing should be described in greater detail.

In order to start a videoconference a subscriber that initiates the videoconference selects other participants of the videoconference from the list of available subscribers and sends them requests for connection ("the handshaking request"). SoPVAC exe cutes the procedure of checking availability of subscribers via CMN allocated to them. After that subscribers with validated active status receive the handshaking re quest. If they accept the request the procedure is executed for the generation and transfer of the relevant session encryption keys using point-to-point protocol based on Diffie-Hellman algorithm [43] (the handshake). ECC [42] protocol is used as an asymmetric protocol along with the generation of ephemeral key pairs [43] to compute secrets and steady-state key pairs as registered for AA in SDSE to ensure operation of the digital signature algorithm. And for traffic encryption, with the help of the relevant session encryption key, we apply the Advanced Encryption Standard

(AES) [46]. We selected this standard due to its high cryptographic robustness and, at the same time, low computational complexity of the algorithm, which is especially important when processing multimedia streams in online mode. In fact, this means AES operates fast enough so as to exclude any significant delays in the transmission of audio and video data, thanks to the execution of encryption/decryption procedures. Still, at the time of writing this article there is only one known algorithm for the implementation of AES attack—the so-called algebraic attack described by Niels Ferguson [47] in 2001. This attack, however, has never been actually performed. And its invalidity was proven at a later stage [48]. Other attacks are related with specific implementations of AES. They are complex for the most part and hardly realized into practice [46].

A tuple of 10 session keys is generated during the handshake for each participant of a videoconference; these keys are used one-at-a-time during the communication session for AES encryption of multimedia traffic and text messages sent by the messenger. The key tuples of each participant, encrypted with the ECC common secret, are provided to the remaining participants of the videoconference.

AN UI of the videoconference host randomly defines the position number of the encryption key in the tuple, which is common for all the participants and which is provided to all the participants. In the course of 5 min the participants use the key to encrypt their outbound traffic, which key is located in the position computed this way in the key tuple generated by them. Upon expiry of this time AN UI of the videoconference host generates and transmits the new position number and provides the same to other participants of the videoconference. Thus, what is transmitted in The OoL network is data streams encrypted with keys which are different for each videoconference participant, and such keys are randomly replaced every 5 min. AN UI of the videoconference host randomly defines the position number of the encryption key in the tuple, which is common for all the participants and which is provided to all the participants. In the course of 5 min the participants use the key to encrypt their outbound traffic, which key is located in the position computed this way in the key tuple generated by them. Upon expiry of this time AN UI of the videoconference host generates and transmits the new position number and provides the same to other participants of the videoconference. Thus, what is transmitted in TheOoL network is data streams encrypted with keys which are different for each videoconference participant, and such keys are randomly replaced every 5 min. In case there are two (2) subscribers in a videoconference SoPVAC creates no entities directly related with the videoconference. Data transmission is executed p2p between nodes of the conversation participants who keep two (2) tuples of AES keys and maintain open one inbound encrypted data stream and one outbound encrypted data stream via I2P network. It is fundamentally important that these are different streams transmitted via different I2P tunnels. Where there are n > 2 participants of a videoconference SoPVAC provides for 3 op erating modes. In the first mode (Fig. 4a) the host and all the participants set up direct p2p connections between each other. This option for a multiple user videoconference requires that the network connection bandwidth of each AN UI equals 2n (n inbound and outbound ones) and this implies additional load on the subscriber nodes, which in this scheme independently send out traffic to the participant networks. This means high network load, I2P usage charges (for security of the transport layer), which shifts the function of a multicast router to the client node. The key advantages of this scheme include its low cost as no additional resources are used in PSDE cloud, and enhanced security as there are no gate-keepers in the course of data transmission and processing.

In the first mode, SoPVAC may be quite efficient, even in case of mediocre network connection where users resort to a voice conference. Broadband network connection is required for efficient video stream processing. To restrict traffic SoPVAC provides for operation using the videoconference virtual routers (VRC).

In the second mode, AN UI host addresses CDApp via CMN, and in SCN network (Table 1) the videoconference virtual router is created (VRC). Its main function is to harvest data streams from videoconference participants and to redirect them to receivers. In this mode VCR does not have encryption keys; it merely receives encrypted data streams from the videoconference participants and transmits them as intended. Thanks to VCR the number of simultaneously open multimedia streams in the sub scriber node is reduced to n + 1 (n inbound and 1 outbound one), in other words, in this mode VCR is a multicast router. Thus, requirements to the speed of AN UI network connection is cut in half, due to insignificant reduction in the level of conversation privacy. Here we have an intermediary node which executes the VCR process, and the entire videoconference traffic goes via this node. A potentially compromised owner of the intermediary is able to intercept the traffic; however, the owner is not able to read the traffic because the owner lacks the relevant encryption keys.

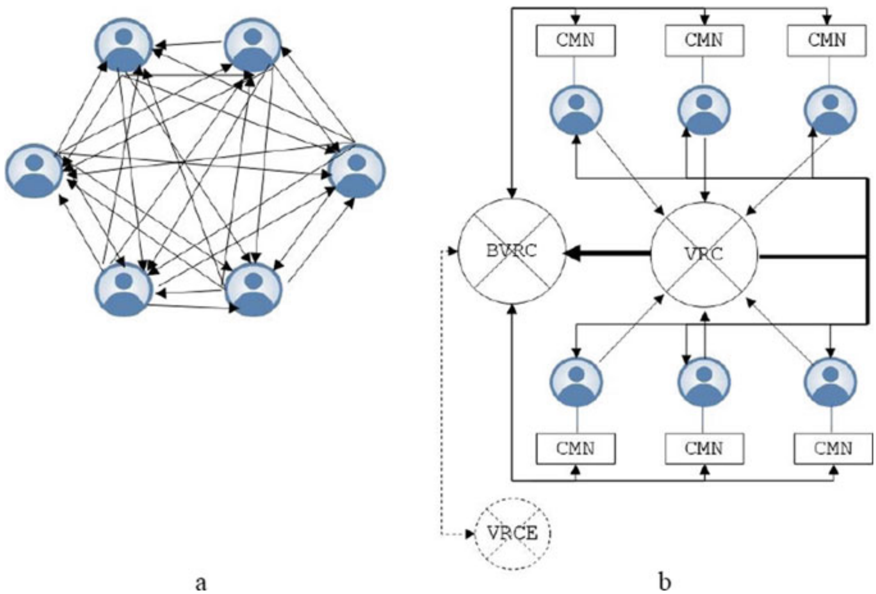



**Fig. 4** SoPVAC multiple user conferences: **a** simple p2p; **b** with VRC

And, finally, the third operating mode implies that VCR, in addition to its routing function, processes the inbound multimedia traffic for participants of the videoconference. This means VCR gets the encryption keys on par with the remaining participants of the videoconference. The ECC key pair as used by VCR to encrypt AES keys is the SCN owner's pair, and this is where VCR software process is launched. The process of AES key dissemination is altered several times. VCR generates only one (1) AES key. Each client transmits only one (1) (active) key from the VCR-generated tuple. Tuples with AES keys are not transmitted between the clients.

VCR harvests multimedia streams from the videoconference participants, generates a single stream made of them, encrypts the resulting stream using its key, and then shares it with all the videoconference participants. Thus, VCR gets full access to the conversation without reducing the load on AN UI to two (2) multimedia streams, thereby reducing requirements to the network connection and hardware of the subscriber node, thanks to the steep decrease in the level of conversation privacy.

Yet another downside of VCR operating modes is the overall deterioration of fault tolerance of SoPVAC videoconferences, since VCR outage automatically terminates a videoconference for all its clients. To tackle this problem, as well as to mitigate the risk of conversation interception relating to each videoconference, three (3) VCRs are simultaneously assigned in SoPVAC (primary, backup, and standby one). They operate in the network of 'floating virtual servers' ("FVS"). In total, PSDE assigns from (1) a list of 10 SCN for VCR operation, the already familiar FVS $f : S_s \rightarrow \left\{ V = \{v_1, \ldots, v_{10}\} | \max_{a,b} S_s \right\}$. Initially two servers are randomly selected among them $random(V) = \{v_1^r, v_2^r\}$. The primary VCR is elevated to $v_1^r$, which gets the first keys and starts to service the videoconference. The backup VCR is elevated to $v_2^r$, which gets the first keys and—5 min later—the second encryption keys for the videoconference. $v_2^r$ does not service the videoconference during the first 5 min; it receives all the videoconference information from $v_1^r$. AN UI maintain 'light contact' with the backup VCR via its CMN, and in case the primary VCR is disconnected they immediately switch to it. CDApp begins to monitor availability of the remaining FVS. These eight (8) SCN generate the third $V_3^r = V \backslash \{v_1^r, v_2^r\}$, which replaces the distributed VCR. VCR passive processes are executed in $V_3^r$ nodes; they get CMN list of the videoconference participants and immediately go to standby mode. In case one of the FVS gets disconnected CDApp requests PSDE cloud for a new one.

Upon expiry of 5 min FVS begin to 'float.' The backup VCR becomes the primary one and begins to service the videoconference with the new encryption keys. The new backup VCR is allocated from FVS, while the new SCN is given to the standby VCR. After that, the procedure of server alteration and encryption key alteration is repeated every 5 min.

Thus, a compromised SCN owner is able to intercept no more than 5 min of a videoconference. In case of the second operating mode an encrypted traffic may be intercepted, and the number of data packets in that traffic is definitely insufficient for decryption [46]. In case of the third operating mode, the time period for altering the keys and servers may be changed downwards as required by the videoconference host, which undoubtedly further cuts any loss of the potential interception, yet it substantially increases the cost of the videoconference. Therefore in the third operating mode it is not appropriate to discuss any vital information. The third operating mode was introduced in SoPVAC only to enable mass communication events, for instance, open access lectures, seminars or webinars. From the security standpoint, the only advantage of using SoPVAC in the third operating mode is total anonymity of all the conversation participants, provided, of course that faces or other identifies of such participants are not caught on camera.

SoPVAC does not record any information about a videoconference or any information about holding of the same, and does not allow any of its participants to record any such information, except for transactions of payment for the resources used in TheOoL distributed ledger. Once a videoconference is over all the SCN executed software processes are terminated without logging, immediately after payment (which is a fully automated process).

## 4 SoPVAC Text Messaging Subsystem

In certain cases it may be necessary to send a message to a user who is currently offline. For this scenario SoPVAC has a built-in text messaging subsystem. In order to transmit messages to an offline user or to record a message for subsequent reading (Fig. 5) the messenger's repository is used in the distributed personal data storage, access to which is executed via DT&SP.

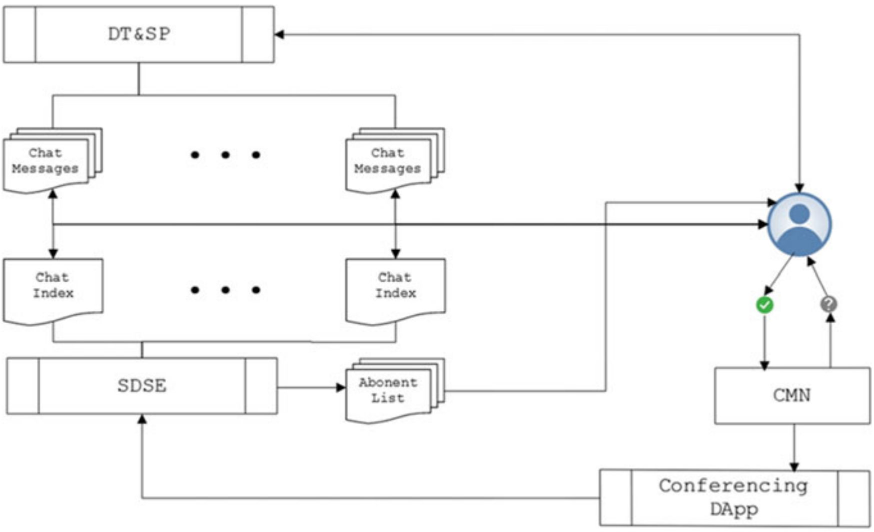



**Fig. 5** SoPVAC text messaging subsystem

In this case it matters who sends the message for storage—the sender or the receiver—since it defines who sets the parameters of storage of messages and pays fees for the storage, as well as who owns such messages in the data repository, which determines the right to dispose of information, for instance, provision of third party access to such information. The key parameters of storage include the storage time, required speed of data access, redundancy coefficient (it defines the number of backup copies, which is two (2) by de fault), and the minimum number of parallel reading threads, which depends on the number of potential readers and which is defined by the owner of information. The owner sets the storage parameters for all messages of the messenger (applicable to all messages by default), and certain rules may be established for a particular conversation. By default, the sender is the message owner in SoPVAC; however, there is an option where message ownership rights may be transferred to another participant of the conversation, if the sender is not willing to pay for maintenance of the conversation in SoPVAC repository. To activate this mode, however, a prior consent of the conversation participants is required.

The method of parametric synthesis of a PID controller of a control system with feedback by high-order objects (above second) taking into account the placement of complex conjugate poles is considered. The method is implemented in the MathCAD software environment. Working capacity of the method showed good results corresponding to the design task on the example of the ACS model by the fourth order object. Transient process was without excessive correction.

Special repository is created for each conversation in DT&SP and its matching index in SDSE. It stores metadata of the messages in order to retrieve them from DT&SP, as well as a link to security descriptors of the group that partakes in the conversation. Thus, any member of the group is able to place their message in the conversation, regardless of availability of the remaining participants. When launched the subscriber AN UI scans indexes of the groups connected to it, receives and reads the messages. Index scanning is performed with the user-defined frequency throughout the entire uptime of AN UI.

AES encryption ensures protection of messages from unauthorized reading. AES key transfer is executed based on Diffie-Hellman protocol using elliptic curves [43]. Message encryption and transmission is performed in the following order:

1. AES random key is generated in the message encrypted.

   Ephemeral pair of ECC keys is generated.

   Table of secrets is put together for transmission of the key from Item 1 to the conversation participants. The following is performed for each conversation participant:

   (a) We retrieve ECC steady-state public key from SoPVAC subscriber list.
   (b) We calculate the common secret for the private key of the ephemeral pair (Item 2) and the steady-state public key of the receiver (Item 3.a).
   (c) We encrypt AES key using the protocol derived in Item 1. The common secret is used as the encryption key (Item 3.b).

(d) We place the encrypted key and the steady-state public key of the receiver into the table.

We place the encrypted message, the ephemeral pair public key, and the table of secrets into the corresponding repository.

In order to read the message the receiver must select in the table of secrets the en try containing their steady-state public key and AES corresponding to their encrypted key. The receiver must then compute the common secret by applying the ephemeral pair public key, and the private key of their steady-state pairs of ECC key, and then decrypt AES key. Then the receiver can decrypt the message.

## 5 Conclusion

In summary, the proposed private communications system ensures high level of privacy due to the decentralized architecture and elimination of the server infrastructure owners from SoPVAC, and thereby elimination of the potential threat of privacy breach by such owners due to their unreliability or errors. Even in the operating modes using temporary centralized resources (VCR) SoPVAC does not allow their owners to reliably intercept conversations, because none of the owners of PSDE technical nodes knows and may know which of the nodes and when that node receives VCR software process for execution. What prevents this is, on the one hand, the random principle of selecting nodes for VFS and their perpetual random alteration and, on the other hand, and even to a greater extent the very organizational nature of The OoL private internet, which takes the form of a DAO. Put it in other way, it is an open peer-to-peer organization where interaction is governed by TheOoL algorithms. Any participants are free to join this organization at any point of time. Both subscribers and owners of the technical nodes automatically become participants of the computing resource market in TheOoL network and they are financially incentivized to offer hardware servers to PSDE cloud. These (let's call them vendors) network members can be of any geographic setting, nationality, and beliefs. Their nodes may be selected to perform any tasks within PSDE on equal terms with hardware servers of potential attackers. Actually, the more vendors operate in TheOoL network the lower the chance for any of them to intercept a specific conversation.

And even if they succeed in intercepting part of a conversation the attackers must, firstly, break through cryptographic data protection and, secondly, the attackers must be able to identify participants of the conversation, which is effectively prevented by user abstraction from PSDE technical infrastructure and user anonymization. This substantially complicates (makes it almost impossible) identification of any intercepted traffic and setting that traffic against specific TheOoL network members in the real world.

## References


1. Silaeva, V.L: Substitution of reality as a socio-cultural mechanism of societal virtualization: Dissertation for Cand. Philosophy Sciences Degree: 09.00.11: p. 115. Moscow 2004. Russian State Library OD, 61:04-9/733; Ziegler, J.G., Nichols, N.B.: Optimum settings for automatic controllers. Trans. Am. Soc. Mech. Eng. **64**, 759–768 (1942)
2. Control Engineering. https://controlengrussia.com//rynok//obshhestvo-5-0-po-yaponski. Last accessed 12 May 2022
3. Nenashev, A. V., Khryashchev, V.V.: The economics of introducing the peer-to-peer system of storage and processing of protected information at an enterprise. In: 2019 XXI International Conference Complex Systems: Control and Modeling Problems (CSCMP), pp. 769–772, Samara, Russia. https://doi.org/10.1109/CSCMP45713.2019.8976720
4. Vassilieva, E.V: "Tele-everything" world and hybrid office model—new global trends after the COVID-19 pandemic. UPRAVLENIE/MANAGEMENT (Russia) 9(3):125–136. (In Russ.). https://doi.org/10.26425/2309-3633-2021-9-3-125-136. Last accessed 12 May 2022
5. Gartner Forecasts Worldwide Security and Risk Management Spending to Exceed $150 Billion in 2021. Gartner. STAMFORD. https://www.gartner.com/en/newsroom/press-releases/2021-05-17-gartner-forecasts-worldwide-security-and-risk-managem. Last accessed 17 May 2021
6. Distributed Systems: Principles and Paradigms / E. Tanenbaum, Maarten van Steen, pp. 877: St. Petersburg: Piter 2003. Fig.—(Computer Science Classics series)
7. Olifer, V., Olifer, N.: Computer Networks: Principles, Technologies, and Protocols: College Textbook 5th edn., p. 992. St. Petersburg: Piter 2016—Fig.—(College Textbook series)
8. The Media Freedom Internet Cookbook. Edited by Christian Möller and Arnaud Amouroux: Organization for Security and Co-operation in Europe (OSCE) 2004, p. 274. Vienna.
9. F. Gont ICMP attacks against TCP (draft-gont-tcpm-icmp-attacks-01). The Internet Society. 2004. https://tools.ietf.org/id/draft-gont-tcpm-icmp-attacks-01.txt. Last accessed 12 May 2022
10. INTERNET STANDARD RFC 792 Internet Control Message Protocol (1981). https://tools.ietf.org/html/rfc792. Last accessed 12 May 2022
11. INTERNET STANDARD RFC 4443 Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification (2006). https://tools.ietf.org/html/rfc4443. Last accessed 12 May 2022
12. PROPOSED STANDARD RFC 3376 Internet Group Management Protocol, Version 3 (2002). https://tools.ietf.org/html/rfc3376. Last accessed 12 May 2022
13. PROPOSED STANDARD RFC 3810 Multicast Listener Discovery Version 2 (MLDv2) for IPv6 (2004). https://tools.ietf.org/html/rfc3810. Last accessed 12 May 2022
14. . B.Yu: Anin. Computer Information Protection, p. 384. St. Petersburg: BKhV-Peterburg 2000. Fig.
15. Baronov, V.V., Kalyanov, G.N., Popov, Yu.N., Titovsky, I.N.: Information Technologies and Enterprise Management, p. 328. IT, Moscow (2009)
16. Khryashchev, V.V., Nenashev, A.V.: Efficiency of the implementation of the peer-to-peer distributed system for securable information storage and processing (TheOoL Project). Model. Syst. Process **3**, 82–89. https://doi.org/10.12737/2219-0767-2021-14-3-82-89. Last accessed 12 May 2022
17. Romanets, Yu.V., Timofeev, P.A., Shangin, V.F.: Information Protection in Computer Systems and Computer Networks (Shangin, F., ed.), 2nd revised and enlarged edition, p. 376. Radio and Communications, Moscow (2001)
18. Information Protection in Computer Systems / eds by Doctor of Economics E.V. Stelmashonok, Cand. Sc. Physics and Mathematics I.N. Vasilyev, p. 163. St. Petersburg: Publishing House of St. Petersburg State University of Economics (2017)
19. List of Data Breaches. Wikipedia. https://en.wikipedia.org/wiki/List_of_data_breaches. Last accessed 12 May 2022
20. High-profile leaks of user data in January 2019—April 2019. Cloud4Y company blog. https://habr.com/ru/company/cloud4y/blog/452880/. Last accessed 12 May 2022

21. Blagoveshchensky, A.: Yandex explained the Russians' passport- and ticket data leak. Rossiyskaya Gazeta (2018). https://rg.ru/2018/07/17/iandeks-obiasnil-utechku-pasportov-i-biletov-rossiian.html. Last accessed 12 May 2022
22. Google reported a data breach that affected more than 52 million users due to an incident. RBK 2018. https://www.rbc.ru/technology_and_media/11/12/2018/5c0f033a9a794766b04420a5. Last accessed 12 May 2022
23. Web 2.0. Material taken from Wikipedia—The Free Encyclopedia. https://ru.wikipedia.org/wiki/%D0%92%D0%B5%D0%B1_2.0. Accessed on 31 March 2022
24. I2P Project Documentation: https://geti2p.net/en/docs. Last accessed 12 May 2022
25. Egger, C., Schlumberger, J., Kruegel, C., Vigna, G.: Practical attacks against the I2P network. In: Stolfo, S.J., Stavrou, A., Wright, C.V. (eds.) Research in Attacks, Intrusions, and Defenses. RAID 2013. Lecture Notes in Computer Science, vol. 8145. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41284-4_22. Last accessed 12 May 2022
26. TOR Project Documentation: https://2019.www.torproject.org/docs/documentation.html.en. Last accessed 12 May 2022
27. Abbott, T.G., Lai, K.J., Lieberman, M.R., Price, E.C.: Browser-based attacks on tor. In: PET'07: Proceedings of the 7th International Conference on Privacy Enhancing Technologies, pp. 184–199. Springer, Berlin, Heidelberg (2007)
28. Bauer, K., McCoy, D., Grunwald, D., Kohno, T., Sicker, D.: Low-resource routing attacks against tor. In: WPES '07: Proceedings of the 2007 ACM Workshop on Privacy in Electronic Society, pp. 11–20. ACM, New York, NY, USA (2007)
29. Casenove, M., Miraglia, A.: Botnet over tor: the illusion of hiding. In: 2014 6th International Conference on Cyber Conflict (CyCon 2014), pp. 273–282 (2014)
30. Link to the repository The Tor Path Simulator: https://github.com/torps/torps. Last accessed 12 May 2022
31. Link to the repository Stinger-tor: https://github.com/whitepacket/Stinger-Tor. Last accessed 12 May 2022
32. Security of Videoconferencing Services. Trend Micro company blog. https://habr.com/ru/company/trendmicro/blog/510300/. Last accessed 12 May 2022
33. Web 3. Material taken from Wikipedia—The Free Encyclopedia. https://ru.wikipedia.org/wiki/Web3. Accessed on 31 March 2022
34. Web 3.0. Material taken from Wikipedia—The Free Encyclopedia. https://ru.wikipedia.org/wiki/Web_3.0. Accessed on 31 March 2022
35. TheOoL White Paper. TheOoL, Inc. https://theool.net/doc/TheOoLWhitePaper.pdf. Last accessed 12 May 2022
36. Michael, S.: Some Reflections on the Nature of Decentralized (Autonomous) Organizations (September 1, 2021). King's College London Law School Research Paper Forthcoming, Available at SSRN: https://ssrn.com/abstract=3915843 or https://doi.org/10.2139/ssrn.3915843. Last accessed 12 May 2022
37. Nenashev, A.V., Tolstenko, A.Yu., Oleshko, R.S.: Model of the peer-to-peer distributed system for securable information storage and processing without traffic prioritization (TheOoL project). In: 2021 Proceedings of the III International Workshop on Modeling, Information Processing and Computing (MIP: Computing-2021), Krasnoyarsk, Russia. https://doi.org/10.47813/dnit-mip3/2021-2899-141-150
38. Satoshi Nakamoto: Bitcoin: A peer-to-peer electronic cash system. http://bitcoin.org/bitcoin.pdf. Last accessed 12 May 2022
39. Li, W., Andreina, S., Bohli, J.M., Karame, G.: Securing Proof-of-Stake blockchain protocols. In: Garcia-Alfaro, J., Navarro-Arribas, G., Hartenstein, H., Herrera-Joancomartí, J. (eds.) Data Privacy Management, Cryptocurrencies and Blockchain Technology. DPM 2017, CBT 2017. Lecture Notes in Computer Science, vol. 10436. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67816-0_17. Last accessed 12 May 2022
40. Saleh, F.: Blockchain without waste: proof-of-stake. Rev. Financial Stud. **34**(3), 1156–1190 (2021). https://doi.org/10.1093/rfs/hhaa075. Last accessed 12 May 2022

41. Sayeed, S., Marco-Gisbert, H.: Assessing blockchain consensus and security mechanisms against the 51% attack. Appl. Sci. **9**, 1788 (2019). https://doi.org/10.3390/app9091788. Last accessed 12 May 2022
42. SEC2: Recommended Elliptic Curve Domain Parameters. Standards for Efficient Cryptography (SEC) Version 1.0. Certicom Corp (2000). https://www.secg.org/SEC2-Ver-1.0.pdf. Last accessed 12 May 2022
43. Schneier, B.: Applied Cryptography: Protocols, Algorithms and Source Code in C, 20th Anniversary edn., p. 784. Wiley, New Jersey (2015)
44. Gstreamer: open source multimedia framework. https://gstreamer.freedesktop.org/. Last accessed 12 May 2022
45. Perkins, C.: RTP: Audio and Video for the Internet, p. 414. Addison-Wesley (2003)
46. Barichev, S.G., Goncharov, V.V., Serov, R.E.: 2.4.2: AES Standard. Rijdael Algorithm. In: The Essentials of Modern-Day Cryptography, 3rd edn., p. 176.: Dialot-MIFI, Moscow (2011)
47. Ferguson, N., Schroeppel, R., Whiting, D.: A simple algebraic representation of Rijndael. In: Selected Areas in Cryptography. Proceedings of SAC 2001, Lecture Notes in Computer Science #2259, pp. 103–111. Springer (2001)
48. Lim, C.-w., Khoo, K.: An analysis of XSL applied to BES (English). In: Fast Software Encryption, vol. 4593, pp. 242–253. Springer, Berlin/Heidelberg. https://doi.org/10.1007/978-3-540-74619-5_16

# Formation of an Optimal Set of Protective Measures to Handle the Information Security Risks of Industrial Control Systems of Gas Production Enterprises

**Alexander Bolshakov, Darya Fomina, and Andrey Rimsha**

**Abstract** This article describes the elements of the generalized structure of a modern automated process control system and the levels of process control. An algorithm has been developed for assessing the risks of information security of automated process control systems and the formation of an optimal set of protective measures for their processing using a metaheuristic approach. The basic functions of the genetic algorithm were given. The necessity of the process of searching for the optimal configuration of protective measures is shown, which will improve the efficiency of information security of automated process control systems under destructive influences. A formalized statement of the optimization problem is given. Testing of the proposed software for the gas producing enterprise Achimgaz has been carried out. The efficiency of the optimization algorithm and the complexity of programs that implements it has been confirmed. The preferable regions for the parameters of the genetic algorithm were determined: the probabilities of crossing over and mutation, the type of crossing-over, the volume of the initial population.

**Keywords** Automated control system · Information security · Information security tools · Gas production enterprise · Risk assessment · Risk treatment measures · Genetic algorithm

## 1 Introduction

Nowadays information systems are essential part of organizations. These organizations will also lose many competitive advantages by losing security. There are many research papers and standards in the field of security risk management (ISRM), including NIST 800–30 and ISO/IEC 27005. However, few research papers are

---

A. Bolshakov (✉) · D. Fomina · A. Rimsha
Peter the Great St. Petersburg Polytechnic University, 29, Polytechnicheskaya Street, St. Petersburg 195251, Russia
e-mail: aabolshakov57@gmail.com



A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_23

devoted to information security risk mitigation. Although the standards explain the general principles and recommendations, they do not provide any details about the implementation of ISRM. Handling information security risks under financial constraints requires a lot of effort.

Automated process control system (APCS) information risk management is an important area of information security. The need to assess and process information risks in a multilevel hierarchical structure requires the development of an appropriate algorithm. For industries in which the protection of automated systems is especially critical, for example, automated control systems operating in the field of fuel and energy complexes, significant material damage is possible due to the failure of expensive equipment, and negative economic, environmental or humanitarian consequences, human casualties are not excluded. Ensuring a guaranteed level of APCS information security requires improving the methods for assessing and processing information security (IS) risks. Therefore, carrying out theoretical research and solving practical problems in the field of assessing and processing risks of IS APCS is an urgent task.

One of the tasks in building an APCS information security system is to compile a set of protective measures. At the same time, it is necessary to take into account not only security requirements, but also the costs of implementing and maintaining an information security system. Usually, when drawing up a set of safeguards, a method based on an expert approach is used, which is reduced to the preferences of experts.

This article proposes the use of a genetic algorithm for selecting and optimizing the configuration of an information security system for automated process control systems. Genetic algorithms are a particular case of evolutionary methods for heuristic solution of optimization and structural synthesis problems. Genetic algorithms are based on finding the best solutions by inheriting and enhancing the useful properties of a set of objects of a certain application in the process of simulating their evolution.

## 2 Formulation of the Problem

The study is aimed at improving the existing methods and techniques in assessing ICS information security risks and developing an algorithm based on a metaheuristic approach that allows choosing a set of optimal measures to reduce risks from actual threats to ICS information security.

The purpose of the study is to increase the security of industrial control systems and ensure the necessary level of information protection using a risk-based approach.

Achieving this goal is directly related to the solution of the following scientific and technical problems:

1. conduct an analysis of the typical structure of the APCS, elements and assets;

2. draw up an APCS scheme to describe various types of assets that are subject to security threats, emphasizing the features of the interaction of assets in solving the problem of information security risk management;
3. based on the existing method described in the source [1], develop an optimization algorithm using an APCS scheme that describes different types of assets and their relationships;
4. form a set of protective measures, determine the main assets of the automated process control system, vulnerabilities and threats to the information security of the automated process control system. At the same time, it is necessary to take into account security requirements, as well as the costs of implementing and maintaining an information security system;
5. create a software tool that implements an optimization algorithm to assess information security risks and form an optimal set of protective measures for risk treatment.

## 3 Analysis of the Typical Structure of Industrial Control Systems

Automated process control system—a human–machine control system that provides automated collection and processing of information necessary to optimize the control of a technological object in accordance with the accepted criterion.

Technological control object (TCO) is a set of technological equipment and implemented on it according to the relevant instructions or regulations of the technological process of production.

TCO and APCS form an automated technological complex.

APCS should provide three properties of the technological process, similar to the three properties of information that IS is designed to provide—these are confidentiality, integrity, availability. In the case of automated process control systems—efficiency, reliability and safety [2].

Efficiency is a reduction in the cost of resources and energy resources necessary for the production of a product [3].

Reliability—ensuring the constancy of the quality of the product, incl. in the conditions of failures of technological equipment, reduction of the frequency of failures [3].

Safety—increasing the service life of technological equipment by choosing the optimal operating modes, reducing human participation in the technological process [3].

The management of a technological object is carried out at various levels of management, each of which has its own tasks and capabilities [4]. A typical process control system consists of three levels: lower, middle and upper. The lower level of the automated process control system consists of field equipment, including sensors, actuators. Their activity consists in obtaining process parameters, converting them into the appropriate form for further transmission to a higher level (sensor functions), as well as in receiving control signals and performing appropriate actions (actuator functions). Field equipment is connected to one of the input/output (I/O) subsystem modules using electrical cables. The I/O subsystem consists of hardware I/O modules. Modules differ in the type of electrical signal they use to communicate with field devices and in the direction in which the signal is transmitted.

If a sensor is connected to the module, then the module inputs a signal into the system and is called an input module; if an actuator is connected, then the module outputs the control action from the system and is called the output module (Fig. 1).

At the middle level are programmable logic controllers (PLC) and communication devices with the object (OCD). OCD are a set of modules that provide interface with sensors, actuators and allow you to receive, process, issue signals of various types in a wide range of voltage, current, power, pulse duration.

The PLC is a microprocessor-based device designed to control industrial processes. The principle of operation of the PLC is to collect signals from sensors and process them with the issuance of control signals to actuators. This processing is carried out in accordance with the embedded control algorithms, with the use of which the PLC and OCD control the lower-level actuators and analyze information from the sensors.

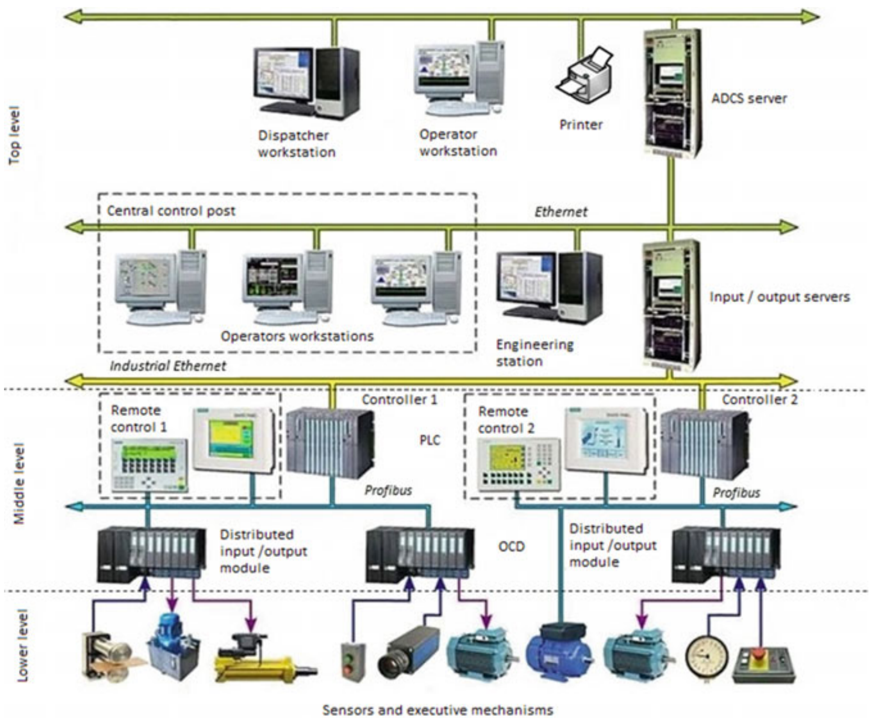



**Fig. 1** Typical structure of process control systems

At the top level is a system for data collection and operational dispatch control (eng., SCADA—Supervisory Control and Data Acquisition) or a distributed control system (DCS). This system is software that is configured and installed on dispatch computers. At this level, control over the production of products is carried out. This process includes the collection of data coming from the production sites, their accumulation, processing and issuance of guiding directives to the lower steps. The upper level of the system consists of input/output servers, an automated dispatch control system (ADCS) server, and an operator workstation. At this level, a person is involved, i.e. operator. It performs local control of technological equipment through a human–machine interface. Often, IP and Ethernet networks are used at the upper levels of the architecture. Communication between different levels is provided by switching nodes, and access to sensors is carried out via field buses using special protocols, such as Profibus.

## 4 Statement of the Problem of Information Security Risk Assessment and Formation of an Optimal Set of Protective Measures for Their Processing

To effectively solve the problem of processing unacceptable IS risks under financial constraints, it is necessary to determine optimization methods that best suit the specifics of the problem being solved. For this, an analysis of existing mathematical methods used for similar problems was carried out.

Optimization problem: it is required to find such an optimal set of risk treatment protective measures for given input parameters, which will provide a minimum of the objective function—the total residual risk when implementing risk reduction measures.

$$\sum_{i=1}^{N_T} r_\lambda^i \rightarrow \min$$

In doing so, the financial constraints $C_{\max}$ on the implementation of risk treatment measures must be taken into account. The cost of implementing risk treatment measures $s_c^j$ must not exceed the financial limits $C_{\max}$ set by the enterprise.

$$\sum_{j=1}^{N_{\text{RTM}}} s_c^j \leq C_{\max}$$

Thus, $r_\lambda^i$—is the residual risk when taking protective measures; $i$—threat number; $N_T$—is the number of threats; $s_c^j$ is the cost of implementing protection measures; $N_{\text{RTM}}$ is the number of risk treatment measures (RTM); $C_{\max}$—financial limits set by the enterprise for the implementation of protective measures.

In the optimization problem, a group of threats is taken to eliminate in order to achieve a minimum overall residual risk assessment. The next condition takes into account the restrictions of the enterprise on the financial costs of the implemented protection measures.

It is supposed to be fair to assume that the appearance of threats is an independent event. In addition, it is assumed that when the selected threat is neutralized in the process of implementing the selected set of RTM, no new threats will appear. The last aspect is a fairly strong assumption. When it is removed, the solution of the optimization problem in general becomes quite complicated. The most preferable is the formulation of the problem of determining the RTM based on the calculation of the minimum residual risk at the current moment [5].

The task of forming optimal protective measures associated with risk reduction, with established restrictions on financial costs, can be solved by enumerating all the proposed solutions, that is, by exhaustive enumeration. The main advantage of the exhaustive search approach is the ease of implementation, the search for the objective functions of a large class, and the calculation of the global optimum. The main disadvantage of this method is the time it takes to find a solution, the time complexity $\left(2^{N_{\text{RTM}}}\right)$. Therefore, the brute force method is only used for a small number of safeguards. With an increase in the number of RTM, the problem cannot be solved by this method in the optimal time.

More complex optimization methods are used to get rid of this disadvantage. In the problem under consideration, the Wald criterion (“maxmin” criterion) is used and there are financial constraints, therefore it is proposed to solve it using the genetic algorithm method.

## 5 Risk Analysis

In the field of information technology and information security, there is the following definition of risk: information security risk is the possibility that this threat will be able to take advantage of the vulnerability of an asset or group of assets and thereby cause damage to the organization.

The risk is considered in the context of some valuable resource (asset).

An asset is anything that has value to an organization. These can be equipment, finance, processes, human and information assets.

The concept of risk is often confused with the concept of threat. A threat is the potential for an adverse event to occur that, with some degree of destructiveness, damages defense mechanisms or directly affects a valuable resource, leading to negative consequences. Consequences can be expressed in the form of financial, material, moral, reputational or other damage [6].

The damage in the implementation of the threat $t^i$ is determined as a value:

$$y_W^i = 3_i^T + \Pi_i^T + \tau_i^T$$

Here $З_i^T$ are indicators of the influence of the threat $t^i$ on the costs of the set of assets $A^i \cdot П_i^T$—indicators of the impact of the threat $t^i$ on the loss of the set of assets $A^i$, measured in percent, $\tau_i^T$—the time of termination or disruption of the operation of the group of assets $A^i$ when the threat $t^i$ is realized, measured in minutes. All this is determined by the expert group based on the Delphi method.

Having assessed the possible damage in the implementation of the threat $t^i$, the owner of the risk needs to assess the possibility of realizing this threat. In this case, the expert method of event trees is used.

The advantage of event tree analysis (ETA) over many other risk analysis methods is its ability to model the sequences and interactions of various protection factors that accompany the occurrence of a triggering event. Thus, the system and its interactions with all protection factors in the development of an unfavorable scenario become visually presented, which contributes to further risk assessment.

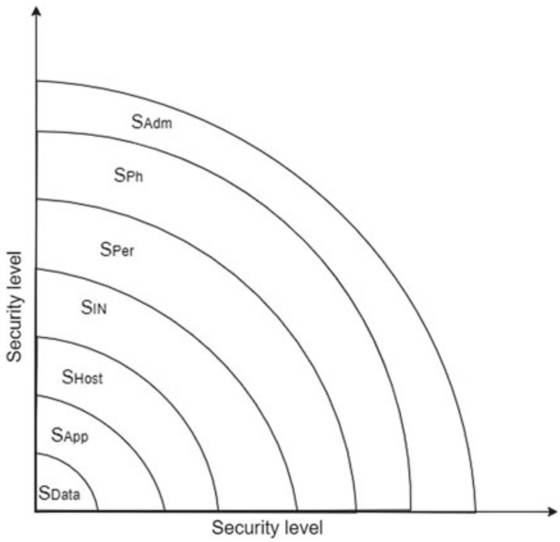
Using ETA, the possibility of neutralizing the threat at each of the existing levels of protection is assessed and, according to the results obtained, the possibility of realizing the threat $t^i$ is determined as:

$$
\begin{aligned}
p^i = & \left(1 - p_1^i\right) \times \left(1 - p_2^i\right) \times \left(1 - p_3^i\right) \\
& \times \left(1 - p_4^i\right) \times \left(1 - p_5^i\right) \times \left(1 - p_6^i\right) \\
& \times \left(1 - p_7^i\right)
\end{aligned}
$$

We talk about levels of protection using the principle of «protection in depth» [7] (Fig. 2).

According to the proposed principle, all risk mitigation measures are divided into protection levels by spheres of influence: administrative measures (financial penalties

**Fig. 2** Protection levels



for violating established information security rules, etc.), physical security (creation of measures to prevent unauthorized entry to the facility), perimeter (firewall to filter traffic from the external network), internal network (monitoring equipment inside the network, etc.), device (using anti-virus protection, etc.), application (monitoring vulnerabilities, etc.), information (data encryption, backup, etc.).

Thus, each risk mitigation measure can be attributed to one of the protection levels. $S_{\mathrm{Adm}}$—administrative measures, $S_{\mathrm{Ph}}$—physical security, $S_{\mathrm{Per}}$—perimeter, $S_{\mathrm{IN}}$—internal network, $S_{\mathrm{Host}}$—device, $S_{\mathrm{App}}$—application or $S_{\mathrm{Data}}$—data.

In addition to assessing the damage from the implementation of the threat and the possibility of the implementation of the threat, it is necessary to determine the assessment of vulnerabilities. Vulnerability assessment systems specialized for industrial control systems do not currently exist [8]. It uses many different methods to assess vulnerabilities, such as the US-CERT and the SANS Institute's Critical Vulnerability Analysis Scale. However, the Common Vulnerability Scoring System (CVSS) is the predominant technique for most vulnerability databases.

CVSS determines the severity value of a particular vulnerability implementation based on the values of its characteristics.

The values of the vulnerability vector $\mathrm{v}^i$ and its resulting score will be calculated using the formula:

$$\begin{aligned}\mathrm{CVSS}\left(\mathrm{v}^i\right) &= \mathrm{CVSS{:}3.1}/\mathrm{AV} \\ &\quad : X/\mathrm{AC} : X/\mathrm{PR} \\ &\quad : X/\mathrm{UI} : X/S \\ &\quad : X/C : X/I \\ &\quad : X/A : X,\end{aligned}$$

where $X$ is the metrics score, $\mathrm{CVSS}\left(\mathrm{v}^i\right)$ is the quantitative vulnerability score $\mathrm{v}_i$, ranging from 0 to 10.

The magnitude of the CVSS vulnerability vector for a particular valuable resource can be calculated by taking the maximum criteria of all vulnerabilities associated with the asset:

$$\begin{aligned}\mathrm{CVS}\left(a^j\right) &= \mathrm{CVSS{:}3.1}/\mathrm{AV} \\ &\quad : \max X/\mathrm{AC} : \max X/\mathrm{PR} \\ &\quad : \max X/\mathrm{UI} : \max X/S \\ &\quad : \max X/C : \max X/I \\ &\quad : \max X/A : \max X\end{aligned}$$

## 6 Description of the Method of Risk Management and Information Security of an Automated Process Control System

Risk management of IS APCS is based on standard GOST ISO/IEC 27,005. In accordance with the standard the information security risk management process includes the steps for setting the context, assessing, treating, accepting, monitoring and reassessing the information security risk. The risk assessment and treatment processes can be performed iteratively. An iterative approach allows to invest time and effort in a balanced way in selecting controls and management, in the same time providing an appropriate assessment of high-level risks.

When establishing the context, the relationship between the assets of different levels of the automated system is determined according to its model. The next step is to identify threats and vulnerabilities for the identified assets and relationships. Risk analysis includes an assessment of the damage from the implementation of the threat, an assessment of the possibility of the implementation of the threat and the definition of measures for the treatment of risks. Financial constraints at this stage are also determined.

Risk assessment is carried out for each individual threat according to the formula:

$$R = P \cdot V \cdot D,$$

where $P$ is the probability of the threat being realized, $V$ is the vulnerability assessment, $D$ is the damage in the event of an attack.

There is a search for the optimal set of protective measures to reduce the risks of IS APCS by using a genetic algorithm at the stage of risk processing. Further, an optimal set of protective measures for risk treatment is formed that satisfies the specified financial constraints at the risk acceptance stage.

Individual APCS assets can be upgraded. In addition, information about new computer security vulnerabilities in existing software is published in public sources. All this leads to the need to monitor and reassess the risk of the system throughout the entire life cycle.

Thus, the APCS risk management method can be presented in the form of the following block diagram (Fig. 3).

The method was proposed by Andrey Rimsha in his Ph.D. dissertation [1]. Based on the existing method, an optimized genetic algorithm was developed, which differs from the one proposed by A.S. Rimsha based on dynamic programming. The stages of risk assessment, processing and acceptance have been modified.

Here, at the stage of processing and accepting the risk, a genetic algorithm was used to search for the optimal set of protective measures to reduce the risks of IS APCS.

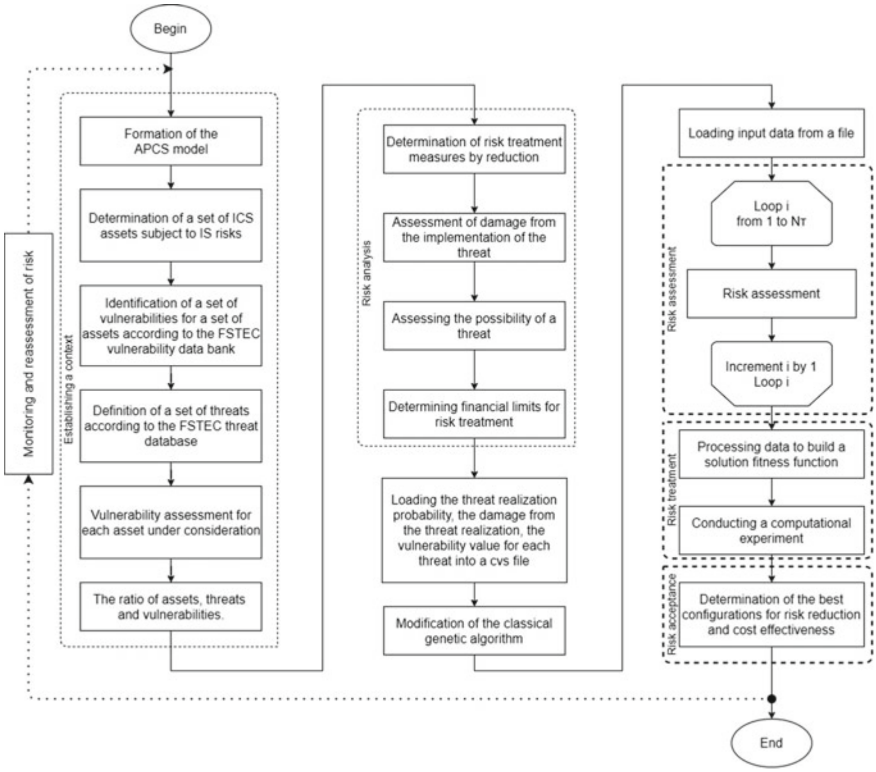



**Fig. 3** Flowchart of the IS APCS risk management process

## 7 Construction of an Optimization Algorithm for Solving the Problem of an Optimal Set of Protective Measures

Due to the generally large dimension of the problem, as well as the need to consider resource constraints and the variety of products, the problem being solved belongs to the so-called NP-complete in the field of discrete optimization. Therefore, it requires large computational costs to solve [9].

It is expedient to solve the considered problem of formation of a set of protective measures based on optimization methods. In this case, for problems of relatively small dimension, it is advisable to use the exhaustive search method, which guarantees the determination of the global extremum. When solving the problem of formation of a set of protective measures of a relatively large dimension, it is necessary to use evolutionary population methods [10–13], to which the genetic algorithm belongs. Below is a description of the genetic algorithm used to solve the problem [14–18].

Note that genetic algorithms can implement different target criteria (fitness functions) and constraints. This property is very useful for formation of an optimal set of protective measures under financial constraints. In the considered problem of optimal set of protective measures it is necessary to introduce appropriate definitions and assumptions for the use of genetic algorithms. So, everyone is a possible solution to the problem, i.e., set of protective measures. An important advantage of genetic algorithms is the ability to use several points for the search space under consideration, thus parallelization of the search is possible. In this case, the genetic algorithm uses information that is contained in the entire set of feasible solutions [19–22].

The main concepts of the genetic algorithm as a method for solving the problem of optimization of the set of protective measures are as follows [23–26].

***Gene*** It represents a unit of so-called hereditary information, i.e., is an integral part of the chromosome, for the internal representation of alternative solutions. In the present work gene is a protective measure.

***Chromosome*** It is an ordered sequence of genes in the form of a coded data structure that determines a decision. The chromosome will be a set of defenses encoded as a binary number. Each bit equal to 1 means that a protective measure with the corresponding number is included in the configuration. Range of code change:

$$G = (d_0 d_1 \ldots d_L)_2 = (0 \ldots 2L)_{10},$$

where $L$ is the number of safeguards available, $d_i$ is whether or not the safeguard is included in the set.

***Population*** It is a set of individuals—potential solutions (the optimal set of protective measures and residual risk assessment), which are represented by chromosomes.

The population is made up of individuals with different chromosomes. Its size is limited by $K$, the maximum number of specimens in a population. Each instance is defined by the set:

$$X = \{G, C, R\},$$

where $G$ is the genetic code, $C$ is the cost of protective measures, $R$ is the total risk, taking into account the protective measures taken.

Let's consider the main operators of the genetic algorithm.

***Crossbreeding*** The function of the crossover is the generation of new descendants. For each pair, a new instance is created that inherits the traits of the parents.

A multipoint crossover was used in the work. With multipoint crossover, k cut points $\mathrm{dot}_i \in \{1, 2, \ldots, N_{\mathrm{gene}}\}$, $i = \underline{1, k}$, $N_{\mathrm{gene}}$ is the number of genes in an individual, are randomly selected. Next, the cut points are sorted from smallest to largest. Parts of chromosomes limited by cut points are exchanged with each other. Ultimately, two new offspring are formed. The part of the individual with the first gene up to the first point of the cut does not take part in the exchange.

Multipoint crossover obliges to introduce several variables, that is, cut points. To create offspring, you need to select individuals with the greatest fitness.

If a multipoint crossing-over has a large even number of crossing points, then it is convenient to depict it for chromosomes as rings. In this case, the chromosome appears as a closed ring, and the choice of crossing points occurs with equal probability over the entire circle.

Four points (borders for exchange) of the crossover are randomly determined. Crossover generates a child by selecting a tour from one parent and retaining the position and guard order from the other parent. For example, for parents $G_1 = (0\ 0\ 1\ 1 \mid 0\ 0 \mid 1\ 1\ 1 \mid 00 \mid 0)$ and $G_2 = (1\ 0\ 1\ 0 \mid 1\ 1 \mid 0\ 1\ 1 \mid 0\ 1 \mid 1)$ we form the child as follows.

The first step is to exchange selected subtours, as a result we get $\mathrm{Ch}_1 = (X\ X\ X\ X \mid 1\ 1 \mid X\ X\ X \mid 0\ 1 \mid X)$ and $\mathrm{Ch}_2 = (X\ X\ X\ X \mid 0\ 0 \mid X\ X\ X \mid 0\ 0 \mid X)$, where "$X$" is an empty position. Next, in place of "$X$" we put the numbers of protective measures from the original parents for which there are no conflicts (no cycle is formed). As a result, we get two children: $\mathrm{Ch}_1 = (0\ 0\ 1\ 1 \mid 1\ 1 \mid 1\ 1\ 1 \mid 0\ 1 \mid 0)$ and $\mathrm{Ch}_2 = (1\ 0\ 1\ 0 \mid 0\ 0 \mid 0\ 1\ 1 \mid 0\ 0 \mid 1)$.

***Mutation*** This is a random change in one or more positions on the chromosome. According to the GA scheme, the mutation operator is executed with a given probability $P_m$. The mutation allows to maintain the genetic diversity of individuals so that the population does not degenerate and the chromosomes do not become similar to each other. The function of mutation is the variation of the genetic code. Randomly inverted two binary digits in the chromosome. Bit inversion [15] was used as a mutation operator for the problem being solved.

***Reproduction*** This is the process by which chromosomes are copied into an intermediate population for further "propagation" in accordance with their objective function values. Thus, one or more descendants with better objective function scores are more likely to get into the next generation. Before selection is applied, it is necessary to calculate the cost of protective measures and the magnitude of the risk, taking into account the set of protective measures for the population. The initial population must consist of at least two specimens. In the work, the reproduction operator is implemented according to the roulette rule.

As a fitness function, the objective function is used

$$\sum_{i=1}^{N_T} r_\lambda^i \rightarrow \min$$

The algorithm consists of the following steps:

1. loading all necessary data from file;
2. formation of the initial population;
3. calculation of the fitness function for the initial population;
4. entering the while loop with the condition of checking the number of populations with the initially specified maximum allowable population value;
5. application of the operator of reproduction;

6. application of the crossing-over operator for selected individuals;
7. mutation of individuals;
8. selection of the best individual according to the lowest value of the fitness function.

The block diagram of the genetic algorithm is shown in the Fig. 4.

The program that implements the genetic algorithm is written in the C++ programming language.

To study the influence of the algorithm settings (size of the initial population, crossover probability, mutation probability, stopping criterion) and operator types (selection of the parent pool, selection of parental pairs, type of crossover, type of mutation, formation of a new population), it is advisable to use a low-dimensional problem. For it, using the brute force method, a global extremum is determined, and it is possible to compare the solution obtained based on the genetic algorithm

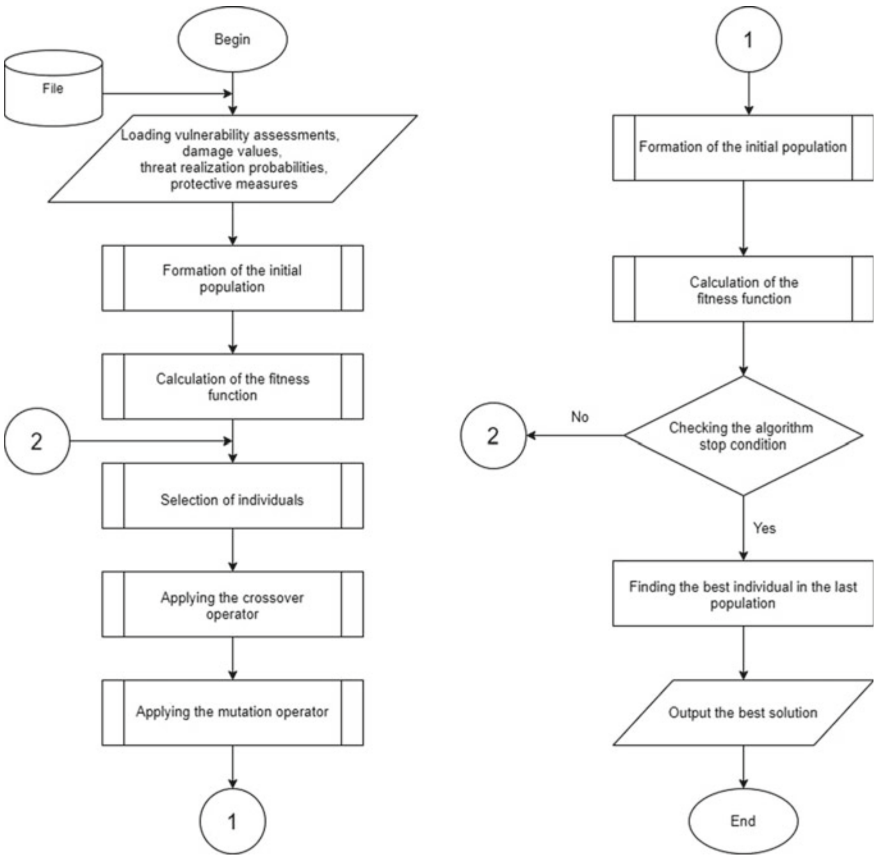



**Fig. 4** Block diagram of the implemented genetic algorithm

to conclude the dependence of the rate of convergence to the optimal value on the settings of the genetic algorithm.

## 8 Test Approbation of the Program According to the Data of the Gas Producing Enterprise "Achimgaz"

First, a list of assets is formed based on the scheme of a typical automated process control system for a gas production enterprise. The list of assets of the middle and upper levels is presented in Table 1.

Using a database of threats and vulnerabilities, typical threats and vulnerabilities are identified. Then a list of protective measures is formed. The list of protective measures is given in Table 2. Suppose that the management of the gas production enterprise has set restrictions on the inability to change the critical process and involve outside companies for continuous maintenance of the process control system, as well as financial restrictions on the implementation of risk treatment measures in the amount of 7 million rubles. The cost of implementation for each element of the set a $S_{\text{RTM}}$ is given.

The calculated probabilities of the threat, the damage in the event of an attack, and the magnitude of the vulnerability for each threat were recorded in a file with the

**Table 1** List of middle and upper level assets

| Assets | Description |
|---|---|
| $a^1_{\text{PLC}}, \ldots, a^{15}_{\text{PLC}}$ | PLC collecting information from sensors and controlling actuators |
| $a^1_{\text{network hardware}}$ | Primary firewall located between industrial and corporate networks |
| $a^2_{\text{network hardware}}$ | Backup firewall located between industrial and corporate networks |
| $a^3_{\text{network hardware}}$ | Access point for communication with subscriber line modules |
| $a^4_{\text{network hardware}}, \ldots, a^9_{\text{network hardware}}$ | Subscriber module (DLU) |
| $a^1_{\text{server}}$ | Main server responsible for the operation of PLC algorithms |
| $a^2_{\text{server}}$ | Server for collecting data on history, reports, events |
| $a^3_{\text{server}}$ | Data collection server for information from sensors of a pipeline intended for transporting gas condensate, data from uninterruptible power supplies |
| $a^4_{\text{server}}$ | Server for transferring data to a corporate network |
| $a^5_{\text{server}}$ | Time server |
| $a^1_{\text{AWP}}$ | Engineering station for administration |
| $a^2_{\text{AWP}}, \ldots, a^6_{\text{AWP}}$ | Dispatcher's stations |

**Table 2** Suggested protective measures

| $S_{MOP}$ | $S$ | Description | Cost, rubles |
|---|---|---|---|
| $S^1_{adm}$ | $S^1$ | The introduction of penalties for violation of the rules | 2000 |
| $S^1_{ph}$ | $S^2$ | Implementation of an access control and management system throughout the facility | 3,020,000 |
| $S^2_{ph}$ | $S^3$ | Installation of burglar alarm sensors in equipment rooms | 105,000 |
| $S^3_{ph}$ | $S^4$ | Installation of electronic locks compatible with the access control system in telecommunication cabinets | 4,010,000 |
| $S^4_{ph}$ | $S^5$ | Installation of metal detector frames and X-ray television introscope at the guard post | 2,630,000 |
| $S^5_{ph}$ | $S^6$ | Organization of video surveillance in rooms with equipment | 520,000 |
| $S^1_{per}$ | $S^7$ | Configuring a White List of device MAC addresses for the firewall | 4000 |
| $S^2_{per}$ | $S^8$ | Creating a DMZ (installing two additional firewalls) | 408,000 |
| $S^3_{per}$ | $S^9$ | Installing Unidirectional Gateways | 2,508,000 |
| $S^1_{IN}$ | $S^{10}$ | Create a standby AD read-only server and restrict access to the primary AD server | 158,000 |
| $S^2_{IN}$ | $S^{11}$ | Installation of a system for monitoring the status of equipment and APCS software based on Zabbix | 58,000 |
| $S^1_{Host}$ | $S^{12}$ | Authorization in the system by means of an access card and a PIN code | 1,008,000 |
| $S^1_{App}$ | $S^{13}$ | Tracking new vulnerabilities and fixing them with antivirus protection tools | 4000 |
| $S^1_{Data}$ | $S^{14}$ | Regular backup of information to a backup server | 4000 |

cvs extension. In total, 14 protective measures $G$ and their cost $C$, 14 vulnerability assessments $V$, 14 damage assessments in the implementation of threats, 14 assessments of the probability of a threat implementation are loaded into the file (Table 3). The cost of implementation for each safeguard was determined.

The risk value is calculated as the sum of the products of the probability of the threat realization $p^i$ by the vulnerability assessment $\mathrm{v}^i$ and the damage in the event of an attack $d^i$. In the initial configuration R_0 = 421,945.8376.

Table 4 shows the result of the program based on the genetic algorithm. The cells of the table present the ratio of the assessment of the effectiveness of the protection system, taking into account the set of protective measures, to the cost of the set of protective measures, with the corresponding percentage of crossing individuals and elitism from the entire population, as well as the power of the population. The best solution is highlighted in yellow. The best solution is determined as the most effective, with the lowest residual risk assessment, and the most appropriate cost of implementing protective measures within the established restrictions.

Variable values of the parameters of the genetic algorithm:

- the number of individuals in the population $N = 50, 200$;
- the probability of crossing over 30, 50, 70, 90%;

**Table 3** Risk.cvs file data

| *G* | *V* | *C* | *D* | *P* |
|---|---|---|---|---|
| 1 | 0.98 | 2000 | 43,939 | 0.73 |
| 2 | 0.84 | 3,020,000 | 117,318 | 0.3 |
| 3 | 1 | 105,000 | 34,707 | 0.86 |
| 4 | 0.68 | 4,010,000 | 32,297 | 0.77 |
| 5 | 0.98 | 2,630,000 | 48,994 | 0.65 |
| 6 | 1 | 520,000 | 28,349 | 0.81 |
| 7 | 0.98 | 4000 | 64,559 | 0.53 |
| 8 | 0.86 | 408,000 | 108,443 | 0.4 |
| 9 | 0.84 | 2,508,000 | 52,560 | 0.69 |
| 10 | 0.8 | 158,000 | 50,658 | 0.77 |
| 11 | 0.88 | 58,000 | 62,272 | 0.61 |
| 12 | 0.84 | 1,008,000 | 118,839 | 0.34 |
| 13 | 1 | 4000 | 71,289 | 0.49 |
| 14 | 0.88 | 4000 | 37,207 | 0.77 |

**Table 4** The result of the genetic algorithm for different values of the variable parameters

| *N* | Percentage of interbreeding individuals from the total population | Mutation percentage | | |
|---|---|---|---|---|
| | | 30% | 50% | 70% |
| 50 | 30% | 69,438.3/6889000 | 46,475.6/7409000 | 78,580.6/7763000 |
| | 50% | 46,475.6/7409000 | 46,475.6/7409000 | 73,332.1/7795000 |
| | 70% | 47,375.3/7921000 | 46,475.6/7409000 | 46,475.6/7409000 |
| | 90% | 46,475.6/7409000 | 47,375.3/7921000 | 52,557.9/7831000 |
| *N* | Percentage of interbreeding individuals from the total population | Mutation percentage | | |
| | | 30% | 50% | 70% |
| 200 | 30% | 49,874.2/7909000 | 46,475.6/7409000 | 73,332.1/7795000 |
| | 50% | 46,475.6/7409000 | 46,475.6/7409000 | 49,874.2/7909000 |
| | 70% | 47,375.3/7921000 | 49,874.2/7909000 | 48,120.7/7799000 |
| | 90% | 73,332.1/7795000 | 78,580.6/7763000 | 52,557.9/7831000 |

- the probability of mutation 30, 50, 70%;
- crossing over: multy-point;
- mutation: bit inversion.

**Table 5** Best configurations for risk reduction after 10 epochs

| Protective measures configuration | Risk considering a set of protective measures, ruble | The cost of protective measures, ruble | Residual risk after the implementation of measures, ruble |
|---|---|---|---|
| 10101111111111 | 3,754,710.9 | 7,409,000 | 46,475.6 |
| 11101111011111 | 374,571.4 | 7,921,000 | 47,375.3 |
| 11100111111111 | 373,825.3 | 7,799,000 | 48,120.7 |
| 11101011111111 | 382,072.4 | 7,909,000 | 49,874.2 |
| 10110111111111 | 361,171.9 | 7,831,000 | 52,557.9 |
| 10101011111111 | 352,508.3 | 6,889,000 | 69,438.3 |
| 11101111111111 | 405,035.1 | 7,795,000 | 73,332.1 |
| 11001111111110 | 375,187.1 | 7,763,000 | 78,580.6 |

With a population power of 50, with 50 and 70% of crossed individuals from the total population, with 50 and 70% mutations, the most appropriate values of the optimal configuration of protective measures were obtained. With the cost of a set of protective measures 7,409,000, the residual risk after the implementation of protective measures will be 46,475.6.

With a population power of 200, the most appropriate values of the optimal configuration of protective measures are obtained with 50% of crossed individuals from the total population and 50% of the mutation. With an increase in the number of individuals, an increase in the running time of the algorithm is observed.

From the studies carried out, it was possible to identify the relationship between an increase in the probability of a mutation and the optimal solution. The optimal value of the mutation probability is 0.5.

For each carrier, according to its genetic code, a set of protective measures is selected, using which the model parameters are changed, and the total risk is calculated. Further, for all variants of the genetic code, a computational experiment is carried out to determine the best configurations for risk reduction and cost efficiency (Table 5).

## 9 Analysis and Conclusions from the Results of Software Testing to Find the Optimal Configuration of Protective Measures

During the analysis, the following conclusions were made:

1. It should be noted that in the proposed algorithm, a larger amount of risk calculations $\left(2^{\text{RTM}}\right)^2$ is carried out compared to the enumeration of all options $2^{\text{MOP}}$. For 10 epochs, 50,069 calculations of the risk value were carried out, for 5 epochs—17,518.

2. The running time of the algorithm was analyzed for various numbers of measures to reduce the risk. For 5 RTMs, the execution time of the genetic algorithm was 0.029 s. For 9 RTMs, the execution time of the genetic algorithm was 0.036 s. For 14 RTMs, the execution time of the genetic algorithm was 0.048 s.
3. At high values of the crossover probability ($P_c = 0.7, P_c = 0.9$) and the mutation probability ($P_m = 0.7$), the algorithm did not always find a solution that satisfies the accuracy. This can be explained by the fact that the solutions prematurely converge to one point.
4. The best results were observed at 50% crossover, that is, the probability of crossing over 0.5, and at 50% mutation. Increasing the values further does not improve performance, and sometimes worsens the search for the optimal solution.
5. The most efficient in terms of the number of generations and the time spent on finding a solution is the use of a population of 50 individuals.
6. The most cost-effective protection system consists of introducing fines for non-compliance with established rules ($C = 2000$), installing intruder alarm sensors in equipment rooms ($C = 105{,}000$), installing metal detector frames and an X-ray television introscope at the guard post ($C = 2{,}630{,}000$), organizing video surveillance in rooms with equipment ($C = 520{,}000$), setting up filtering by the white list of MAC addresses ($C = 4000$), installing two additional firewalls ($C = 408{,}000$), installing data transmission gateways from one LAN to another only in one direction ($C = 2{,}508{,}000$), creating a backup AD read-only server and restricting access to the main AD server ($C = 158{,}000$), establishing a Zabbix system for monitoring the status of hardware and programs ($C = 58{,}000$), obtaining rights in the system using a PIN code and card of access ($C = 1{,}008{,}000$), monitoring of newly formed vulnerabilities and their elimination using anti-virus ($C = 4000$), regular backup of information cations to the backup server ($C = 4000$). With the resulting set of protective measures, the total residual risk when implementing risk treatment measures will decrease by 46,475.6, and the cost of implementing measures is 7409000 rubles, which is within the financial limits set by the enterprise.

Thus, the cost of implementing the RTM is 7409000 rubles, which satisfies the established financial constraints.

## 10 Conclusions

The article considers an important area of information security—management of information risks of industrial control systems. The need to assess and process information risks in a multi-level hierarchical structure required the development of an appropriate algorithm. An analysis of the information security of the APCS was carried out, the levels of process control and the main functions of the APCS, the main threats to the information security of the APCS were considered.

The purpose of the work was to improve the efficiency of information security of automated process control systems in the face of destructive impacts through the use of a risk-based approach. On the basis of the metaheuristic approach, an algorithm was developed for assessing the risks of information security of automated process control systems and the formation of an optimal set of protective measures for their processing.

A software tool in the C++ programming language has been developed to implement the algorithm for assessing the risks of information security of automated process control systems and the formation of an optimal set of protective measures for their processing. The program receives input data from a file, and then converts the received input parameters of the task into elements of the algorithm: gene, chromosome, population, fitness function. In the process of research, in order to achieve the best optimization result, different parameters and characteristics of the algorithm were varied, such as the number of iterations, population size, mutation and crossover probabilities.

## References


1. Rimsha, A.S.: Method and Algorithms for Managing Information Security Risks of Automated Process Control Systems for Critical Infrastructures. Ph.D. thesis (2022)
2. Pishchik, B.N.: Safety of process control systems. Comput. Technol. **18**, 170–175 (2013)
3. Alexandrovskaya, L.N., Aronov, I.Z., Kruglov, V.I., Kuznetsov, A.G., Patrakov, N.N., Sholom, A.M.: Safety and Reliability of Technical Systems, p. 376 (2004)
4. Rimsha, A.S.: Method, model and software package for the analysis of information risks in the automated process control system of gas producing enterprises. Math. Methods Eng. Technol. Sat. Tr. Int. Sci. Conf. **12**(2), 124–129. SPbPU, St. Petersburg (2019)
5. Bolshakov, A.A., Rimsha, A.S.: Mathematical support for analysis and management of information risks in the automated process control system of gas producing enterprises. Math. Methods Eng. Technol. **10**, 73–78. St. Petersburg (2021)
6. Azhmukhamedov, I.M., Vybornova, O.N., Knyazeva, O.M.: Risk analysis of information security. Federal State Budgetary Educational Institution of Higher Professional Education Astrakhan State Technical University, Astrakhan (2015)
7. Abdelghani, T.: Implementation of defense in depth strategy to secure industrial control system in critical infrastructures. Am. J. Artif. Intell. **3**(2), 17–22 (2019)
8. Fedorchenko, A.V., Chechulin, A.A., Kotenko, I.V.: Research of open databases of vulnerabilities and assessment of the possibility of their application in systems for analyzing the security of computer networks. Inf. Control Syst. **5**, 72–79 (2014)
9. Protalinskiy, O., Shvedov, A., Khanova, A.: Life cycle management of power grid companies' equipment. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Cyber-Physical Systems: Design and Application for Industry 4.0. Studies in Systems, Decision and Control, vol. 342, pp. 265–274. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-66081-9_21
10. Skobtsov, Y.: Prospects of the interdisciplinary course "computational intelligence" in engineering education. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Cyber-Physical Systems: Design and Application for Industry 4.0. Studies in Systems, Decision and Control, vol. 342, pp. 431–441. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-66081-9_33

11. Skobtsov, Y., Chengar, O., Skobtsov, V., Pavlov, A.: Synthesis production schedules based on ant colony optimization method. In: Kacprzyk, J., et al. (eds.) Proceedings of the 6th Computer Science Online Conference 2017 (CSOC2017). Artificial Intelligence Trends in Intelligent Systems. Advances in Intelligent Systems and Computing, vol. 573, pp. 456–465. Springer International Publishing, Switzerland, Cham (2017)
12. Karpenko, A.P., Svianadze, Z.O.: Meta-optimization based on self-organizing map and genetic algorithm. Opt. Mem. Neural Netw. **20**(4), 279–283 (2011)
13. Chistyakova, T.B., et al.: Decision support system for optimal production planning polymeric materials using genetic algorithms. In: 2016 XIX IEEE International Conference on Soft Computing and Measurements (SCM), pp. 257–259. IEEE (2016)
14. Mitra, K.: Genetic algorithms in polymeric material production, design, processing and other applications: a review. Int. Mater. Rev. **53**, 275–297 (2008)
15. Man, K.F., Tang, K.S., Kwong, S.: Genetic algorithms in production planning and scheduling problems. In: Genetic Algorithms. Advanced Textbooks in Control and Signal Processing, pp. 259–280. Springer, London (1999)
16. Kucukkoc, I., Zhang, D.Z.: Integrating ant colony and genetic algorithms in the balancing and scheduling of complex assembly lines. Int. J. Adv. Manuf. Technol. **82**, 265–285. Springer (2016)
17. Tahmasebi, P., Hezarkhani, A.: A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation. Comput. Geosci. **42**, 18–27 (2012). https://doi.org/10.1016/j.cageo.2012.02.004. Accessed 9 Mar 2021
18. Dagaeva, M., Katasev, A.: Fuzzy rules reduction in knowledge bases of decision support systems by objects state evaluation. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Cyber-Physical Systems: Modelling and Intelligent Control. Studies in Systems, Decision and Control, vol. 338, pp. 113–123. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-66077-2_9
19. Zaborovskij, V., Polyanskiy, V., Popov, S.: On the problem of computability of bounded rationality cognitive solutions. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, vol. 333, pp. 15–23. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-63563-3_2
20. Bolshakov, A., Nikitina, M., Kalimullina, R.: Intelligent system for determining the presence of falsification in meat products based on histological methods. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, vol. 333, pp. 179–201. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-63563-3_15
21. Dykin, V.S., Musatov, V.Y., Varezhnikov, A.S., Bolshakov, A.A., Sysoev, V.V.: Application of genetic algorithm to configure artificial neural network for processing a vector multisensory array signal. In: 2015 International Siberian Conference on Control and Communications (SIBCON), pp. 719–722 (2015). https://doi.org/10.1109/SIBCON.2015.7147049. ISBN: 978–147997102–2
22. Valls, V., Ballestín, F., Quintanilla, S.: A hybrid genetic algorithm for the resource constrained project scheduling problem. Eur. J. Oper. Res. **185**(2), 495–508 (2008)
23. Galkin, A., Istomin, V.: Automation of assembly batches installation in hot rolling mills. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M.V. (eds.) Cyber-Physical Systems. Studies in Systems, Decision and Control, vol. 350, pp. 53–62. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-67892-0_5
24. Moshev, M., Meshalkin, V., Romashkin, M.: Development of models and algorithms for intellectual support of life cycle of chemical production equipment. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Cyber-Physical Systems: Advances in Design & Modelling. Studies in Systems, Decision and Control, vol. 259, pp. 153–165. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-32579-4_12
25. Belov, V., Moshev, E.R.: Functional model for the formation of individual metal control programs of boiler equipment. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.)

Cyber-Physical Systems: Modelling and Intelligent Control. Studies in Systems, Decision and Control, vol. 338, pp. 323–334. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-66077-2_26

26. Dorrer, M., Dorrer, A., Zyryanov, A.: Numerical modeling of business processes using the apparatus of GERT networks. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, vol. 333, pp. 47–55. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-63563-3_5

# Application of PID Control Principles in the Tasks of Modeling the Movement of Wheeled Vehicles Equipped with an Anti-Lock Braking System

**Grigory Boyko, Alexey Fedin, and Jozef Redl**

**Abstract** The steady increase in the speed and intensity of road traffic, and as a result, a decrease in the distance between road users in the traffic flow, inevitably leads to an increase in the number of accidents, thus, the relevance of the issue considered by the authors of the chapter is beyond doubt, but, with Over time, this issue will become more and more relevant. Improving the efficiency of the braking system of ground wheeled vehicles in real operating conditions requires the optimization of methods for solving the mathematical model of the control element of the brake force distribution system of ground wheeled vehicles. The scientific novelty of this study lies in the use of fuzzy logic elements in the algorithm for the functioning of the electronic control unit of the brake force distribution system in relation to the tasks of braking ground wheeled vehicles.

**Keywords** Wheeled vehicle · Braking process · Anti-lock braking system · Simulation · Control · Optimization · Fuzzy logic

## 1 Introduction

The growing role of artificial intelligence, fusion robotics with industry and business, implementation «Internet of things» in automated production led to the fourth industrial revolution. But Industry 4.0 is only part of a larger process. Today, the Japanese government is talking about the birth of a super-intellectual society—Society 5.0 [1].

The use of various sensors that can collect and process Big Data in a car is one of the trends in the Society 5.0 strategy. The car drives around the city and collects data about the surrounding area, helping the driver make the right decisions in various

G. Boyko (✉) · A. Fedin
Volgograd State Technical University, Volgograd, Russian Federation
e-mail: boyko@vstu.ru

J. Redl
Slovak University of Agriculture in Nitra, Nitra, Slovak Republic
e-mail: jozef.redl@uniag.sk

 
A. G. Kravets et al. (eds.), *Society 5.0*, Studies in Systems, Decision and Control 437,
https://doi.org/10.1007/978-3-031-35875-3_24

situations, and, for example, in braking mode, minimizes driver errors, especially at the time of controlled wheel slip in braking mode [2–5].

The concept of controlled wheel slip during braking has been known for about a century. So, already in the 30 s of the XX century anti-lock braking systems (ABS) were used in aircraft landing gear. Also, the concept of controlled wheel slip in braking mode is used in railway traffic problems. However, this concept became widespread among ground wheeled transport only in the last quarter of the twentieth century. The introduction of anti-lock braking systems into serial models of cars was associated as well with the requirements to improve road safety in the face of an increase in speed limits and traffic density as the possibility of implementing technical solutions that appeared during the period of rapid development of electronics and information technologies [6–8].

In addition to anti-lock braking systems, controlled wheel slip is used in the clutch control systems (Traction Control), exchange rate stability (ESP), etc. At the present time, various options for implementing the principles of automatic control theory in electronic units of braking systems are known, all of which have their advantages, as well as disadvantages. On the other hand, over the past two centuries, there has been an increasing importance and prevalence of automatic control in various technical systems. Starting from the third quarter of the XIX century. There are strict provisions of the theory of automatic control [9, 10].

Initially, the principles of the theory of automatic control were applicable to linear or linearized forms of problems; later, methods were developed and continue to develop methods applicable to nonlinear problems. An important step in the development of an automatic control system is the determination of the control principle and the choice of the type of controller. The prerequisites for solving the problem of parametrization of all possible stabilizing controllers related to a particular control object began to appear in the works of researchers in the last quarter of the twentieth century and originally referred only to finite-dimensional linear stationary systems with feedback. Subsequently, studies were carried out in relation to both non-stationary, and nonlinear, and systems with infinite dimensions. Let's approach the issue of modeling an anti-lock braking system with an emphasis on the automatic control system [11, 12].

## 2 Features of Modeling the Automatic Control System of the Anti-Lock Braking System

The task of the automatic control system of the anti-lock braking system is to maintain the value of the argument S, corresponding to the maximum value of the function φ(S), by transferring the braking torque to the brake disk. Note that for different types of pavements, the maximum of the function φ(S) is in the vicinity of the value of the argument 0.2, while the maximum values of the functions themselves differ significantly. In the simplest automatic control systems (ACS) of ABS was proposed to use the PID control principles on fuzzy control with regulation based on the deviation of the wheel longitudinal slip coefficient from the target value of 0.2. In general ACS for ABS are known based on the principles of PID control, adaptive, optimal, robust, fuzzy, intelligent control, as well as their various combinations and hybrids of both single input and multi-input types [13, 14]. The simplest systems in terms of describing the mathematical model and the technical implementation of the solution, as a rule, have the worst adequacy in relation to real processes. Solutions with greater adequacy can be significantly more complex in terms of mathematical modeling and require deep knowledge from developers in several different areas of several different scientific disciplines at once. Thus, designing an automatic control system for an anti-lock braking system with an abundance of existing solutions remains a non-trivial task.

The main goal of the anti-lock braking system of a car is to reduce the braking distance to a minimum while maintaining controllability during braking [15]. The main input data for calculating key parameters are the angular velocity (acceleration) of the wheel and the linear velocity (acceleration) of the car. The key parameters that determine the operation of the ABS system are the coefficient of longitudinal slip S and the coefficient of longitudinal adhesion of the wheel to the road surface $\varphi$. To determine the angular velocity of the wheel, mainly pulsed magnetic sensors are used. To determine the linear speed of the vehicle, several methods are known based on processing the signals of the gearbox output shaft speed sensor [16]; simultaneous processing of signals from several sensors of the angular velocity of the vehicle wheels [16]; accelerometer data processing; data processing of the GPS module, as well as various combinations of the indicated methods. Based on the received data, the software of the anti-lock braking system control unit (ECU ABS) distributes braking forces. The efficiency of the anti-lock braking system is determined both by the error in the readings of the indicated sensors and by the adequacy of the process of processing the input data by the computing device [17]. This implies the necessity of striving for a high level of accuracy of correspondence of approximations to real values of parameters in modeling and calculating the movement of an automobile wheel. A satisfactory result of the primary processing of sensor signals can currently in most cases be achieved by various software implementations of the Kalman filter method [18]. For subsequent data processing, methods using a PID controller, optimal control methods with the fulfillment of the Lyapunov stability condition, nonlinear control methods, etc. [19, 20] can be used. The fuzzy logic method, reviewed in this chapter, is one of the most satisfying to the adequacy of modeling the movement of an automobile wheel in the braking mode in real time.

As an example of an mathematical model of a feed-back brake system presents a block diagram of the algorithm for the operation of a model of a brake system with ABS, which includes an operator for solving a system of differential equations (see Fig. 1) [16].

The adequacy of the solutions of the mathematical model to the real process is determined, firstly, by the correspondence of the equation of the real process to the physics of the phenomenon, secondly, by the accuracy of the mathematical description of the process modules, and, thirdly, by the correspondence of the results

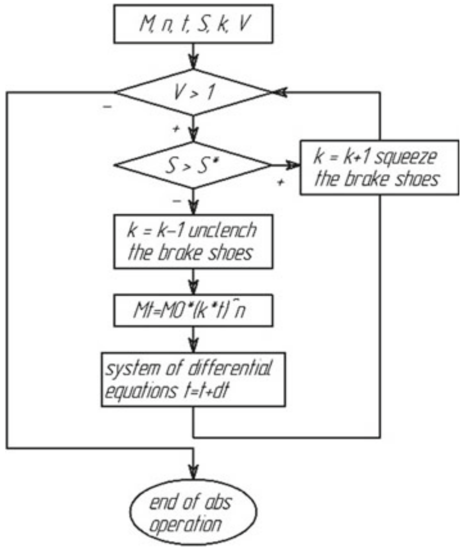



**Fig. 1** Block diagram of the algorithm of the brake system model with ABS

of computational and full-scale experiments with the same initial data, within made assumptions and errors of calculations and measurements [16].

When modeling a controlled process, one more thing is added to the requirements of uniqueness, stability and adequacy—the actual accumulated time for solving the model equations must be ahead of the time of the process being modeled. Moreover, the lead time should be no less than the duration of time for the adoption and execution of the decision by the corresponding control devices.

## 3 Mathematical Model of Car Wheel Braking

Models which involve the solution of differential equations require amount computer time to calculate the simulated process. For example, the solution of the system of nonlinear differential equations describing the motion of a wheel in the emergency braking mode given in [21] is Lyapunov stable. System of differential equations describing the process of braking on automobile wheel looks like:

$$\begin{cases} M = f(t), & (1) \\ \dot{\omega} = \frac{1}{j} \cdot [-M + \mu \cdot m \cdot g \cdot r], & (2) \\ \omega = \omega_0 + \int \dot{\omega} dt, & (3) \\ \dot{v} = -\mu \cdot g, & (4) \\ v = v_0 + \int \dot{v} \cdot dt, & (5) \\ \lambda = 1 - \frac{\omega \cdot r}{v}, & (6) \\ \mu = \frac{f_0 \cdot \lambda}{a \cdot \lambda^2 + b \cdot \lambda + c}. & (7) \end{cases}$$

where is

| | |
|---|---|
| $t$ | current time value; |
| $M$ | braking torque applied to the wheel; |
| $\omega_0, v_0$ | initial values of angular and linear speed; |
| $\dot{\omega}, \omega$ | angular acceleration and rotational speed; |
| $\dot{v}, v$ | linear acceleration and wheel speed; |
| $j, m$ | wheel moment of inertia and mass per wheel; |
| $g$ | gravity factor; |
| $r$ | dynamic radius of the wheel; |
| $\mu$ | coefficient of the longitudinal traction of a wheel with a road surface; |
| $\lambda$ | coefficient of the longitudinal slipping; |
| $a, b, c$ | empirical coefficients (descriptors of the road surface); |
| $f_0$ | tire-to-surface friction coefficient on full-lock skidding. |

This is following next physical phenomena:

(1) the dependence of the braking torque on time;
(2) change in the angular acceleration of the wheel;
(3) change in the rotational speed of the wheel;
(4) change in the deceleration of the translational motion of the mass concentrated in the center of the wheel and creating a normal load in the contact patch;
(5) change in the linear speed of the wheel;
(6) the ratio between the translational speed and the circumferential speed of the tread in the center of the contact spot is the coefficient of relative slippage;
(7) dependence of the tire-to-surface friction coefficient in the longitudinal direction on the coefficient of relative slippage and road surface.

However, in this case, in the solution of this system of equations, instability may arise, which is a consequence of the numerical solution methods used, in particular, with the choice of the integration step. With a decrease in the integration step, the accuracy of calculations increases, but at the same time, the calculation time increases. Existing numerical methods (Euler, Runge–Kutta, Dorman-Prince) assume the expansion of the desired function in a Taylor series, or (Adams) use Newton's second interpolation formula [22, 23]. In practice, the use of numerical methods with a given degree of calculation accuracy, for example, Runge–Kutta of the fourth order, only partially solves the problem of lack of computing power [23]. The problem of increasing the counting rate while maintaining a given calculation accuracy can be solved by using an ECU element of higher performance, for example, general-purpose graphic processors [24], which, accordingly, will lead to an increase in the price of the ECU. When developing the ABS model, one should take into account the currently economically unacceptable high cost of high-performance graphics processors for mass production of control units.

One of the possible alternative methods for representing the model of the wheel braking process in real time, considering the applied braking forces and allowing to exclude systems of differential equations as an intermediate stage of calculations, is the fuzzy logic method using linguistic variables. The use of linguistic variables used in fuzzy logic will require much less power of the computing element of the ECU [25].

## 4 Implementation of a One-Wheel ABS Model Under the Control of a Fuzzy Controller

In the calculation part, we will determine the braking parameters of an automobile wheel both using a numerical method and a deliberately small integration step (to ensure high accuracy of the results obtained) and using a fuzzy controller with the same initial data.

Next, it will be necessary to compare the results obtained and draw the appropriate conclusions.

Table 1 presents the initial data for the calculation. When solving the given system numerically, the improved Euler method and the calculation step equal to 0.005 s will be used.

As a version of the controller, an alternative to the PID controller, we use a fuzzy model of the controller. Let's leave the one-wheel model from the previous section unchanged, we will replace only the subsystem with the regulator.

To create a fuzzy controller, we use the Fuzzy library logic Toolbox of the MATLAB package, for work we will select the Fuzzy block logic controller (see Fig. 2).

To simplify the model, we use a model with one input and three output values determined by the range of the input value.

**Table 1** The initial data and the constant values adopted for the solution of the system

| Name | Designation | Value |
|---|---|---|
| Linear speed value at the initial moment of time | $v_0$ | 60 km/h |
| Linear and angular acceleration at the initial moment of time | $\dot{\omega}_0, \dot{v}_0$ | 0 |
| Wheel moment of inertia | $j$ | $0.9\ \mathrm{m} \cdot \mathrm{s}^2$ |
| Mass per wheel | $m$ | 350 kg |
| Gravity factor | $g$ | $9.8\ \mathrm{m} \cdot \mathrm{s}^2$ |
| Dynamic radius of the wheel | $r$ | 0.3 m |
| Empirical coefficients characterizing the road surface (dry surface) | $a; b; c; f_0$ | 0.342; 0.612; 0.046; 0.7 |
| Braking torque applied to the wheel at the initial moment of time | $M_0$ | 0 |
| Angular speed value at the initial moment of time | $\omega_0 = \frac{v_0}{r}$ | |
| The law of braking torque variation | $1000 \cdot \mathrm{t}$ | |

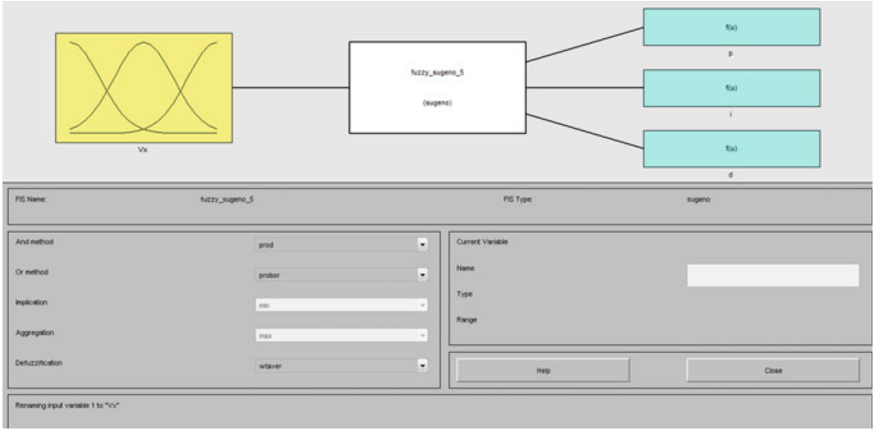



**Fig. 2** Model of a PID-like fuzzy controller

We will apply the speed parameter to the input, depending on the value of which at the output we will obtain the values Ki, Kp and Kd—the coefficients for the PID component of the designed controller.

Let's set the rules and get the visualization in Simulink (see Fig. 3).

Now compatible in Simulink fuzzy and PID components.

To improve the adequacy of the controller, we use two input parameters: the speed value for the fuzzy component and the deviation value for the PID component (see Fig. 4).

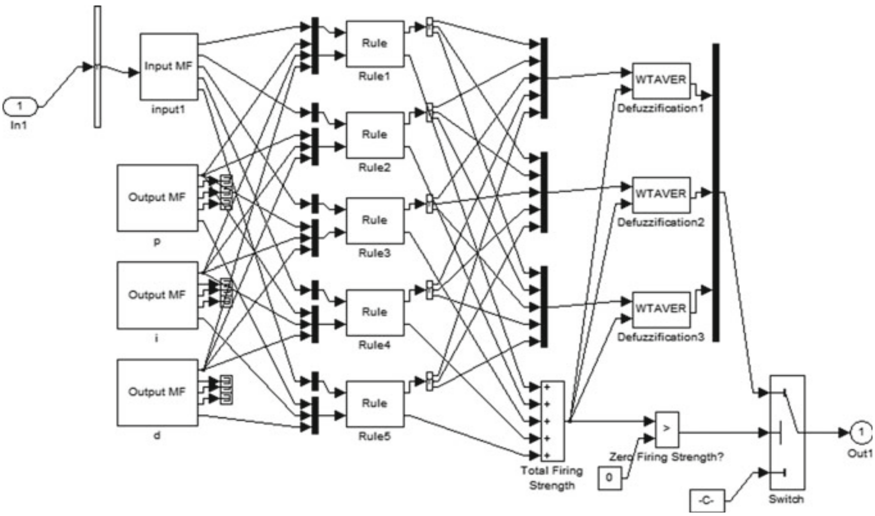



**Fig. 3** Visualization of the fuzzy controller in Simulink FIS wizard

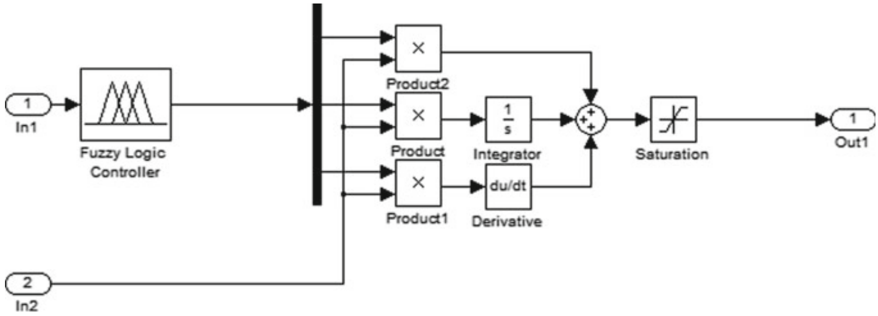



**Fig. 4** Fuzzy PID-like controller in MATLAB Simulink

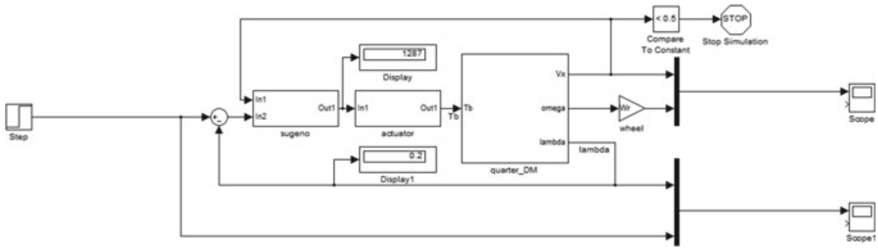
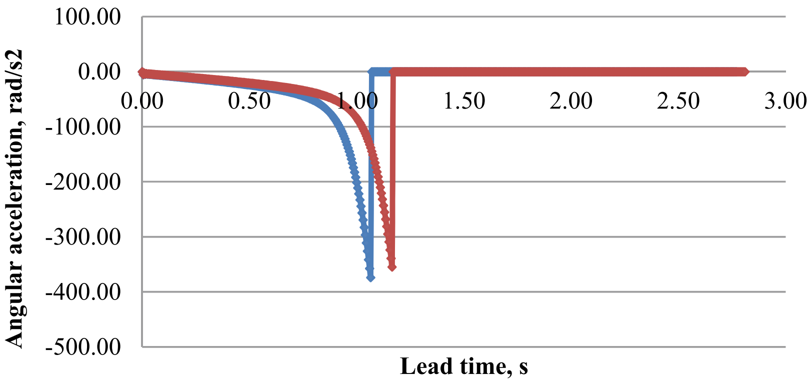



**Fig. 5** Implementation in MATLAB Simulink ABS models controlled by a fuzzy PID controller

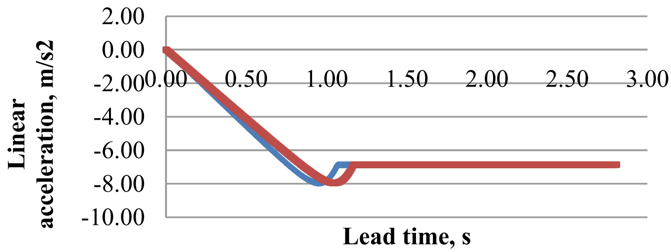
As a result, we obtain a PID-like fuzzy controller with coefficients P-, I-, D that change depending on the speed mode (see Fig. 5).

The calculation results are presented in Figs. 6 and 7. Figure 6 shows the dependence of the braking parameter on the process time, moreover, the blue line corresponds to the numerical solution, and the red line corresponds to the solution using a fuzzy controller. Figure 7 shows the differences between the obtained solutions, and the exact solution is taken as the solution obtained using the numerical method, the difference is expressed as a percentage.

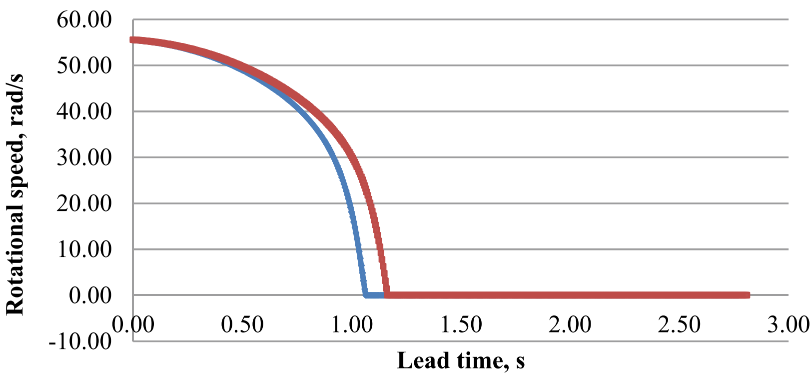
As can be seen from the graphs, over the entire range of the braking process, the fuzzy PID-like controller provides a satisfactory match to the exact values of the braking parameters.




a) angular acceleration of a car wheel

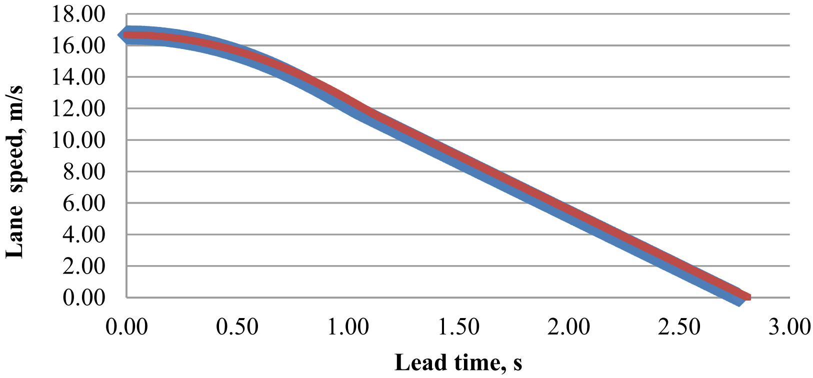



b) linear acceleration of a car wheel

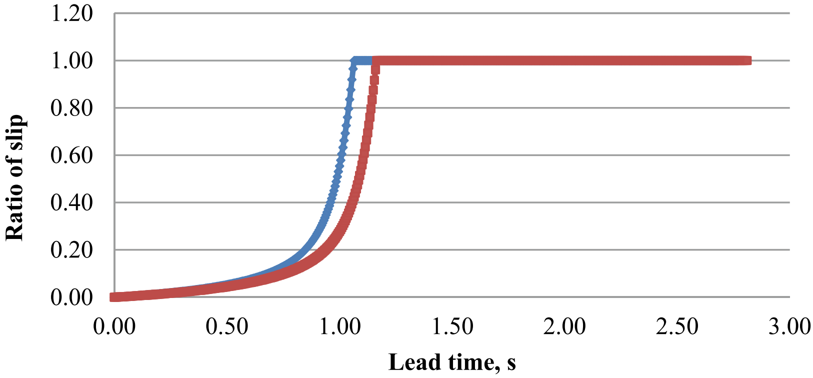



c) the angular velocity of the car wheel

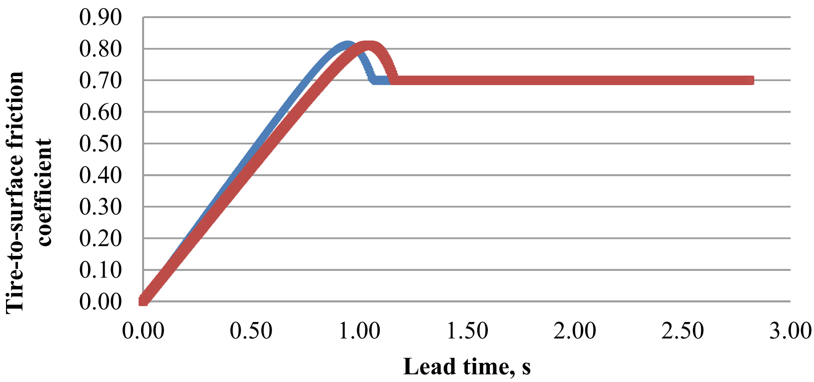
**Fig. 6** Braking parameters of a car wheel (blue line—numerical solution, red line—solution obtained using a fuzzy controller)




d) linear speed of a car wheel

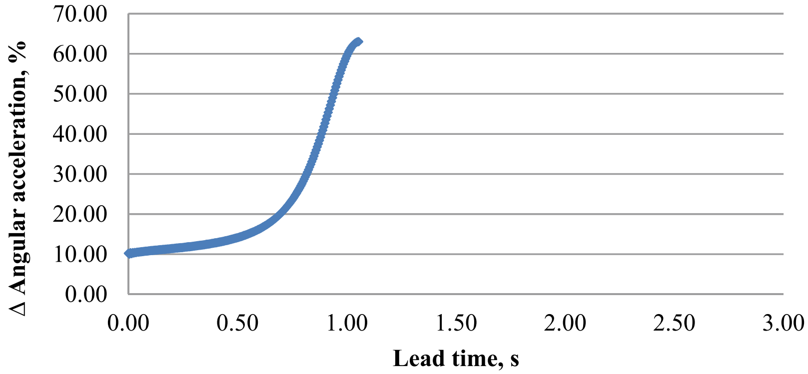



e) relative slippage coefficient

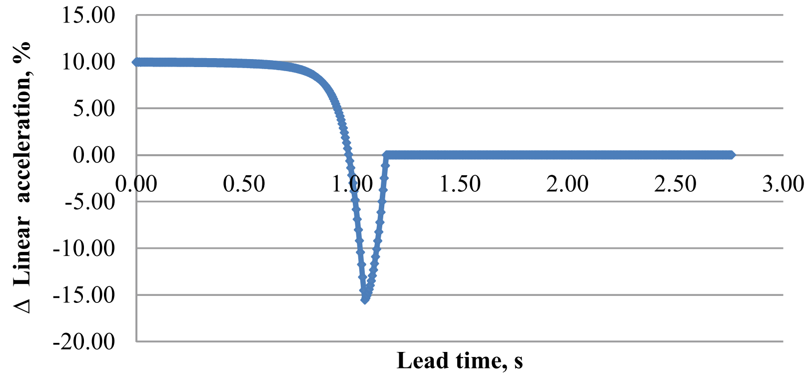



f) coefficient of adhesion of the wheel to the road

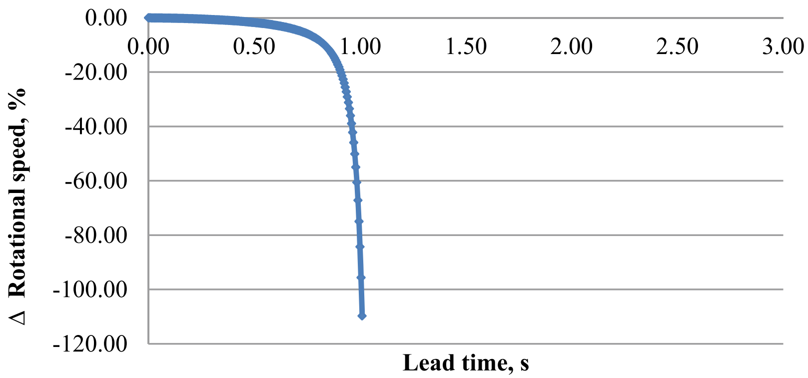
**Fig. 6** (continued)




a) Δ of the angular acceleration of the car wheel

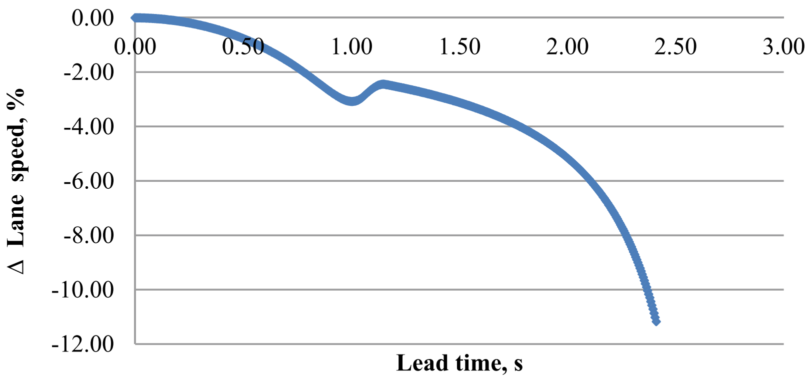



b) Δ linear acceleration of a car wheel

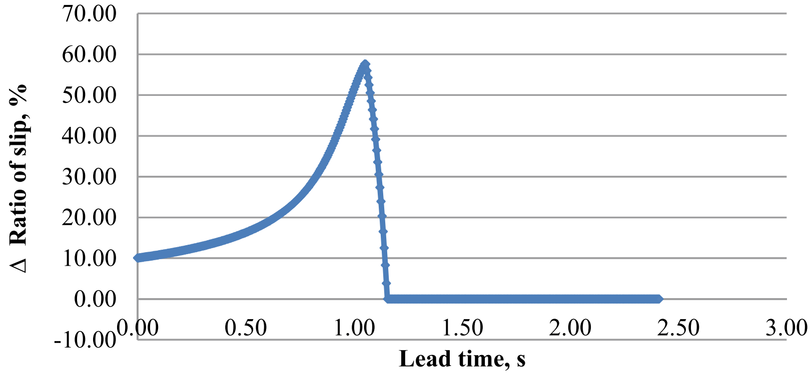



c) Δ of the angular velocity of the car wheel

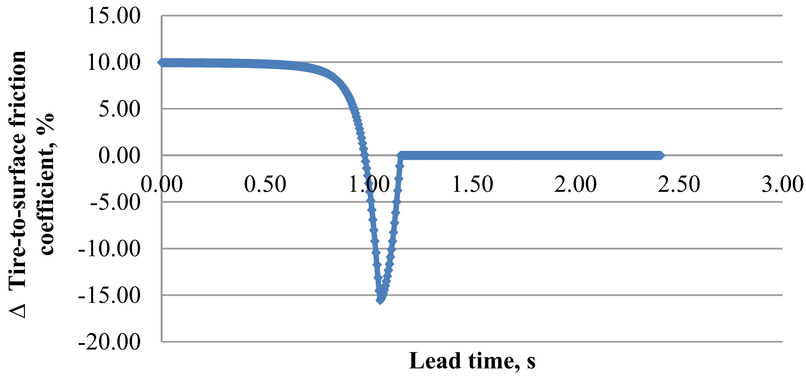
**Fig. 7** Differences between the braking parameters of an automobile wheel calculated in different ways




d) Δ of the linear speed of the car wheel




e) Δ skin coefficient of relative slip




f) Δ coefficient of adhesion of the wheel to the road

**Fig. 7** (continued)

## 5 Conclusion

With regard to the process of computer simulation of the movement of an automobile wheel in braking mode in real time, it is currently possible to recommend the following.

In some cases, the result of using a fuzzy controller is quite difficult to predict and it is not always possible to obtain the required accuracy of regulation. In this regard, it is necessary to make a preliminary analysis of the system and reasonably make a choice of the regulator. As the results of the study have shown, where the regulation is provided, it is preferable to use classical PID controllers, and in relation to the issue of modeling the movement of a car with ABS, the use of an adaptive fuzzy controller is reasonable and adequate;

When developing a fuzzy controller, the introduction of an additional input parameter in addition to the deviation value improves the quality of the model. In the proposed simplified model, the value of the vehicle speed is used. However, it is preferable to use the first or second derivative of either vehicle speed, wheel speed, or deviation change. The use of numerical values (through the values of derivatives, or according to the accelerometer data) and signs of accelerations in the fuzzy model makes it possible to consider the prediction of the system's behavior;

To improve the adequacy of the model, one should take into account the time delay, taking into account the action of the hydraulic brake system and data processing by the processor. It is necessary to predict in real time the value of the slip coefficient at the next point in time and use it to control the error. To do this, in the model, at the output of the slip coefficient signal, the subsystem should be switched on, taking into account the time delay.

## References


1. Society 5.0, Digital Transformation and Disasters. Past, Present and Future. https://doi.org/10.1007/978-981-19-5646-1
2. Tsugawa, S.: Significance of Automated Driving in Japan, pp. 793–818. Smart Cities, Foundations, Principles, and Applications (2017)
3. Hashimoto, N., Thompson, S., Kato, S., Boyali, A., Tsugawa, S.: Necessity of automated vehicle control customization: experimental results during lane changing. Trans. Res. Recordthis Link Disabled **2672**(22), 1–9 (2018)
4. Hashimoto, N., Wu, Y., Sato, T.: Human Factor Issues in Remote Operator of Automated Driving System for Services—One Operator to N Units of Automated Vehicles in Automated Vehicle Services. Communications in Computer and Information Science this link is disabled, 1655 CCIS, pp. 368–374 (2022)
5. Gao, T., Han, X., Tomita, K., Kamimura, A.: Nonlinear model-assisted control for autonomous parachute precision landing recovery system. In: 2022 61st Annual Conference of the Society of Instrument and Control Engineers of Japan, SICE 2022, pp. 1158–1164 (2022)
6. Galkin, A., Sysoev, A.: Controlling traffic flows in intelligent transportation system. In: Kravets, A.G., Bolshakov, A.A., Shcherbakov, M. (eds.) Society 5.0: Cyberspace for Advanced Human-Centered Society. Studies in Systems, Decision and Control, vol. 333. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-63563-3_8

7. Kravets, A.G., Skorobogatchenko, D.A., Salnikova, N.A., Orudjev, N.Y., Poplavskaya, O.V.: The traffic safety management system in urban conditions based on the C4.5 algorithm. In: 2018 Moscow Workshop on Electronic and Networking Technologies, MWENT 2018—Proceedings, art. no. 8337254, pp. 1–7 (March 2018)
8. Abe, M.: Vehicle Handling Dynamics. Theory and Application, 2nd edn, p. 322. Butterworth Heinemann, Elsevier Ltd., Oxford, UK (2015)
9. Balkwill, J.: Performance Vehicle Dynamics. Engineering and Applications. Department of Mechanical Engineering and Mathematical Sciences, p. 360. Oxford Brookes University, Elsevier Inc. (2018)
10. Arnold, M., Burgermeister, B., Führer, C., Hippmann, G., Rill, G.: Numerical methods in vehicle system dynamics: state of the art and current developments. Veh. Syst. Dyn. **49**(7), 1159–1207 (2011)
11. Bhaumik, S., Date, P.: Nonlinear Estimation: Methods and Applications with Deterministic Sample Points (2019). https://doi.org/10.1201/9781351012355
12. Chung, E., Dumont, A.-G.: Transport Simulation: Beyond Traditional Approaches (2019). https://doi.org/10.1201/9780429093258
13. Fedin, A., Kalinin, Y., Marchuk, E.: ANN in car antilock braking systems modeling. In: 2019 3rd School on Dynamics of Complex Networks and their Application in Intellectual Robotics, DCNAIR 2019, pp. 49, 8875513 (2019)
14. Fedin, A.P., Kalinin, Y.V., Marchuk, E.A.: Antilock braking system fuzzy controller optimization with a genetic algorithm in a form of cellular automaton. In: Conference Proceedings—4th Scientific School on Dynamics of Complex Networks and their Application in Intellectual Robotics, DCNAIR 2020, pp. 78–81, 9216912 (2020)
15. Tujrin, S., Boyko, G., Revin, A., Fedotov, V.: Research to longevity brake lines on the exploitations. Transp. Prob. **10**(2), 75–81 (2015)
16. Revin, A.A., Dygalo, V.G.: Complex simulation in the projection cycle of automobiles and their systems. Avtomobil'naya Prom-st. **11**, 29–30 (2002)
17. Dygalo, V., Lyaschenko, M., Boyko, G., Dygalo, L.: Application of virtual physical modeling technology for the development of elements of the autonomous (unmanned) vehicles' systems. IOP Conf. Ser.: Mater. Sci. Eng. **315**(1), 012007 (2018)
18. Guiggiani, M.: The Science of Vehicle Dynamics: Handling, Braking, and Ride of Road and Race Cars (2018). https://doi.org/10.1007/978-3-319-73220-6
19. Piegat, A.: Fuzzy Modeling and Control. Monograph, XIV, p. 728 (2001). https://doi.org/10.1007/978-3-7908-1824-6
20. Mirzaei, M., Mirzaeinejad, H.: Fuzzy scheduled optimal control of integrated vehicle braking and steering systems. IEEE/ASME Trans. Mechatron. **22**(5), 8025611, 2369–2379 (2017)
21. Balakina, E.V., Zotov, N.M., Fedin, A.P.: Modeling of the motion of automobile elastic wheel in real-time for creation of wheeled vehicles motion control electronic systems. IOP Conf. Ser.: Mater. Sci. Eng. **315**(1), 012004 (2018)
22. Pawlus, W., Robbersmyr, K.G., Karimi, H.R.: Mathematical modeling and parameters estimation of a car crash using data-based regressive model approach. Appl. Math. Model. **35**(10), 5091–5107 (2011)
23. Pacejka, H.B., Besselink, I.Y.: Magic formula tyre model with transient properties. Suppl. Veh. Syst. Dyn. **27**, 234–249 (1997)
24. Saiteja, P., Jeyanthi, S.: Fuzzy logic simulation for brake-by-wire control system. Lecture Notes in Mechanical Engineering, PartF9, pp. 135–147 (2017)
25. Saiteja, P.: Fuzzy logic simulation for automatic speed control system. Lecture Notes in Mechanical Engineering, pp. 499–509 (2019)